Substructural Local Search in Discrete Estimation of Distribution Algorithms

Cláudio F. Lima

Department of Electronics and Informatics Engineering
University of Algarve

clima.research@gmail.com

May 29, 2009

Outline

Topics

1. Background and Motivation.
2. Estimation of Distribution Algorithms.
    - Extended Compact Genetic Algorithm (eCGA).
    - Bayesian Optimization Algorithm (BOA).
3. Substructural Local Search in eCGA.
4. Substructural Local Search in BOA.
5. Model Accuracy in BOA.
6. Improved Substructural Local Search in BOA.
7. Summary and Conclusions.

Optimization Problems

- Are part of our daily life.
  - Looking for the shortest route with a GPS navigator.
  - Solving Sudoku puzzles.
  - Building schedules that must obey different constraints.
- Many current challenges in science are related to some optimization problem.
  - Problems are increasing in dimension and complexity.
- Efficient methods are required to solve such problems.
  - Methods must solve problems quickly and accurately.
  - But should also scale well for larger problem instances.
- Estimation of distribution algorithms (EDAs) provide scalable optimization through probabilistic modeling.

Motivation

- EDAs can solve many challenging problems in an efficient and scalable manner.
  - Significantly outperforming other methods.
- Although effective at exploring the search space, they inherit a common drawback from traditional global search methods.
  - Slower convergence to optimal solutions when compared to appropriate local searchers.
- This observation often leads to the combination of global and local search approaches.
- EDA research lacks on methods for designing and combining competent global and local-search methods.
  - That can exploit problem decomposition.

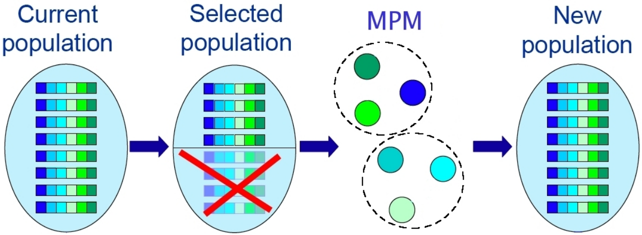
Estimation of Distribution Algorithms

EDAs

- Population-based search and optimization procedures.
- Population of promising solutions is modeled to generate new candidate solutions.
- “Similar” to genetic algorithms: replace genetic operators by building and sampling from a probabilistic model (PM).
- PMs used can range from simple univariate probability vectors to more complex Markov or Bayesian networks.
- Popular EDAs:
  - Population-based incremental learning (PBIL).
  - Extended compact genetic algorithm (ECGA).
  - Bayesian optimization algorithm (BOA).

Extended Compact Genetic Algorithm (eCGA)

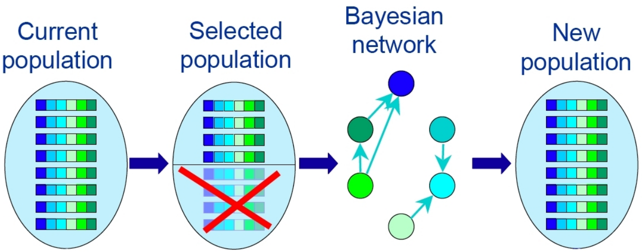
- Uses marginal product models (MPMs) to partition problem variables into non-overlapping clusters.



$$p(\mathbf{X}) = p(X_1 X_5 X_6) p(X_2 X_3 X_4)$$

Bayesian Optimization Algorithm (BOA)

- Uses Bayesian networks (BNs) to represent conditional dependencies between variables.



$$p(\mathbf{X}) = p(X_1|X_5X_6)p(X_6|X_5)p(X_5)p(X_3|X_2X_4)p(X_2)p(X_4)$$

Model Learning

- Both eCGA and BOA learn the structure of the probabilistic model at each generation.
- A simple learning algorithm is used.
  - Start with a simple model structure (no dependencies).
  - Increase model complexity while the likelihood of the model w.r.t. the data increases.
- Likelihood is quantified by a scoring metric.

EDA's great advantage over other methods

- Learned probabilistic models provide valuable information.
  - Get a better insight about the optimization problem.
  - Speedup the search process in EDAs even more.

Model-Based Efficiency Enhancement Techniques

Modeling solution quality in EDAs

- Probabilistic modeling in EDAs allow to identify patterns of good solutions.
- But association between solution components and their corresponding quality is also possible!
- The probabilistic model can also estimate the quality of solutions.
  - Saving costly function evaluations.

Induction of global neighborhoods

- The probabilistic models reveal important dependencies between variables that can be exploited in local search.

Substructural Local Search (SLS)

Traditional Local Search

- Typically neighborhood operators have fixed structure.
- Implicit tradeoff between generalization and efficiency.

SLS in EDAs

- Use substructural neighborhoods to perform local search in EDAs.
- Topology of neighborhoods is defined by the learned probabilistic models.
- Exploits underlying problem structure while not loosing generality of application.

Substructural Local Search in eCGA

SLS in eCGA

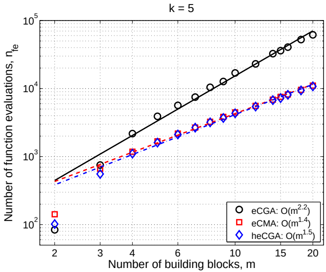
- Performed at each generation after model learning.
- Pick the best solution in the population.
- Each cluster of k variables in the probabilistic model is a substructural neighborhood.
- For each substructural neighborhood, evaluate all possible 2^k solutions and choose the best for searching the remaining neighborhoods.
- Model is updated based on the result of local search.
  - Equivalent to the insertion of the solution in the population.

Experimental Setup

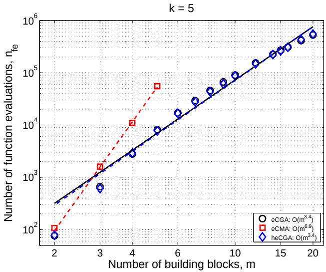
- Algorithms are tested on *adversarial problems* that represent different dimensions of problem difficulty.
  - Decomposable problems of bounded difficulty.
- Problems are composed by several subfunctions of size k.
  - Each subfunction is a very difficult problem which requires the detection and propagation of building-blocks (BBs).
  - Scaling between different subfunctions can be uniform or exponential.
  - Gaussian noise is also incorporated into the problem.
- Different types of modular interaction are also considered.
  - No interactions between problem variables.
  - Non-overlapping interaction.
  - Overlapping interaction.
  - Hierarchical interaction.

Scalability Results

Problem w/ uniform scaling

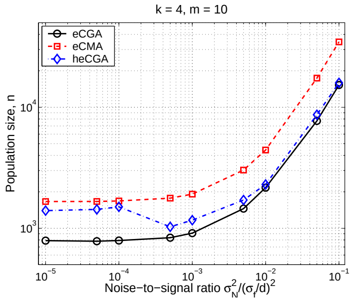


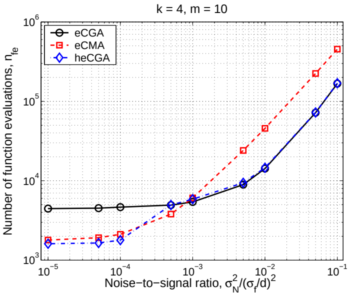
Problem w/ exponential scaling



- For uniform scaling the mutation approach works better.
- For exponential scaling recombination approach excel.
- eCGA+SLS performs well on both problems.

Results for a Noisy Problem





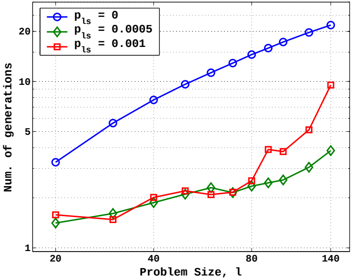
- eCGA+SLS follows the behavior of the best approach for each situation (low vs. high noise variance).
- Implicitly switches between a global and local search operator—*adaptive time continuation*.

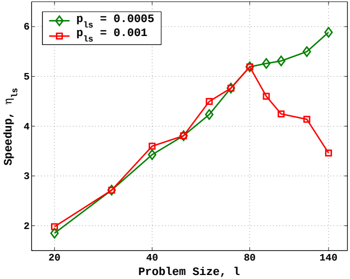
Substructural Local Search in BOA

SLS in BOA

- Performed at each generation after model learning.
- Applied to a proportion of the population (typically small).
- Each variable and its corresponding parent variables form a substructural neighborhood—*parental neighborhood*.
- For each substructural neighborhood, choose the best configuration according to the estimated fitness.
  - BNs also incorporate fitness information.
- Order for visiting neighborhoods is defined in a way that larger neighborhoods are considered first.
  - Larger neighborhoods contain more accurate statistical information.

Scalability Results

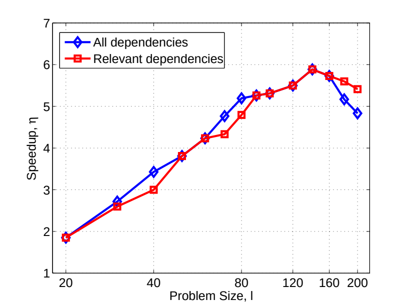




- SLS in BOA reduces number of iterations required.
- Provides significant speedups in number of evaluations.
- But speedup slows down for larger problem instances...

Understanding Why?

All vs. relevant neighborhoods



- Eliminate redundant neighborhoods. Ex: $\{X_1\}$ is already included in $\{X_1 X_2\}$
- Using only relevant neighborhoods doesn't help much.
- Looking closely at model structure one can see excessive complexity $\rightarrow$ *model overfitting*.

Model Structural Accuracy in BOA

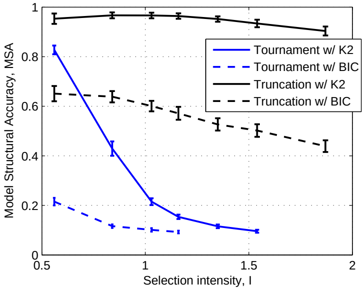
Measuring model accuracy

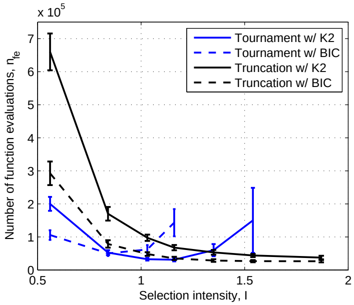
- Model structural accuracy (MSA) is defined as the ratio of correct edges over the total number of edges in the BN.
- An edge is correct if it connects two variables that are linked according to the problem definition.

Features compared

- Selection methods considered are Tournament and Truncation.
- Scoring metrics considered are K2 and BIC.
- Everything else in BOA is identical.

Selection-Metric Comparison



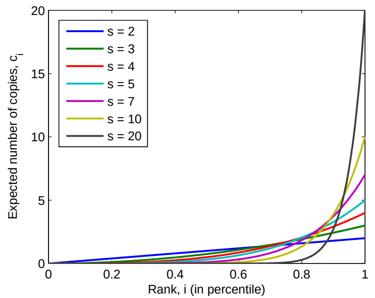


Observations

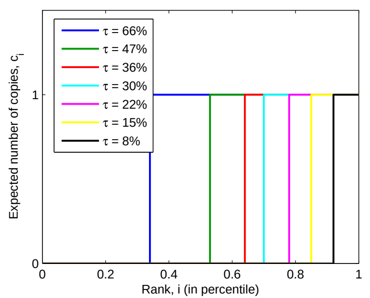
1. Truncation performs better than tournament selection.
2. K2 metric performs better than BIC metric.

Distribution of the Expected Number of Copies

Tournament selection



Truncation selection

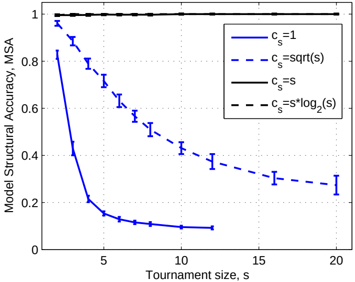


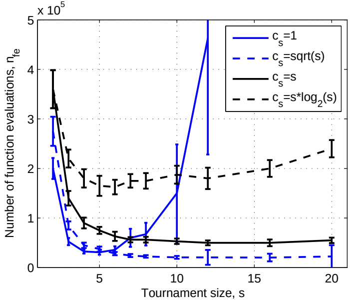
- Tournament selection generates the mating pool according to a power distribution, which leads to model overfitting.

Adaptive Scoring Metric

Complexity penalty for the K2 metric

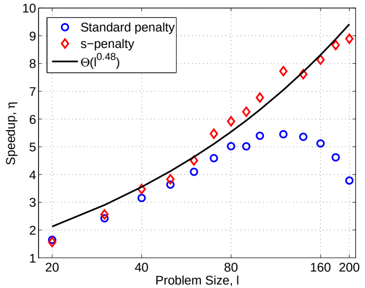
$$p(B) = 2^{-0.5 c_s \log_2(n) \sum_{i=1}^{\ell} |L_i|} \quad \longrightarrow \quad c_s = \{1, \sqrt{s}, s, \text{and } s \log_2(s)\}$$





Improved SLS in BOA

Influence of model accuracy

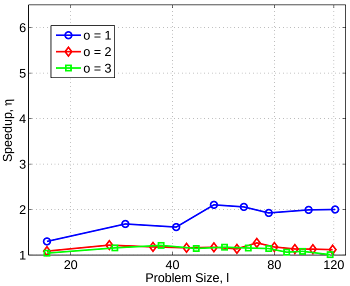


- More accurate models improve local search performance.
- Speedup consistently increases for larger ℓ.
- Speedup obtained approaches theoretical bound of $\theta(\sqrt{\ell})$.
- For $\ell = 200$, BOA+SLS takes almost an order of magnitude less evaluations than BOA.

Overlapping Difficulty

Decomposable problems with overlapping subfunctions

- Subfunction of size $k = 5$ and overlap of $o = \{1, 2, 3\}$.

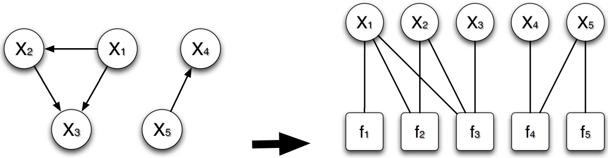


- Speedups do not carry over for problems with overlap.
- SLS does not take into account the context of overlapping interactions.
- A more general SLS method is required to tackle these problems.

Loopy Substructural Local Search in BOA

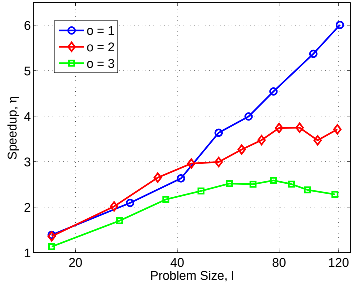
Loopy SLS

- Local search based on loopy belief propagation (LBP).
- LBP uses message-passing algorithms in graphical models (typically a factor graph).
- Exchange messages between factor nodes (squares) and variable nodes (circles) until reaching a stable situation.



Results for Loopy SLS

Problem w/ overlapping BBs



- Speedups improve significantly.
- For higher overlapping the speedup stagnates at some point.
- Overlapping effect is similar to noise which is known to be extremely hard for local search approaches.

- Even for high overlap, loopy SLS reduces the number of evaluations of BOA for less than a half.

Summary & Conclusions

- Thesis investigates SLS for discrete EDAs (eCGA, BOA).
- It is shown that incorporating SLS is advantageous for several types of boundedly-difficult problems.
- Empirical results demonstrate that SLS can substantially reduce the number of function evaluations.
  - Providing speedups superior to 10.
  - Speedup scale with problem size as $O(\sqrt{\ell})$.
- For problems where SLS does not improve the results, EDA performance is not compromised.
  - Because implicitly allows to switch between a global and a local search operator based on problem features.
- Study of model accuracy in BOA is also an important contribution to the field.