Current and Planned Projects at MEDAL

Martin Pelikan
Missouri Estimation of Distribution Algorithms Lab (MEDAL)
Department of Mathematics and Computer Science
University of Missouri in St. Louis
pelikan@cs.umsl.edu
Welcome
Program

9:30am  Martin Pelikan  (Missouri Estimation of Distribution Algorithms Lab (MEDAL), St. Louis, MO)  
Quo Vadis MEDAL: Current and Planned Projects at MEDAL

10:05am  Kenneth P. Turvey  (Missouri Estimation of Distribution Algorithms Lab (MEDAL), St. Louis, MO)  
XCS in Dynamic Environments

10:40am  Coffee Break

10:55am  Kumara Sastry  (Illinois Genetic Algorithms Lab (IlliGAL), Urbana, IL)  
Efficiency Enhancement of Estimation of Distribution Algorithms

11:30am  David E. Goldberg  (Illinois Genetic Algorithms Lab (IlliGAL), Urbana, IL)  
Little Models, Big Results

12:05pm  Lunch Break

1:30pm  Cezary Janikow  (University of Missouri, St. Louis, MO)  
Representations and Heuristics in GP

2:05pm  Moshe Looks  (Washington University, St. Louis, MO) and Ben Goertzel  (Novamente LLC)  
Towards Competent Genetic Programming - What are the Missing Ingredients?

2:40pm  Shigeyoshi Tsutsui  (Hannan University, Matabara, Osaka, Japan)  
Node Histogram vs. Edge Histogram: A Comparison of PMBGAs in Permutation Domains

3:15pm  Coffee Break

3:30pm  Mark Jakiela  (Washington University, St. Louis, MO)  
Open collaborative design of tangible artifacts - simulation and optimization?

4:05pm  Open Discussion
This Talk

- Current and planned MEDAL projects
  - Optimization via probabilistic modeling.
  - Learning classifier systems.
  - Neural nets with evolutionary algorithms.
Optimization

- Task
  - Given set of candidate solutions & evaluation function.
  - Find the best candidate solution(s).

- Three goals
  - Robustness
    Applicability to a broad range of problems.
  - Scalability
    Scalable performance in tough problems.
  - Practicality
    From academia to industry.
Optimization via Probabilistic Modeling

- Estimation of distribution algorithms (EDAs)
  - Combine evolutionary computing and machine learning.
  - Maintain a population of candidate solutions.
  - Build probabilistic model of high-quality candidates.
  - Sample built model to generate new candidates.
Past Work on EDAs

- Many powerful EDAs
  - Scalable solutions for previously intractable problems.
  - Powerful efficiency enhancement techniques.
  - Important applications in various areas.
  - Most work for fixed-length, discrete vector solutions, but important work in other domains, too.

- Hierarchical BOA (Pelikan, Goldberg, 2001)
  - Solutions are fixed-length, discrete vectors.
  - hBOA scalably solves nearly decomposable and hierarchical problems, $O(n^2)$ evaluations or faster.
  - Can be used for other representations if solutions can be mapped into fixed-length discrete vectors.
Challenges

- Extend current EDAs to capture other regularities
  - Example: Motifs=repeated substrings (Looks, 2006).

- Match progress for fixed-length, discrete representations in other representations.
  - Real-valued / variable length vectors.
  - Permutations and schedules.
  - Program codes / labeled trees.

- Efficiency enhancement.

- Applications.
Efficiency Enhancement

- Efficiency enhancement (EE)
  - Low-order polynomial scalability is great, but sometimes not enough.
  - Use EE methods to further improve efficiency.
- Example: Parallelization
Efficiency Enhancement: Past Work

- Past work
  - Parallelization.
    - Can effectively use large parallel computers.
  - Fitness evaluation relaxation.
    - Can have speedups of 30-50 on single processor.
  - Hybridization.
    - Can solve intractable problems with good hybrids.
  - Prior knowledge utilization.
    - Can incorporate various kinds of prior knowledge.
  - Incremental and sporadic model building.
    - Leads to significant speedups without affecting scalability.
Example: Parallelization

- Diamond - speedup for 2D spin glass 20x20, N=4000
- Dot - predicted speedup for 2D spin glass 20x20, N=4000
- Circle - speedup for 2D spin glass 25x25, N=6000
- Triangle - predicted speedup for 2D spin glass 25x25, N=6000
- Triangle with dot - speedup for 2D spin glass 30x30, N=8000
- Square with dot - predicted speedup for 2D spin glass 30x30, N=8000
Example: Hybridization
Example: Sporadic Model-Building

![Graph showing the relationship between problem size and CPU speedup with SMB.](image)
Challenges in Efficiency Enhancement

- Design new efficiency enhancement techniques.
- Combine efficiency enhancement techniques to multiply speedups
  - What combinations make sense?
  - How to do come these techniques well?
  - Expected speedups?
- Examples
  - Parallelization of model building and hybridization complement each other.
  - Parallelization of model building and sporadic model building multiply.
Applications

- Computational physics
  - Spin glasses, structure optimization (molecules).
- Complexity theory
  - Graph problems.
- Classification
  - Feature subset selection, feature extraction.
- Medicine and bioinformatics
  - Mostly classification, but also other problems.
- Scheduling and OR problems
Learning Classifier Systems

- LCS combine evolutionary computation and reinforcement learning.

- Task
  - Environment defines states, actions, rewards.
  - Learn reward predictions for state/action pairs.
  - Use inspiration from evolutionary computation.
LCS Challenges

- LCS in dynamic environments
  - How do LCS work in dynamic environments?
  - Can we make LCS work better in such environments?
- Transfer of ideas from EC to LCSs
  - Can we learn from EC or EDAs to make better LCSs?
  - Example: Efficiency enhancement, model building.
- Applications of LCS
  - Classification.
  - Bioinformatics.
Neural Nets: Neuroevolution

- Neural networks (NNs)
  - Maps inputs to outputs.
  - Powerful way to model complex functions.
  - Incorporating evolutionary computing into NNs leads to important results.
NN Challenges and Applications

- Challenges
  - How to design good network topology?
  - How to choose good connection weights?
  - How to do “linkage learning” in neural nets?
  - Many neurons with few connections or few neurons with many connections?

- Applications
  - Control (in simulated environments).
  - Modeling/regression.
Example: Enforced Subpopulations

Environment
Example Environment: RARS

- Robot Auto Racing Simulation
  http://rars.sourceforge.net/
Summary

- Projects in 3 areas
  - Optimization using probabilistic modeling.
  - Learning classifier systems.
  - Neural networks.
- 3 research directions
  - Design and analysis.
  - Efficiency enhancement.
  - Applications.
Thank You

- Questions or comments?
- Enjoy the event!