

06061 Abstracts Collection
Theory of Evolutionary Algorithms
— Dagstuhl Seminar —

Dirk V. Arnold¹, Thomas Jansen², Jonathan E. Rowe³ and Michael D. Vose⁴

¹ Dalhousie Univ., CA

dirk@cs.dal.ca

² Univ. Dortmund, DE

thomas.jansen@udo.edu

³ Univ. of Birmingham, GB

j.e.rowe@cs.bham.ac.uk

⁴ Univ. of Tennessee, US

vose@cs.utk.edu

Abstract. From 05.02.06 to 10.02.06, the Dagstuhl Seminar 06061 “Theory of Evolutionary Algorithms” was held in the International Conference and Research Center (IBFI), Schloss Dagstuhl. During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations given during the seminar as well as abstracts of seminar results and ideas are put together in this paper. The first section describes the seminar topics and goals in general. Links to extended abstracts or full papers are provided, if available.

Keywords. Evolutionary algorithms, evolutionary computation, theory

06061 Executive Summary – Theory of Evolutionary Algorithms

The 2006 Dagstuhl Seminar “Theory of Evolutionary Algorithms” carried forward a series of Dagstuhl seminars that started in 2000 and has become an established event in the community. In the week from 05.02.2006 to 10.02.2006, 56 researchers from 12 countries discussed their recent work and recent trends in evolutionary computation.

Joint work of: Arnold, Dirk V.; Jansen, Thomas; Rowe, Jonathan E.; Vose, Michael

Extended Abstract: <http://drops.dagstuhl.de/opus/volltexte/2006/599>

Convergence on the Sphere Model – Progress Rate and Beyond

Anne Auger (ETH Zürich, CH)

Progress-rate theory has been introduced with Evolution Strategies and guided its development. In this talk we show some limits of the predictions based on this theory.

We explain for the sphere function how surprisingly, negative progress can lead to convergence towards the optimum. We explain why positive progress rates give convergence in mean, negative progress rates divergence in mean and show that almost sure convergence can take place despite divergence in mean.

Hence step-sizes associated with negative progress can actually lead to almost sure convergence. Based on these results we provide an alternative progress rate definition. We present Monte Carlo simulations to investigate the discrepancy between both progress rates and therefore both types of convergence. This discrepancy vanishes when dimension increases. The observation is supported by an asymptotic estimation of the new progress rate definition.

Keywords: Evolution strategies, convergence analysis, progress rate

Joint work of: Auger, Anne; Hansen, Nikolaus

Evaluation Structures in Coevolution

Anthony Bucci (Brandeis University - Waltham, USA)

One characteristic of the problems typically approached with coevolutionary algorithms is their lack of an explicit measure of value. Unlike typical optimization or evolutionary algorithms, which rely on the existence of an objective function, coevolutionary algorithms must develop metrics through time while simultaneously discovering entities which optimize them. With a few notable exceptions, generally the means for developing such metrics has been a black art. My focus here is therefore on the problem of evaluation.

With this problem in mind, I pose the question: what should a coevolutionary algorithm be doing? My simple answer is that an algorithm should be discovering and maintaining as many innovations as possible. In an attempt to approach the question precisely, in this talk I will develop a theoretical tool called an evaluation structure. Such is a structured mathematical object built from a population and representing salient aspects, or measurements, of entities' capabilities. Examples include the Pareto covering order and Ficici's notion of measurement table. I use this tool to distinguish two types of algorithm transitions: natural ones, which preserve but elaborate existing structure; and unnatural ones, which upset structure. I prove that, under certain conditions, unnatural transitions are necessary in the sense that an algorithm which does not make them will never reach a solution. Under these conditions I dub such transitions "innovative," as they are more or less radical changes in how we perceive the population to be structured which are required to reach solution. I conclude with several empirically testable hypotheses suggested by the theory.

Keywords: Coevolution, coevolutionary algorithms, Pareto coevolution, solution concepts, theory

See also:

<http://www.cs.brandeis.edu/~abucci/>

Using the Underlying Objectives for Evaluation in Coevolution

Edwin D. de Jong (Utrecht University, NL)

In coevolution, evaluation is based on interactions with (co-)evolving individuals. Since the set of individuals used in evaluation evolves over time, the evaluation function effectively develops as part of the evolutionary process. This provides a potential for open-ended evolution.

To guarantee that adaptive evaluation results in overall progress, coevolution must be set up such that the resulting dynamics is sufficiently stable. To this end, Pareto-coevolution views the opponents in interactions ('tests') as objectives. The resulting number of objectives can be very large. We therefore consider whether the number of objectives can be reduced by combining multiple tests onto a single objective.

We apply a dimension extraction algorithm to the Game of Nim. We found that the number of objectives is reduced from 36 to 4. Thus, extraction of the underlying objectives can permit efficient evaluation without loss of accuracy. The Dimension Extracting Coevolutionary Algorithm (DECA) is applied to test problems, and found to perform well on the compare-on-one problem. Application to the more realistic Tartarus problem is shown to be feasible, but yields little improvement. An open question is the relation between the number of underlying objectives (an intrinsic property of a problem) with other measures; and which features limit the performance of current coevolution methods on problems of practical interest.

Keywords: Coevolution, underlying objectives, dimension extraction, Game of Nim

An Analysis of Co-evolutionary Dynamics

Kenneth A. De Jong (George Mason University - Fairfax, USA)

A new technique for analyzing co-evolutionary dynamics is described that is based on the notion of "best response" curves. The usefulness of this technique is illustrated for both cooperative and competitive co-evolutionary settings.

Keywords: Co-evolution, dynamical systems, basins of attraction, best response curves

Joint work of: De Jong, Kenneth A.; Popovici, Elena

Functional Decomposition of NSGA-II and Various Problem-Solving Strategies

Kalyanmoy Deb (Indian Inst. of Technology - Kanpur, IND)

NSGA-II is a commonly-applied tool for multi-objective optimization. In this presentation, the NSGA-II procedure will be functionally decomposed into three tasks.

Thereafter, various modifications to these three tasks will be made to demonstrate that the NSGA-II framework can generate a number of efficient and interesting optimization strategies for solving various problem-solving tasks. This decomposition should allow researchers and applicationists to try various other modifications systematically to make NSGA-II suitable for other kinds of problem-solving tasks. Finally, NSGA-II framework will be expanded to develop an ‘omni-optimizer’, which is capable of solving four different types of optimization problems involving single and multiple objectives and single and multiple optima.

Keywords: Multi-objective optimization, multi-modal optimization, NSGA-II

See also:

<http://www.iitk.ac.in/kangal/pub.htm>

A Rigorous Runtime-Analysis of the Compact Genetic Algorithm for Linear Functions

Stefan Droste (Universität Dortmund, D)

Estimation of distribution algorithms (EDAs) solve an optimization problem heuristically by finding a probability distribution focused around its optima. Starting with the uniform distribution, points are sampled with respect to this distribution and the distribution is changed according to the function values of the sampled points. Although there are many successful experiments suggesting the usefulness of EDAs, there are only few rigorous theoretical results apart from convergence results without time bounds. Here we present first rigorous runtime analyses of a simple EDA, the compact genetic algorithm, for linear pseudo-boolean functions on n variables. We prove a general lower bound for all functions and a general upper bound for all linear functions. Simple test functions show that not all linear functions are optimized in the same runtime by the compact genetic algorithm.

On Complexity of Optimized Crossover for Binary Representations

Anton Ereemeev (Sobolev Institute of Mathematics - Omsk, RUS)

In this talk, the computational complexity of producing the best possible offspring in a crossover, complying the principle of respect (see e.g. N. Radcliffe, 1991) is considered. The focus is on the gene transmitting crossover operators, where all alleles present in a child are transmitted from its parents. These operators are studied on the class of Boolean linear programming problems, where the Boolean vector of variables is used as the solution representation.

The well-known optimized crossover operator for the maximum independent set problem (C.C. Aggarwal, J.B. Orlin, R.P. Tai, 1994) is taken as a starting point. By means of efficient reductions of the optimized gene transmitting crossover problems (OGTC) we show the polynomial solvability of the OGTC for maximum weight set

packing problem, minimum weight set partition problem and simple plant location problem.

The OGTC for linear Boolean programming problem with logarithmically upper-bounded number of non-zero coefficients per constraint is shown to be efficiently reducible to the maximum weight independent set problem on 2-colorable hypergraph with 2-coloring given as an input. Both of these problems are NP-hard.

Keywords: Genetic algorithms, optimized crossover, complexity

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2006/593>

A Mathematical Modelling Technique for the Analysis of the Dynamics of Simple Continuous EDA

Marcus Gallagher (The University of Queensland, AU)

We describe a mathematical model for the infinite-population dynamics of a simple continuous EDA: UMDAc. Using this model, it is possible to numerically generate the dynamics of the algorithm on a fitness function of known form. The technique is compared with existing analysis and illustrated on a number of simple test problems. The model is also used to examine the effect of adding an amplification constant to the variance parameter of the UMDAc model.

Keywords: Estimation of distribution algorithms

Joint work of: Gallagher, Marcus; Yuan, Bo

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2006/594>

On The Effect of Populations in Evolutionary Multi-objective Optimization

Oliver Giel (Universität Dortmund, D)

Multi-objective evolutionary algorithms (MOEAs) have become increasingly popular as multi-objective problem solving techniques.

An important open problem is to understand the role of populations in MOEAs. We present a simple bi-objective problem which emphasizes when populations are needed. Rigorous runtime analysis point out an exponential runtime gap between a population-based algorithm (SEMO) and several single individual-based algorithms on this problem. This means that among the algorithms considered, only the population-based MOEA is successful and all other algorithms fail.

Step Length on Linear Fitness Functions—Self-Adaptation and Beyond

Nikolaus Hansen (ETH Zürich, CH)

We investigate the evolution strategy (ES) with σ -self-adaptation on a linear fitness function. We find

- For the $(1, 2)$ -ES the selection does not change the σ -distribution. Consequently $\log \sigma$ conducts an unbiased random walk.
- Given a symmetric population distribution in the $(\mu, 2\mu)$ -ES the selection does not bias σ , that is $\log \sigma$ is constant in mean.
- In general, given a symmetrical population distribution, the step-length distribution is symmetric such that the i -th best and i -th worst individual have the same distribution.

We state two postulates for any search algorithm in continuous domain on the linear fitness function.

1. The step length logarithm (in gradient direction) should increase linearly in time.
2. The step length in gradient direction should increase faster than the step length on a random function (where the function value is independent of the search point).

The second postulate corresponds to a desired selection effect in the algorithm.

A Generic Property of the Infinite Population Genetic Algorithm

Christina Hayes (Montana State University - Bozeman, USA)

We study an infinite population model for the genetic algorithm, where the iteration of the algorithm corresponds to an iteration of a map G . The map G is a composition of a selection operator and a mixing operator, where the latter models effects of both mutation and crossover. We examine the hyperbolicity of fixed points of this model. We show that for a typical mixing operator all the fixed points are hyperbolic.

Joint work of: Hayes, Christina; Gedeon, Tomas

The First-Hitting Times of Evolutionary Algorithms for the One-Max Problem

Jun He (University of Birmingham, GB)

Two important problems in the theoretical study of evolutionary algorithms are: (i) to investigate the relationship between time complexity and population size; and (ii) to compare the time complexity of an algorithm when using crossover or not.

This study could help us gain insight into understanding the roles of population and crossover, and designing efficient algorithms.

However, even for some basic fitness functions, their time complexity is still unknown for population-based evolutionary algorithms with crossover. It is necessary to make a systematic and comprehensive analysis of evolutionary algorithms on different fitness functions. This paper aims at estimating the first hitting time of a family of EAs using uniform crossover, bitwise mutation and elitism selection for the One-Max problem. The exact lower bounds have been derived for the algorithms without and with uniform crossover; and these lower bounds can be achieved by problem-specific algorithms. The result demonstrates that the crossover can shorten the first hitting time in this case, and selection pressure and population diversity also play a crucial role.

Keywords: Time complexity, population, bitwise mutation, uniform crossover

Joint work of: He, Jun; Witt, Carsten; Yao, Xin

Computational Efficient Covariance Matrix Update and the Multi-objective Variable Metric Evolution Strategy

Christian Igel (Ruhr-Universität Bochum, D)

The covariance matrix adaptation (CMA) with rank-one update is combined with elitist $(1 + 1)$ -selection. An iterative Cholesky update for the covariance matrix is developed replacing the computational demanding and numerically involved decomposition of the covariance matrix. The Cholesky update can replace the decomposition only for the update without evolution path and reduces the computational effort from $O(n^3)$ to $O(n^2)$. The resulting $(1 + 1)$ -Cholesky-CMA-ES is an elegant algorithm and the perhaps simplest evolution strategy (ES) with covariance matrix and step size adaptation. Based on the $(1 + 1)$ -CMA-ES, a new multi-objective evolution strategy, the MO-CMA-ES, is developed. In the MO-CMA-ES a population of individuals that adapt their search strategy as in the elitist CMA-ES is maintained. These are subject to multi-objective selection. The selection is based on non-dominated sorting using the contributing hypervolume as second sorting criterion. Both the $(1 + 1)$ -CMA-ES and the MO-CMA-ES inherit important invariance properties, in particular invariance under rotation of the search space, from the original CMA-ES.

Keywords: Evolution strategies, covariance matrix adaptation, multi-objective optimization

Joint work of: Igel, Christian; Hansen, Nikolaus; Sutton, Thorsten

In Between Progress Rate and Stochastic Convergence

Jens Jägersküpper (Universität Dortmund, D)

This talk is about analyzing evolutionary algorithms for the continuous domain of high dimensional Euclidean space in a way known as *probabilistic runtime analysis*.

It is shortly discussed why this can be thought of as a way in between the progress-rate approach and proofs of stochastic convergence.

As an example of such an analysis, a lower-bound theorem is presented and the ideas of the somewhat surprisingly simple proof are given. It reads: A $(1+\lambda)$ ES using isotropic mutations and an *1/5-rule-like* mutation-strength adaptation needs $\Omega(b \cdot n / \ln(1 + \lambda))$ steps to reduce the approximation error (w.r.t. a fixed point in the search space, for instance an optimum) to an 2^{-b} -fraction (of the initial one) with a probability of $1 - \exp(-\Omega(n))$ – whatever function is to be optimized.

Keywords: Continuous optimization, probabilistic runtime analysis, evolution strategies, adaptation

On Turing complete T7 and MISC F-4 program fitness landscapes

William Langdon (University of Essex, GB)

We use the minimal instruction set F-4 computer to define a minimal Turing complete T7 computer suitable for genetic programming (GP) and amenable to theoretical analysis.

Experimental runs and mathematical analysis of the T7, show the fraction of halting programs is drops to zero as bigger programs are run.

Keywords: Genetic programming

Joint work of: Langdon, William B; Poli, Riccardo

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2006/595>

See also:

<http://www.cs.essex.ac.uk/technical-reports/2005/csm-445.pdf>

Controlled Regularity Fitness Landscapes and Local Holder Exponents

Evelyne Lutton (INRIA Rocquencourt, F)

Intuition as well as theoretical analysis suggest that the performance of an EA strongly depends on the irregularity of the fitness function. Several irregularity measures have been derived for discrete search spaces, in order to numerically characterize this type of difficulty source for EA. These characterizations are mainly based on Hölder exponents. Previous studies used a global characterization of fitness regularity (namely the global Hölder exponent), with experimental validations being conducted on test functions with uniform irregularity. We will present an extension of this analysis in two ways: results have been now established for continuous search space, and local instead of global irregularity has been considered. In addition, we present a way to modify the genetic topology to accommodate for variable regularity: the mutation radius, which controls

the size of the neighbourhood of a point, should vary according to the local irregularity of the fitness function. These results are explained through a simple theoretical analysis which gives a relation between the local Hölder exponent and the optimal mutation radius. It poses several questions with respect to on-line measurements and usage of regularity in EAs.

Keywords: Hölder exponent, irregular fitness landscape, adaptive mutation radius

Semantic Building Blocks in Boolean GP

Nicholas Freitag McPhee (University of Minnesota - Morris, USA)

A persistent question is that of exactly how relevant components are brought together in “standard” GP using subtree crossover, especially since crossover can place sub-trees in new contexts that can be very different from those they came from. Here we present a means of fully and exactly characterizing both the semantics of subtrees and contexts for boolean valued GP trees, which can be used to track the proportions of both the subtrees and contexts that crossover is using to construct new individuals. Using this tool we find that the number of contexts that are fully determined (i.e., locations where crossover will have no effect on the semantics of the resulting tree) grows very quickly, especially when little or no improvements are being made on the fitness.

Keywords: Genetic programming, semantics, crossover, bloat, building blocks

Joint work of: McPhee, Nicholas Freitag; Hutchison, Tyler; Ohs, Brian

How fast does the stationary distribution of the markov chain modelling EAs on the homogeneous populations for small rate?

Boris S. Mitavskiy (University of Birmingham, GB)

The state space of the Markov chain modelling an evolutionary algorithm is quite large especially if the population space and the search space are large. I shall introduce an appropriate notion of “coarse graining” for such Markov chains. Indeed, from the mathematical point of view, this can be called a quotient of a Markov chain by an equivalence relation over the state space. The newly obtained Markov chain has a significantly smaller state space and its stationary distribution is “coherent” with the initial large chain. Although the transition probabilities of the coarse-grained Markov chain are defined in terms of the stationary distribution of the original big chain, in some cases it is possible to deduce interesting information about the stationary distribution of the original chain in terms of the quotient chain. I will demonstrate how this method works. I shall also present some simple results and open questions.

Keywords: Markov chains, evolutionary algorithms, coarse graining, quotients of irreducible Markov chains, concentration on the uniform populations

Joint work of: Mitavskiy, Boris S.; Rowe, Jonathan E.

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2006/596>

The Factorized Distribution Algorithm and the Minimum Relative Entropy Principle

Heinz Mühlenbein (Fraunhofer Institut - St. Augustin, D)

We assume that the function to be optimized is additively decomposed (ADF). Then the interaction graph G_{ADF} can be used to compute exact or approximate factorizations. For many practical problems only approximate factorizations lead to efficient optimization algorithms. The relation between the approximation used by the FDA algorithm and the minimum relative entropy principle is discussed. A new algorithm is presented, derived from the Bethe-Kikuchi approach in statistical physics. It minimizes the relative entropy to a Boltzmann distribution with fixed β . We shortly compare different factorizations and algorithms within the FDA software. We use 2-d Ising spin glass problems and Kaufman's $n - k$ function as examples.

Keywords: Junction tree, minimum relative entropy, maximum likelihood, Bethe-Kikuchi approximation

Joint work of: Mühlenbein, Heinz; Höns, Robin

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2006/597>

See also:

<http://www.ais.fraunhofer.de/muehlen/pegasus/publications.html#MueHoe95a>

Emergent Behaviour, Population Based Search and Low-pass Filtering

Riccardo Poli (University of Essex, GB)

In this talk I presented a model of emergent coordinated behaviour for a population of interacting entities (e.g. animals or individuals in an evolutionary algorithm). The model is a modified spring mass model where the masses are active, in the sense that they can perceive the environment and generate external forces. As a result of the interactions (assuming the springs are relatively stiff) the population behaves like a single animal moving under the effect of a force which is the vector sum of the external forces generated by each entity. Under the assumption that the force generated by each individual is proportional to the gradient of a resource distribution $f(x)$, the resultant force controlling the single emergent animal is proportional to the gradient of a modified food distribution. This is the result of applying a filtering kernel to $f(x)$. The kernel is typically a low-pass filter.

This model can be applied to genetic algorithms and other population-based search algorithms. For example, we have found kernels (via genetic programming) using which the single animal model can track the motion of the centre of mass of genetic algorithms and particle swarm optimisers accurately for many generations. Also, I have shown a simple proof that on any 2-bit problem, the effect of crossover is that of reducing the amplitude of the derivatives (slopes) of the fitness function.

Finally, taking inspiration from this, I have proposed a fitness function, OneMix, which has an area with high frequency variations containing the global optimum which a genetic algorithm with high crossover cannot “see” and, so, is deceived and attracted to a local optimum, while with low crossover this does not happen.

Keywords: Emergent behaviour, genetic algorithms, low-pass filters

Modelling Hard Optimisation Problems

Adam Prügel-Bennett (University of Southampton, GB)

An approach to modelling instances of hard optimisation problems is presented by amalgamating together neighbouring configurations to form single states. This uses methods developed for constructing Barrier trees of combinatorial optimisation problems. The resulting models have the same structure of local minima and barriers as the problem instance they are based on. The model problems are sufficiently small to allow many algorithms to be studied analytically. This makes it possible to find optimal parameter settings for the algorithms. By studying different algorithms it becomes possible to gain a deep understanding of how heuristic algorithms perform on realistic hard problems.

Keywords: Optimisation problems, barrier trees, heuristic search, MAXSAT, simulated annealing, optimal annealing schedules

The Never-Hitting Time in EDAs

Jonathan L. Shapiro (Manchester University, GB)

EDAs are evolutionary algorithms which learn probability models from selected data, and use these models to generate the new population.

It has been shown in previous work that simple EDAs such as PBIL and UMDA may converge to a state in which the optimum is never found (for example Gonzales et. al. 2001, Shapiro, 2003). In this work, the probability of never finding the optimum (the “never-hitting probability”) is considered. It has been shown previously, both in empirical studies and theoretically, that this probability depends on the control parameters and the problem size in a very sensitive and problem dependent way (Shapiro, 2003). For example, in UMDA, the population size must exceed a threshold for the optimum to be found, and this threshold scales with the number of variables exponentially in some problems, and polynomially in others.

It is argued that this behaviour can be understood (for binary variables) by considering the dynamics of the trace of the expected covariance matrix, which is a measure of the diversity of the sampled population. It is shown that for any EDA which uses maximum likelihood (ML) to set the model parameters (e.g. MIMIC, BOA, UMDA, PBIL) on a flat fitness landscape, this statistic decays exponentially in time, with a time constant approximately equal to the population size. Thus, on any landscape with a

flat subspace, diversity will be lost in the population at a constant rate. If this rate of diversity loss exceeds the rate of movement towards the optimum, the algorithm will fail to find the optimum. This gives a qualitative explanation of the sensitive dependence on parameters of the never-hitting probability, and suggests the prediction that this sensitivity will hold for all EDAs which use maximum likelihood, such as BOA. This result can be used to derive a rigorous lower bound for the minimum population size required to solve the “needle-in-a-haystack” problem, using any EDA which uses ML to set model parameters.

Keywords: Estimation of distribution algorithms, rigorous population-size bounds, population diversity

Topological Concepts for Fitness Landscapes — or: General Abstract Nonsense

Peter F. Stadler (Universität Leipzig, D)

What do GA, GP, EA, ES, MOO, etc have in common? In all cases we are given a set of X of ‘configurations’ a set Z of ‘values’ that are at least partially ordered, a function $f: X \rightarrow Z$ assigning a value to each configuration and collection M of ‘operators’ that map subsets of X to other subsets of X . The essence of the action of M can be captured by a generalized closure functions $c: P(X) \rightarrow P(X)$ defined as $c(A) = \{x | x \text{ can be reached from } A \text{ by a single action of an operator in } M\}$.

The set-valued set function c defined a generalization of a topology on X that defined in a natural way notions of neighborhood, separation, connectedness, sub- and product spaces and so on. This in turn implies meaningful notions e.g. of local minima, barriers and saddlepoints that in particular also apply to recombination.

Keywords: Fitness landscapes, generalized topology, search operators, closure function, barrier tree

Generalised Recombination

Christopher Stephens (Universidad Nacional Autonoma - Mexico, MEX)

An exact microscopic model for the dynamics of a genetic algorithm with generalised recombination is presented. Generalised recombination is a new form of exchange of genetic material from parents to offspring that generalises and subsumes standard operators, such as homologous crossover, inversion and duplication, and in which a particular gene in the offspring may originate from *any* parental gene. It is shown that the dynamics naturally coarse grains, the appropriate effective degrees of freedom being schemata that act as building blocks.

It is shown that the schema dynamics has the same functional form as that of strings and we derive a corresponding Exact Schema theorem.

To exhibit the qualitatively new phenomena that can occur in the presence of generalised recombination, and to understand the biases of the operator, we derive a complete, exact solution for a two-locus model without selection, showing how the dynamical behaviour is radically different to that of homologous crossover. Inversion is shown to potentially introduce oscillations in the dynamics, while gene duplication leads to an asymmetry between homogeneous and heterogeneous strings. All non-homologous operators lead to allele “diffusion” along the chromosome. We discuss how inferences from the two-locus results extend to the case of a recombinative GA with selection and more than two loci providing evidence from an integration of the exact dynamical equations for more than two loci.

Keywords: Recombination, building blocks, inversion, duplication, Geiringer

Joint work of: Stephens, Christopher; Poli, Riccardo

A Statistical Learning Approach for GP

Olivier Teytaud (Université Paris Sud, F)

In the framework of symbolic regression, we study the convergence of the program chosen by GP to a program that (i) match unseen examples also (ii) has a not-huge size.

We show that the consistency (satisfaction of (i)) is satisfied under mild classical assumptions, but that (ii) requires a modification of the fitness.

Keywords: Genetic programming, statistical learning theory

Joint work of: Amil, Merve; Bredèche, Nicolas; Gagné, Christian; Gelly, Sylvain; Schoenauer, Marc; Teytaud, Olivier

See also:

<http://www.lri.fr/teytaud/eabloat.pdf>

A New Quartet Tree Heuristic for Hierarchical Clustering

Paul M.B. Vitanyi (CWI - Amsterdam, NL)

We present a new quartet heuristic for hierarchical clustering from a given distance matrix.

We determine a dendrogram (ternary tree) by a new quartet method and a fast heuristic to implement it. We do not assume that there is a true ternary tree that generated the distances and which we wish to recover as closely as possible. Our aim is to model the distance matrix as faithfully as possible by the dendrogram. Our algorithm is essentially randomized hill-climbing, using parallelized Genetic Programming, where undirected trees evolve in a random walk driven by a prescribed fitness function. Our method is capable of handling up to 60–80 objects in a matter of hours, while no existing quartet heuristic can directly compute a quartet tree of more than about 20–30 objects without running for years.

The method is implemented and available as public software at www.complearn.org. We present applications in many areas like music, literature, bird-flu (H5N1) virus clustering, and automatic meaning discovery using Google.

Keywords: Genetic programming, hierarchical clustering, quartet tree method

Joint work of: Cilibrasi, Rudi, Vitanyi, Paul M.B.

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2006/598>

Recombining Building Blocks in Simple Building-Block Functions and Organisms

Richard A. Watson (University of Southampton, GB)

Many attempts to illustrate the combination of building blocks have been unconvincing, and in general, there are many good reasons to believe that such notions are redundant. However, the motivation for interest in this possibility is a good one: if an ‘escape’ from optimisation at the relatively microscopic scale of bits were possible the combinatorial advantage would be significant. Some of the prior work on simple concatenated building-block functions assumed that desirable blocks should be found without utilising selection on individual bits; as if precluding the utility of selection on bits, as in fully-deceptive trap functions, maximised the advantage of selecting on blocks. Consequently most work continued with the assumption that the size of blocks, k , must be small, and the population size had to be proportional to 2^k . In fact, with these assumptions, the advantage of crossover is mostly due simply to preserving common alleles, and ‘macro-mutation’ at the loci where the parents differ is sufficient to find good blocks (as in the case of uniform crossover). However, in trap functions that are only partially-deceptive, selection at the bit scale can find good blocks (in polynomial time) in at least some individuals even when k is large, and a population size that is proportional to 2^k is not required. Given large k , the fact that selection on bits is at least sometimes misleading is sufficient to prevent a mutation hill-climber from finding optimal genotypes. And simply preserving the common alleles of the parents in a recombinative process without combining blocks from different individuals will also require time (at least) exponential in k to acquire a new desirable block. In contrast, a GA with two-point crossover (and merely a moderate population size), can follow gradients in bit-sequence space to find good blocks, and combine good blocks to find good genotypes. Thus, simply by using large partially-deceptive subfunctions, a GA with two-point crossover can be used to demonstrate that the combination of building blocks is required. As the original intuition suggested, it is the escape from the microscopic scale of the bits that provides the combinatorial advantage; but ignoring selection at the microscopic scale in previous models has been self-defeating. I suggest that in natural populations, genes correspond to very large building blocks containing thousands of nucleotides that are grouped both functionally and physically. The interaction of the adaptation occurring at the two different scales, point mutation on nucleotides and recombination of genes, could provide a significant component to the advantage of sex.

Simulated Annealing Beats Metropolis in Combinatorial Optimization

Ingo Wegener (Universität Dortmund, D)

The Metropolis algorithm is simulated annealing with a fixed temperature.

Surprisingly enough, many problems cannot be solved more efficiently by simulated annealing than by the Metropolis algorithm with the best temperature.

The problem of finding a natural example (artificial examples are known) where simulated annealing outperforms the Metropolis algorithm for all temperatures has been discussed by Jerrum and Sinclair (1996) as “an outstanding open problem.” This problem is solved here. The examples are instances of the well-known minimum spanning tree problem. Moreover, it is investigated which instances of the minimum spanning tree problem can be solved efficiently by simulated annealing. This is motivated by the aim to develop further methods to analyze the simulated annealing process.

See also: Ingo Wegener (2005): Simulated annealing beats Metropolis in combinatorial optimization. In Automata, Languages and Programming (ICALP 2005), Springer LNCS 3580, pages 589-601.

Robustness in Compositional Coevolution

R. Paul Wiegand (Naval Research - Washington, USA)

Though recent analysis of compositional (cooperative) coevolutionary algorithms (CCEAs) casts doubt on their suitability for static optimization tasks, our experience at the Naval Research Laboratory is that the algorithms perform quite well in multiagent learning settings. This is due in part because CCEAs may be quite suitable to finding behaviors for team members that result in good (though not necessarily optimal) performance but which are also robust to changes in other team members. Understanding what robustness means and how robust properties of a problem affect CCEA performance will help us develop a more constructive view of compositional coevolution. This talk aims to begin this process by meeting two main objectives.

First, I describe a general framework for clearly defining robustness, offering a specific definition for our studies. Second, I examine the hypothesis that CCEAs exploit this robustness property during their search. Using an existing theoretical model to gain intuition about the kind of problem properties that attract populations in the system, I provide a simple empirical study and a proposal for establishing a theoretical foundation justifying this intuition in a practical setting.

Keywords: Coevolution, compositional coevolution, robustness, optimization

Runtime Analysis of a Simple Ant Colony Optimization Algorithm

Carsten Witt (Universität Dortmund, D)

Ant Colony Optimization (ACO) has become quite popular in recent years. In contrast to many successful applications, the theoretical foundation of this randomized search heuristic is rather weak. Building up such a theory is demanded to understand how these heuristics work as well as to come up with better algorithms for certain problems. Up to now, only convergence results have been achieved showing that optimal solutions can be obtained in a finite amount of time. We present the first runtime analysis of a simple ACO algorithm that transfers many rigorous results with respect to the expected runtime of a simple evolutionary algorithm to our algorithm. In addition, we examine the choice of the evaporation factor, which is a crucial parameter in such an algorithm, in greater detail and analyze its effect with respect to the runtime.

Keywords: Randomized search heuristics, ant colony optimization, runtime analysis

Joint work of: Neumann, Frank; Witt, Carsten

Full Paper: <http://drops.dagstuhl.de/opus/volltexte/2006/592>