One Thing at a Time: Evaluating Profit and Legality, or On the Contributions of Losers: Exploring the Benefits of Retaining Infeasible Solutions During Search for a Constrained Optimum

Steven Orla Kimbrough, kimbrough@wharton.upenn.edu
Ming Lu, milu@wharton.upenn,edu
David Harlan Wood, wood@cis.udel.edu

20 & 24 October, 2004, Atlanta, Denver

Id

Abstract

Constraints pose severe challenges to optimization by evolutionary computation. Genetic operators (mutation, crossover, etc.) are not guaranteed to maintain feasibility and typically in practice a large number of infeasible solutions are generated, leading to poor performance of evolutionary algorithms.

We have developed what we call a feasible-infeasible two-population (FI-2Pop) genetic algorithm that circumvents these problems in a principled and surprisingly successful way. Two populations are maintained, one for feasible solutions, one for infeasible solutions. Descendants are placed in whichever population fits them (depending on whether they are feasible or not).

In this talk we motivate and discuss the FI-2Pop GA, describe its

impressive performance (to date), and discuss how the infeasible solutions contribute to the search process.. We also discuss how the algorithm may be used to obtain approximate shadow prices and other useful sensitivity information, valuable for managerial decision making. Finally, we present theoretical considerations that help to explain the algoritm's success.

Papers are available from the authors.

Papers

All available at: `http://opim-sun.wharton.upenn.edu/~sok/comprats/index.html#evolutionarycomputation`

NB, participation and contributions throughout by Ming Lu.

"Introducing Distance Tracing of Evolutionary Dynamics in a Feasible-Infeasible Two-Population (FI-2Pop) Genetic Algorithm for Constrained Optimization," Steven Orla Kimbrough, Ming Lu, and David Harlan Wood.

"Exploring the Evolutionary Details of a Feasible-Infeasible Two-Population GA," Steven Orla Kimbrough, Ming Lu, and David Harlan Wood. PPSN VIII, September 18-22, 2004, Birmingham, UK.

"Exploring a Financial Product Model with a Two-Population Genetic Algorithm" Steven O. Kimbrough, Ming Lu, and Soofi M.

Safavi, *Proceedings of the 2004 Congress on Evolutionary Computation (CEC2004)*, June 19-23, 2004, Portland, OR, pages 855-862.

"Exploring a Two-Population Genetic Algorithm" Steven O. Kimbrough, Ming Lu, David H. Wood, and D.J. Wu. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2003)*, E. Cantu-Paz and et al., eds., Springer, LNCS 2723, pages 1148–1159.

"Exploring a Two-Market Genetic Algorithm" (Steven O. Kimbrough, Ming Lu, David Harlan Wood and D.J. Wu), in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2002)*, W. B. Langdon and E. Cantu-Paz and et al., eds., Morgan Kaufmann, SF, CA, 2002, pp. 415-21.

Part 0:

The Short Version

Background: Interest in Stochastic Search Algorithms

- Especially for constrained optimization. (a) solution, (b) interpretation (b-i) sensitivity analysis (b-ii) candle-lighting analysis (optimize the objective *and* optimize the constraints)
- Strengths of weak methods: generality, approach to NP-hard problems, not letting the best be enemy of the good
- No Free Lunch theorem(s): there can't be a dominating heuristic.

  NB Metaheuristic?
- So how can we get better heuristics? (Our fundamental question)

  Our theme: more information. (a) don't waste what you have, (b) gather more

Topic: Constrained Optimization

- Problem: Genetic operators do not respect feasibility.
- What to do about infeasible solutions generated during the search process?
- Many techniques have been proposed and used, but none of them is fully satisfactory.
- Most common technique—penalizing infeasible solutions—is also not fully satisfactory.
- Here, a new idea.

Main Idea: Evolve 2 Populations, One Feasible, One Infeasible

- Feasible population: Evaluated only on objective function.
- Infeasible population: Evaluated only on constraint violations.
- A standard GA for each population.
- Breeding only within a population.
- Offspring are placed in the feasible or infeasible population, depending on whether they are feasible or infeasible.

How Well Does It—FI-2Pop GA—Work?

- PPSN04 paper: Examined four challenge problems from the OR literature on nonlinear optimization. FI-2Pop GA did very well and compared favorably with GENOCOP. See summary slide.
- For an example of the challenge problems, see Yuan (follows), which is mix-integer and nonlinear in both the objective and the constraints.
- See 3 previous papers (cited here) for other examples.
- In sum: The FI-2Pop GA has been excellent at finding good solutions, does well on average, and has a comparatively low standard deviation.

Problem #1: Yuan's formulation is as follows

Objective function:

$$\min_{\mathbf{x},\mathbf{y}} z = (y_1 - 1)^2 + (y_2 - 2)^2 + (y_3 - 1)^2 - \ln(y_4 + 1) + (x_1 - 2)^2 + (x_2 - 2)^2 + (x_3 - 3)^2 \quad (1)$$

$$y_1 + y_2 + y_3 + x_1 + x_2 + x_3 \leq 5 \quad (2)$$

$$y_3^2 + x_1^2 + x_2^2 + x_3^2 \leq 5.5 \quad (3)$$

$$y_1 + x_1 \leq 1.2 \quad (4)$$

$$y_2 + x_2 \leq 1.8 \quad (5)$$

$$y_3 + x_3 \leq 2.5 \quad (6)$$

$$y_4 + x_1 \leq 1.2 \quad (7)$$

$$y_2^2 + x_2^2 \leq 1.64 \quad (8)$$

$$y_3^2 + x_3^2 \leq 4.25 \quad (9)$$

$$y_2^2 + x_3^2 \leq 4.64 \quad (10)$$

$$x_1, x_2, x_3 \geq 0 \quad (11)$$

$$y_1, y_2, y_3, y_4 \in \{0, 1\} \quad (12)$$

At optimality the (minimizing) value of the objective function is $z^* = 4.5796$, with the decision variables set at $\mathbf{x}^* = (0.2, 0.8, 1.908)^T$ and $\mathbf{y}^* = (1, 1, 0, 1)^T$.

Anything Else?

- Yes! Candle-lighting. ("It is better to light one candle than to curse the darkness.")
- Idea: Use solutions from a constrained optimization model to inform the decision maker regarding, e.g., opportunities for tightening or loosening constraints.
- Think: Shadow prices and reduced costs.
- Together, the traces of the two populations (feasible, infeasible) contain much valuable information of this kind. See discussion in the paper.

Why Does (Should) It Work?

- Observation: Usually, optimal solutions are on or near the boundary of the feasible region.

- Observation: The infeasible population probes and explores the boundary of the feasible region.

- Intuition: Infeasible solutions, especially those created from the feasible population after a number of generations, contain information and valuable genetic material.

  There must be some rate of exchange between an infeasible solution and a randomly-generated feasible solution.

- Observation: Doubling the size of a population does not double the rate of evolution.

- Observation: As a run progresses, both populations often tend to cluster near optimal regions.

Future Work (A lot!)

- Broader, deeper testing with real-world examples.
- Theory: especially on efficiency of using information.
- Extensions: e.g., interpolation, breeding across populations, hybrid algorithms, etc.
- Further development of candle-lighting ideas.

Part 1:

FI-2Pop GA for Constrained Optimization

GAs and Constrained Optimization

- Constrained optimization is difficult: genetic operators don't respect feasibility.

- Approaches to handling infeasible solutions: death penalty (wasteful), penalty functions (but which?), special encoding to guarantee feasibility (rarely available), repair of infeasible solutions (problem-specific; computationally costly, often).

- Why bother with GAs? Some metaheuristic is often required. GAs are easy to use.

- Elsewhere: GAs leave useful information as a side-effect (candle-lighting analysis).

Our New Idea: Separate Feasibility and Optimality Evaluations

- GECCO 2002: Two market GA. Excellent results on knapsack problems, tested against standard penalty methods.
- But: knapsack, while NP-hard, is thought to be atypically easy.
- But: there are standard GA packages out there for comparison.
- And: an improvement, refinement, on the concept: two population GA.

New Results Pertaining to GECCO 2002 Paper

- Previously: "2 market GA" versus standard penalty methods. Result: 2 market GA generally does better, and notably better on apparently harder problems.
- New: FI-2Pop GA versus Genocop4 on the same problems. 20 runs with (full) populations of 50 for 10,000 generations (feasible and infeasible).
- Next page: comparison of best solution found across 20 runs. Upshot: Genocop4 does better than standard constraint methods, *when it can find a feasible solution.* FI-2Pop GA always finds a feasible solution, performs better than either method from GECCO 2002, and almost always does better than Genocop4.
- In only one case does FI-2Pop GA fail to find the optimal solution.

problem	2-pop	Genocop4	optimal
hp2	3186	3186	3186
pb6	776	n/a	776
pb7	1034	1035	1035
pet7	16537	16524	16537
sento1	7772	n/a	7772
sento2	8722	8720	8722
weing7	1094945	1093705	1095445
weing8	624319	n/a	624319
weish12	6339	6339	6339
weish17	8633	8633	8633
weish21	9074	9067	9074
weish22	8947	n/a	8947
weish25	9939	9922	9939
weish29	9410	n/a	9410

GECCO 2003 Paper

- Beyond knapsack to nonlinear and mixed integer problems.
- Comparison with performance of a standard GA package, Genocop II/III.
- Refined concept: two population GA.

  Feasible population. Infeasible population. Evaluation alternates. Children are placed where they belong.

Group 1 Problems, from Michalewicz

Problem	Min/Max	Objective	# Variables	# Linear Inequalities	# Nonlinear Inequalities
1	min	Linear	8	3	3
2	min	Polynomial	7	0	4
3	min	Quadratic	10	3	5
4	min	Polynomial	2	0	2
5	min	Linear	2	0	2
6	min	Polynomial	2	0	2
7	min	Quadratic	2	0	2
8	min	Quadratic	2	1	1
9	max	Nonlinear	20	1	1
10	max	Nonlinear	50	1	1

Nonlinear, with floating point variables. Arithmetic crossover, non-uniform mutation.

Group 1 Results: With Untuned 2 Market GA

Prob.	Min/ Max	Best Known or Optimal*	Genocop II/III	Two-Market GA		
				Best of 10	Median	Std.
⇒1	min	7049.330923*	7268.650	∅	∅	∅
2	min	680.6300573*	680.640	680.6374	680.7566	0.085123
⇒3	min	24.3062091*	25.883	25.18437	25.61935	0.825277
4	min	0.25*	close	0.25	0.25	0
5	min	-5.5079*	close	-5.50773	-5.50708	0.001042
6	min	-6961.81381*	close	-6960.95	-6957.44	1.908346
7	min	5*	close	5	5	0
8	min	1*	close	1	1	0
⇒9	max	0.80351067	0.80351067	0.80288	0.7888	0.12853
10	max	0.8348	0.83319378	0.809211	0.743844	0.054728

Group 2 Problems: GLOBAL Library at GAMS World

Problem	Min/Max	Objective	# Variables	# Linear Inequalities	# Nonlinear Inequalities
11 (test11)	max	Nonlinear	2	0	2
12 (chance)	min	Linear	3	2	1
13 (circle)	min	Linear	3	0	10
14 (ex3_1_4)	min	Linear	3	2	1
15 (ex7_3_1)	min	Linear	4	6	1
16 (ex7_3_2)	min	Linear	4	6	1
17 (ex14_1_1)	min	Linear	3	0	4
18 (st_e08)	min	Linear	2	0	2
19 (st_e12)	min	Nonlinear	3	3	0
20 (st_e19)	min	Polynomial	2	1	1
21 (st_e41)	min	Nonlinear	4	0	2

Group 2 Results

Problem	Best Known or Optimal*	Genocop III		Two-Population GA	
		Best of 10	Std.	Best of 10	Std.
11	?	0.115047	4.87E-05	0.115047	4.16E-08
12	29.8943781591	29.89549	0.034976	†	0.032017
13	4.57424778502	4.574318	0.027615	4.574248	6.67E-05
14	-4.0000	-4	0.032482	-4	0.000131
15	0.341739553124	0.3558	0.141607	‡	0.00033
16	1.08986397147	1.09145	0.033588	1.089952	2.56E-05
17	0.0000	1.44E-10	2.85E-06	4.69E-05	1.36E-05
18	?	0.741782	9.93E-09	0.741782	5.58E-08
19	?	-4.5099	0.02252	-4.51347	0.000995
20	?	-118.705	0.021621	-118.705	0.000187
21	?	645.626	4.398663	641.8244	2.663347

† = 29.8943786178232, ‡ =**0.319777729979837**. NB. untuned, lower variance.

Infeasible Children Decrease over Time

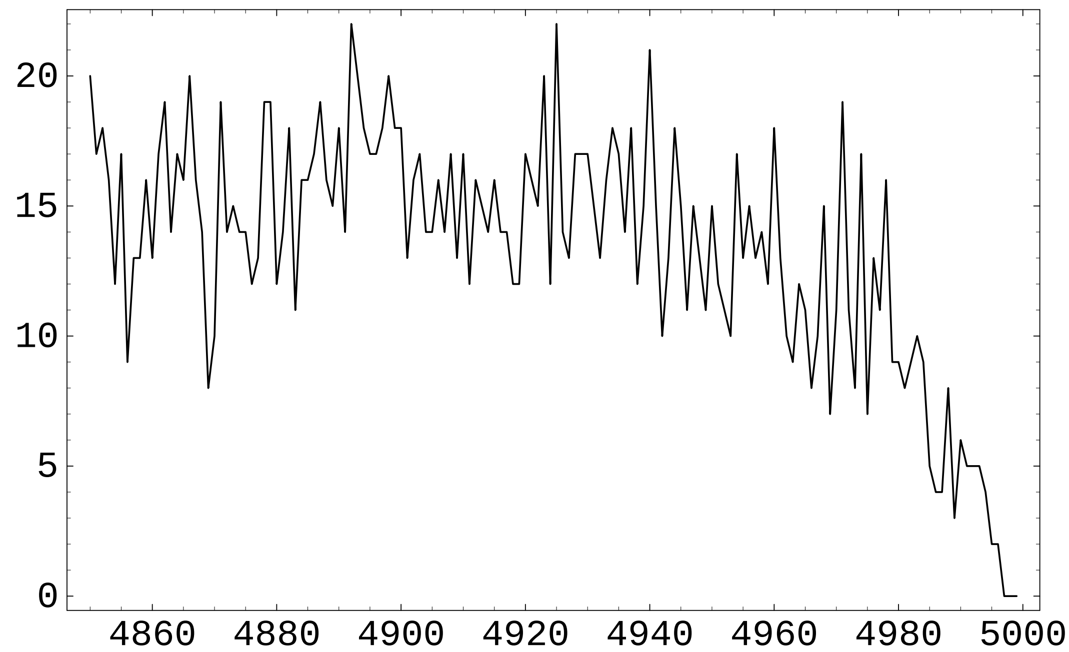


Figure 1: For problem ex7_3_2 with population size 50: count of number of infeasible solutions created from the feasible population, by generation, 4850–4999.

Continuing & Increasing Contribution by the Infeasible

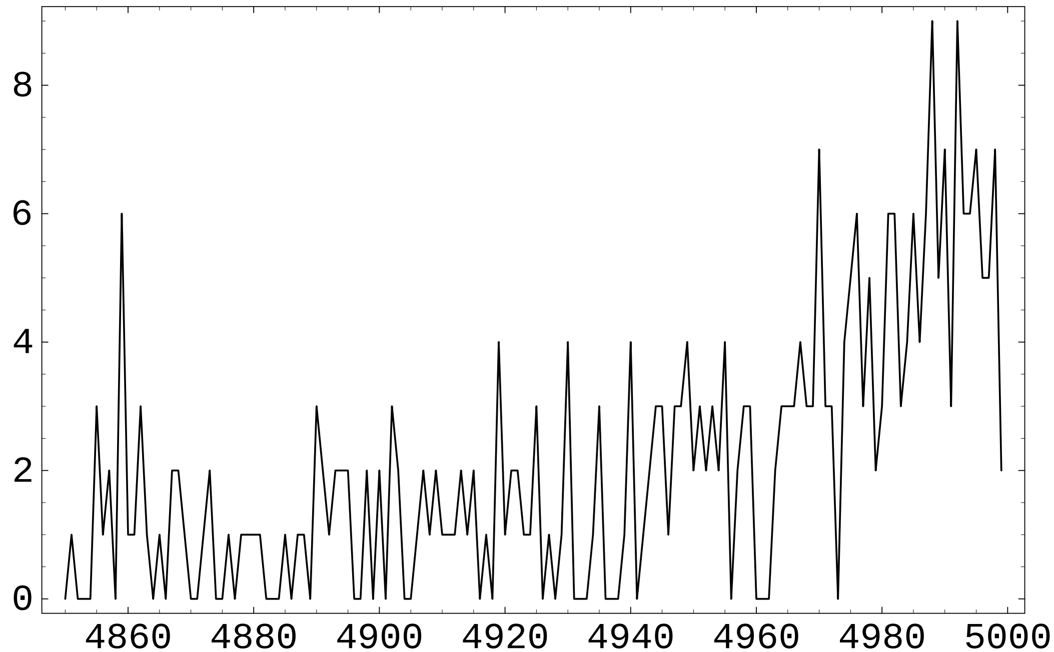


Figure 2: For problem ex7_3_2 with population size 50: count of number of feasible solutions created from the infeasible population, by generation, 4850–4999.

Everyone Is Descended from a Loser

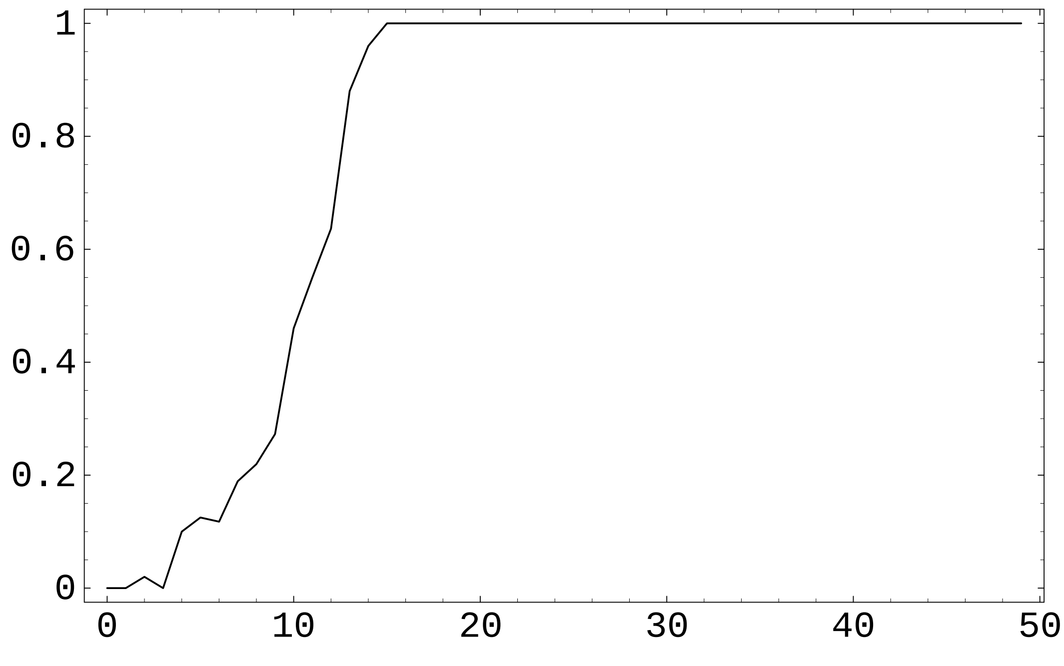


Figure 3: For problem ex7_3_2 with population size 50: fraction, by generation, of feasible individuals with an infeasible ancestor.

Summary: Why/How Does It Work?

- The continuing contributions of losers.
- The infeasible population (after some selection) retains valuable genetic information.
- There is some optimal rate of exchange between having an extra feasible solution and having an extra infeasible solution.

Late Breaking Results: Comparison with PSO

- See: Parsopoulos, K.E., Vrahatis, M.N., “Particle Swarm Optimization Method for Constrained Optimization Problems,” in P. Sincak, J. Vascak, V. Kvasnicka, J. Pospichal (eds.), *Intelligent Technologies - Theory and Applications: New Trends in Intelligent Technologies*, pp. 214-220, IOS Press (Frontiers in Artificial Intelligence and Applications series, Vol. 76), 2002, ISBN: 1-58603-256-9. `http://www.math.upatras.gr/~kostasp/papers/eisci.pdf`. And generally, `http://www.math.upatras.gr/~kostasp/public.html`.

- Used PSO with dynamic penalty functions to solve a number of test problems.

- We examined 3 of their 6 problems.

Problem Summary

Test Problem	Min/ Max	Objective Function	# Vars.	# Linear Inequalities	# Nonlinear Inequalities
#4	Min	Quadratic	5	0	6
#5	Min	Quadratic	5	0	6
#6	Min	Quadratic	6	2	0

Problems #4 & #5

$$\min f(x) = 5.3578547x_3^2 + 0.8356891x_1x_5 + 37.293239x_1 - 40792.141$$

subject to $0 \leq 85.334407 + 0.0056858T_1 + T_2x_1x_4 - 0.0022053x_3x_5 \leq 92$, $90 \leq 80.51249 + 0.0071317x_2x_5 + 0.0029955x_1x_2 + 0.0021813x_3^2 \leq 110$, $20 \leq 9.300961 + 0.0047026x_3x_5 + 0.0012547x_1x_3 + 0.0019085x_3x_4 \leq 25$, $78 \leq x_1 \leq 102$, $33 \leq x_2 \leq 45$, $27 \leq x_i \leq 45$, $i = 3, 4, 5$,

where in #4, $T_1 = x_2x_5$ and $T_2 = 0.0006262$ and for #5, $T_1 = x_2x_3, T_2 = 0.00026$. Best known solution for #4 is $f^* = -30665.538$. #5 has an unknown best solution.

Problem #6

$$\min f(x,y) = -10.5x_1 - 7.5x_2 - 3.5x_3 - 2.5x_4 - 1.5x_5 - 10y - 0.5\sum_{i=1}^{5} x_i^2$$

subject to $6x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 - 6.5 \leq 0$, $10x_1 + 10x_3 + y \leq 20$, $0 \leq x_i \leq 1$, $i = 1, \ldots, 5$, $0 \leq y$.

Best known solution for #6 is $f^* = -213.0$.

Results Compared to Genocop

Settings as in GECCO 2003 paper: population size 50, 10,000 generations, 10 runs each.

Test Problem	Best Known or Optimal	Genocop III		Two Population GA	
		Best	Std.	Best	Std.
#4	-30665.538	-30663.029297	45.381845	-30665.509108	0.019961
#5	?	-30986.876953	35.950656	-31026.424543	0.029210
#6	-213	-212.3456573	3.684304	-213	0

Two-market GA does better in all three cases: best and for std.

PSO (mostly infeasible) vs. feasible 2 pop. GA

Unspecified computational effort. Sum V.C.= sum of violated constraints. PSO solutions generally not feasible.

Test Problem	Method	Best Solution (Sum V.C.)	Mean (Sum V.C.)	2 Pop GA
#4	PSO-In	-31543.484 (1.311)	-31526.304 (1.297)	-30665.509
	PSO-Co	-31542.578 (1.311)	-31528.289 (1.326)	
	PSO-Bo	-31544.459 (1.311)	-31493.190 (1.331)	
#5	PSO-In	-31544.036 (0.997)	-31523.859 (0.958)	-31026.425
	PSO-Co	-31543.312 (0.996)	-31526.308 (0.965)	
	PSO-Bo	-31545.054 (0.999)	-31525.492 (0.968)	
#6	PSO-In	-213.0 (0.0)	-213.0 (0.0)	-213.0
	PSO-Co	-213.0 (0.0)	-213.0 (0.0)	
	PSO-Bo	-213.0 (0.0)	-213.0 (0.0)	

Set Covering Problems

Th. Bäck, M. Schütz and S. Khuri: "A Comparative Study of a Penalty Function, a Repair Heuristic, and Stochastic Operators with the Set-Covering Problem," J. M. Alliot, E. Lutton, E. Ronald, M. Schoenhauer and D. Snyers, editors: *Artificial Evolution*, pp. 3-20, Springer, Berlin, 1996. `http://ls11-www.cs.uni-dortmund.de/people/baeck/papers/ea95.ps.gz`

Links to Thomas Bäck's website:

`http://ls11-www.cs.uni-dortmund.de/people/baeck/`

`http://ls11-www.cs.uni-dortmund.de/people/baeck/ea_application.html`

(The downloadable files are actually in .ps format, not ps.gz. You need rename files and you can use Adobe Acrobat to convert them to pdf files.)

7 set covering problems from Beasleys OR-library

We apply our two-population GA on 7 set covering problems (all minimizing) from Beasleys OR-library. For detailed information and the data file for these test problems, please refer to the following website: http://mscmga.ms.ic.ac.uk/jeb/orlib/scpinfo.html. m=number of constraints. n=number of (0-1) decision variables.

Problem	m	n	density
Scp41	200	1000	2
Scp42	200	1000	2
Scpe1	50	500	20
Scpe2	50	500	20
Scpe3	50	500	20
Scpe4	50	500	20
Scpe5	50	500	20

Mixture of Results Compared to Bäck et al.

Test Problem	Best-Known /Optimal	Two Population GA			Bäck et al.	
		Best	Avg.	Stdev.	Penalty	Repair
Scp41	429	1069	1385	230.1267	1030	439
Scp42	512	1208	1497	181.1648	1208	536
Scpe1	5	6	7	0.92582	6	8
Scpe2	5	6	6.857143	0.638877	7	7
Scpe3	5	6	7	0.755929	7	7
Scpe4	5	7	7.285714	0.451754	7	6
Scpe5	5	6	6.928571	1.083268	8	7

Computational effort: Bäck et al.: population size = 100, generations=5000. We did 50, 50 at 5,000 each. But, cost of repair?

Results from CEC 2004 Paper

- Challenging financial services model for configuring mortgage refinancing deals. Formulated to compare with heuristics.
- Mixed integer, nonlinear. Can't be solved by any other system we could find. (Not GAMS & Co., not GENOCOP).
- Minimizing with previous heuristic $z^+ \approx 90000$
- Relaxed version is linear (but ignores 42 binary variables). Solved by CPLEX to 66282, but solution is (highly) infeasible.
- FI-2Pop GA found many close *and feasible* solutions, the best at $z^+ = 66479$ or within 0.3% of the solution to the relaxed problem!
- NB Difficulty in finding a feasible solution. The first one evolved from the infeasible population in generation 5619.

Results from PPSN 2004 Paper

Generations	Violation	InF→Fea	Fea→InF	z^+	medz_{InF}	$\sigma^2 z_{\mathrm{Fea}}$	$\sigma^2 z_{\mathrm{InF}}$
0–99	-0.2824	3.5400	7.3000	5.222503	7.123	2.302	6.839
900–999	-0.2005	3.4100	6.6200	4.594130	6.577	0.840	8.928
1900–1999	-0.0453	3.3100	6.4000	4.581232	9.468	1.015	7.713
2900–2999	-0.0858	3.0400	6.4800	4.579938	5.926	0.426	3.302
3900–3999	-0.0501	2.7000	6.3300	4.579845	5.103	0.251	1.775
4900–4999	-0.0126	3.2900	4.8200	4.579653	5.245	0.253	0.948

Table 1: Yuan Results: Averages over 100 generations. Violation$=-1\cdot$sum of absolute violations of constraints (averaged over each solution for 100 generations). InF→Fea=number of feasible offspring from the infeasible population (by generation, averaged over 100 generations). Fea→InF =number of infeasible offspring from the feasible population (by generation, averaged over 100 generations). z^+=best objective function value found in the feasible population (by generation, averaged over 100 generations). medz_{InF}=median objective function value in the infeasible population (by generation, averaged over 100 generations). $\sigma^2 z_{\mathrm{Fea}}$=variance of objective function values in the feasible population (averaged over all solutions in 100 generations). $\sigma^2 z_{\mathrm{InF}}$=variance of objective function values in the infeasible population (averaged over all solutions in 100 generations).

PPSN-2004: Comparison with Genocop

problem	Best known	Genocop III		Two-population GA	
		Best of 10	Std.	Best of 10	Std.
Hesse (min)	-310	-306.972	16.185309	-309.9907	0.044785738
Pooling (max)	450	433.6981	37.02992564	444.157	15.08449889
Colville (min)	10122.69643	10126.6504	108.3765361	10122.8412	0.715469975

Table 2: Comparison of the Two-Population GA with Genocop III

NB Mean and variance a probably better measures of comparison. The FI-2Pop GA consistently does well, does better.

Results from a Portfolio Problem

- 300+ continuous variables, 20+ binary variables
- Standard portfolio problem: quadratic in the objective, binaries in constraints
- Regularly beat GAMS

Comment: Preservation and Second Chances

- Repair strategies can be understood as ways of preserving genetic information in infeasible solutions.

- The two-market GA is also a way of preserving genetic information in infeasible solutions.

- Loci may often interact, e.g., w(ab) $>$ w(AB) $>$ w(aB) $>$ w(Ab). [w(.) $=$ fitness]

- Evolution of populations affords chances of discovering, e.g., a path from AB to aa. The adjunct, infeasible population affords this in constrained optimization.

In sum: remarkably good and robust performance from a simple, untuned approach

Next steps:

- More problems.
- Tuning. Exchange rate discovery; dynamic allocation between feasible and infeasible.
- Multi-population GAs.
- Broader comparison, e.g., with other metaheuristics.
- Mathematical and empirical analysis.

Briefly on Candle-Lighting Analysis

- "It is better to light one candle than to curse the darkness."
- Optimizing the objective function: tells you the best deal you can get given the model.
- Candle-lighting analysis is optimizing the constraints (and objective function coefficients): tells you how to get a better deal.
- Candle-lighting. Well on beyond sensitivity analysis (robustness checking). Use a model to direct search and investment.
- The trace (history) of the FI-2Pop GA, combined with SQL queries, has consistently yielded (a) nearly feasible solutions with improved objective values; (b) feasible solutions with mildly degraded objectives and improved slackness. See PPSN and CEC papers especially.

Part 2:

Comments Regarding Theory

Questions & Observations

- Why does the FI-2Pop GA work (so well so often)?
- What lessons lurk for the design of better metaheuristics?

  NB: There can be no optimal heuristic. The question is what we can learn from successes about the design of heuristics.
- No tidy theory yet. Will proceed with a series of observations.
- See CIST (distance tracing; draft) paper for revealing graphics on a run of the Yuan problem. Introduced rescaled distance measure for describing runs of a GA (population-based metaheuristic):

$$d(\mu_x, \mu_y) = (\vec{\mu}_x - \vec{\mu}_y)C^{-1}(\vec{\mu}_x - \vec{\mu}_y)^T \tag{13}$$

General Argument (or Observations)

1. Inductive bias is inevitable.

2. Some commitments have to be made. And it is not always clear what the attendant biases are.

3. Strategies for meliorating bias and making better commitments: (a) don't waste information, (b) hedge your bets, (c) learn (acquire information) on-line, (d) learn off-line by sampling the space of problems.

4. This requires: (a) a source of information, (b) an adjustable mechanism.

Observation #1: Penalty functions waste information.

- Assume: optimal region is on or near the boundary of the feasible region. (No assumption about convexity, connectedness, etc.)
- Goal is to produce a solution that is in the (arbitrarily small) optimal region.
- Assume: The closer a solution is to the optimal region, the higher the probability is that a mutation, etc. will yield a solution in the optimal region. $H(d) =$ score a hit from distance d.

  $P(H(d) = 1) = \int_{\mu+d}^{\mu+d+\delta} f(x)dx$ & $P(H(d) = 1) > P(H(d') = 1)$ iff $d < d'$.

1-Dimensional Model

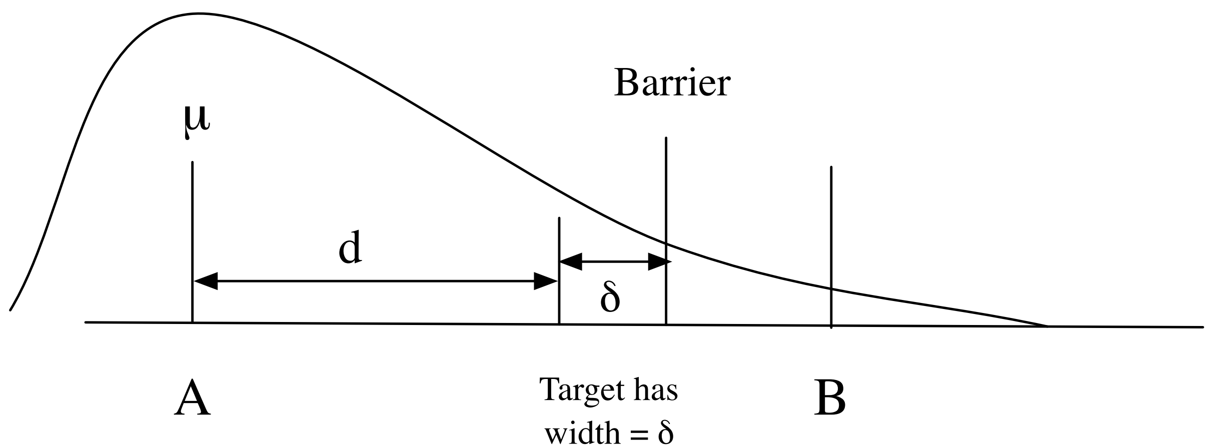


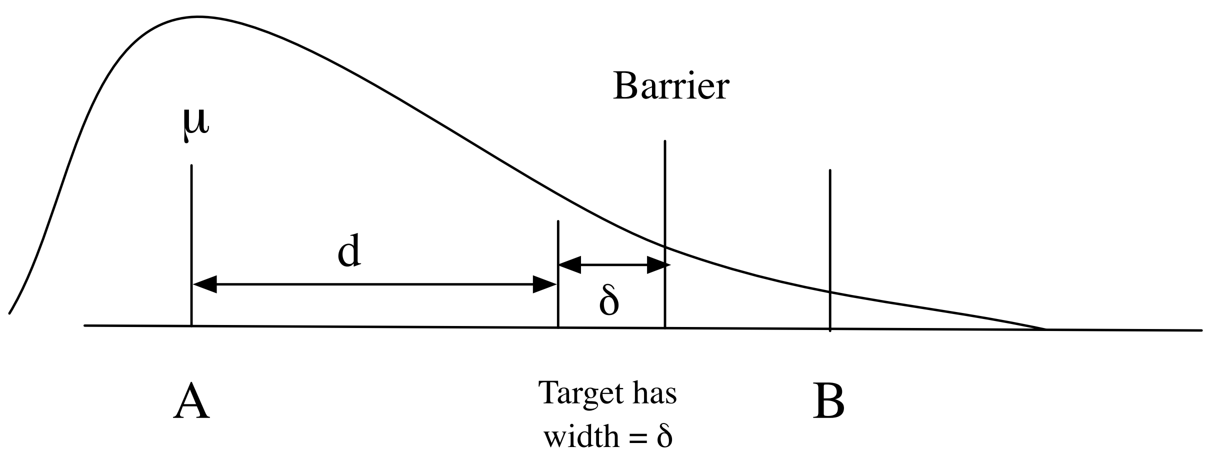
Figure 4: 1-Dimensional Model of Stochastic Search in a Feasible Region. $C(d)$: event of getting closer from distance d. Moving from A to B is a good thing.

Comments on the 1-Dimensional Model

- $P(C(d) = 1) = \int_{\mu}^{\mu+d} f(x)dx + \int_{d+\delta}^{d+\delta+d} f(x)dx$ for the FI-2Pop GA.
- With a death penalty: $P(C'(d) = 1) = \int_{\mu}^{\mu+d} f(x)dx$
- C stochastically dominates C'
- The point generalizes to populations of solutions, to distance noisily recognized by selection, to multiple dimensions. This is a simple model of an essential aspect of a GA for constrained optimization.

Observation #2: The One-Thing-at-a-Time Principle

- Recall Figure 4, page 53. Given an infeasible solution at B, how should it be directed? Answer: to minimize distance to the barrier. Objective function considerations are only likely to misdirect it. Bad news for penalty functions.



Comments

1. In the FI-2Pop GA selection of the two populations obeys the one-thing-at-time (aka: Swiss army knife) principle. Feasible solutions are directed (selected) purely to improve the objective function. Infeasible solutions are directed purely to minimize distance to the boundary (more than one way to do this). This is a <u>good thing</u>.

2. Dividing the total population in 2 is a form of hedging and a form of division of labor. In the FI-2Pop GA each population is specialized to a single task (improve on the objective or improve on feasibility).

3. Note that the degradation from division is less than half, since selection is proportional to the square root of the population size.

4. Recall requirement (b) an adjustable mechanism. If we learn where to put more resources, can we do something about it? M-PEPs, multi-population evolution programs, offer a natural response: enlarge or shrink the appropriate specialized populations.

5. Compare roles of the infeasible population early and late in a run. Early on, infeasible solutions are not descended from feasible solutions. They "explore the boundary" and when they diffuse over, provide variance in the feasible population, which may or may not be used.

6. Later on in the run, many or most of the infeasible solutions are descended from feasible ancestors. The GA may be thought of as engaging in *repair* on these solutions—without having to have domain-specific information!

Envisioning Principles for Design of Heuristics

1. Don't waste information.
2. Find ways to collect more information.
3. Find ways to employ more information.
4. Bias and tradeoffs are inevitable among competing goals; make them explicit; find ways to adjust based on information; find ways to hedge.
5. Specialize (division of labor, one thing at a time, Swiss army knife) in measuring competing goals.

Comment: Multiple, specialized populations afford both collection of information and a means to implement a response to having more information.

Envisioning Principles for Design of Heuristics: Examples

1. What are the competing values we wish to measure?

   Example 1: (a) objective function versus (b) feasibility.

   Example 2: (a) exploration versus (b) exploitation.

2. How (for example) will we measure these values?

   Example 1: (a) objective function value. (b) sum of constraint violations.

   Example 2: (a) mutation and crossover rates; population variance. (b) mutation and crossover rates; coefficient of selection.

3. What are our instruments for measuring these values?

   All examples: separate populations, specializing in each measurement.

4. How do we allocate resources among the instruments?

   All examples: by size of population.

5. How do we decide how to allocate resources based on the instruments?

   All examples: by reading the instruments (population characteristics) and learning how to make profitable adjustments.

   May use hedging and/or received information to decide where and how much.

6. How do we operate the instruments?

   All examples: by updating the populations, e.g., by GA, by hill climbing,. . . Some version of GA is natural, but many possibilities are interesting.

   Important issues: When do you move individuals between populations? (FI-2Pop: when they go feasible/infeasible; can be generalized). When do you combine individuals from different populations? (future research).

Future Research

A lot remains to be done! High on our list:

- Mating individuals from different populations. May be thought of as a form of interpolation. Very promising if the boundary lies between and the optimal is on or near it.
- Actually "turning the knobs", either from sampling information or from information gleaned during a run.
- Multiobjective optimization and goal programming
- Theoretical issues, e.g., why is approach to a boundary sometimes hard (easy) from the infeasible (feasible) side?
- Refinements, e.g., common scaling of constraint violations (see above?).

Finally, elevator speech

1. The FI-2Pop GA has shown considerable success and holds much promise for handling difficult constrained optimization problems. It is (uniquely?) capable in supplying candle-lighting information.
2. Experience with the FI-2Pop GA for constrained optimization teaches (at least suggests) important lessons regarding the design of metaheuristics. Among them:
   - Swiss army knife principle: handling one goal at a time may afford more efficient use of information as well as the capacity to exploit information regarding the profitability of various tradeoffs.
3. The GA in particular, but really population-based metaheuristics generally, cohere well with these findings.