Online Population Size Adjusting Using noise and Sub-Structural Measurements

Tian-Li Yu, Kumara Sastry, David E. Goldberg

Illinois Genetic Algorithms Laboratory
Department of General Engineering
University of Illinois at Urbana-Champaign
http://www-illigal.ge.uiuc.edu
ksastry@uiuc.edu

Supported by **AFOSR** F49620-03-1-0129, **NSF/DMR at MCC** DMR-99-76550, NSF/ITR at CPSD DMR-0121695, DOE at FS-MRL DEFG02-91ER45439.

Motivation

- Facetwise population-sizing models
  - Adequate supply of sub-structures [Goldberg *et al*, 2001]
  - Good decisions between competing substructures [Goldberg *et al*, 1992, Harik *et al*, 1997]
  - Accurate sub-structure identification [Pelikan *et al*, 2003]
- Linkage (sub-structure) learning methods
  - Implicit (*e.g.*, LLGA, BOA)
  - Explicit (*e.g.*, eCGA, DSMGA)
- Utilize sub-structure information in estimating population size adaptively

Outline

- Background
- Population-sizing models
- Design structure matrix GA (DSMGA)
- Population-size estimation
- Population-size adjustment
- Summary & conclusions

Related Work

- Smith and Smuda (1993)
  - Target selection loss $\widehat{L}$
  - Adjust the population size by the factor of $\frac{\widehat{L}}{L_t}$
- Harik and Lobo (1999)–Parameterless GA
  - eCGA (Lobo, 2000)
  - hBOA (Pelikan and Lin, 2004)

How Should I Size the Population Size?

1. Ensure an adequate supply of raw sub-solutions
   - Population size should be large enough to ensure at least one copy of all competing BBs in the population
2. Make decisions well among competing sub-solutions
   - Population size should be large enough to prefer the best BBs over worse ones with a high probability.
3. Identify and mix sub-solutions well
   - Population size should be large enough to ensure that the "innovation time" is less than "takeover or convergence time".

[Goldberg (2002). Design of Innovation]

Building-Block Supply

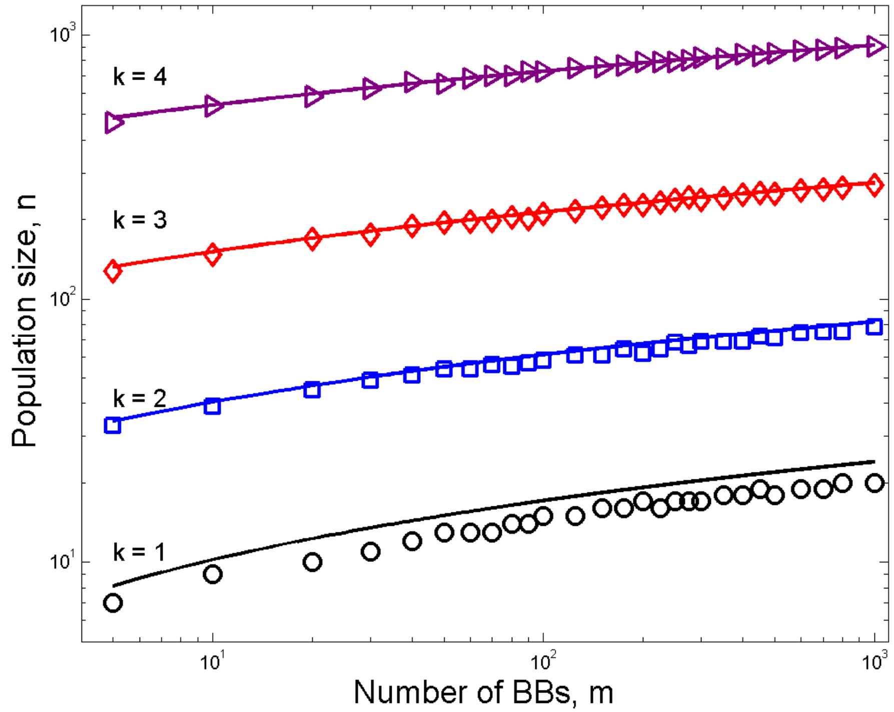
- Ensure the presence of competing schema
  - Temporally through genetic operators
  - Spatially (Initial population) [Holland, 1975; Goldberg, 1989; Reeves, 1993; Goldberg, Sastry, & Latoza, 2001; Sastry, O'Reilly, Goldberg, & Hill, 2003]

Sub-structure size

$$n \approx \chi^{k} \left(k \log \chi + \ln m \right)$$

competing sub-components

components



Gambler's Ruin Population Sizing Model

- Combines decision-making and supply models [Harik, Cantú-Paz, Goldberg, and Miller, 1997]

Signal between competing BBs

Fitness variance of a BB

Failure probability

Sub-Component size (BB size)

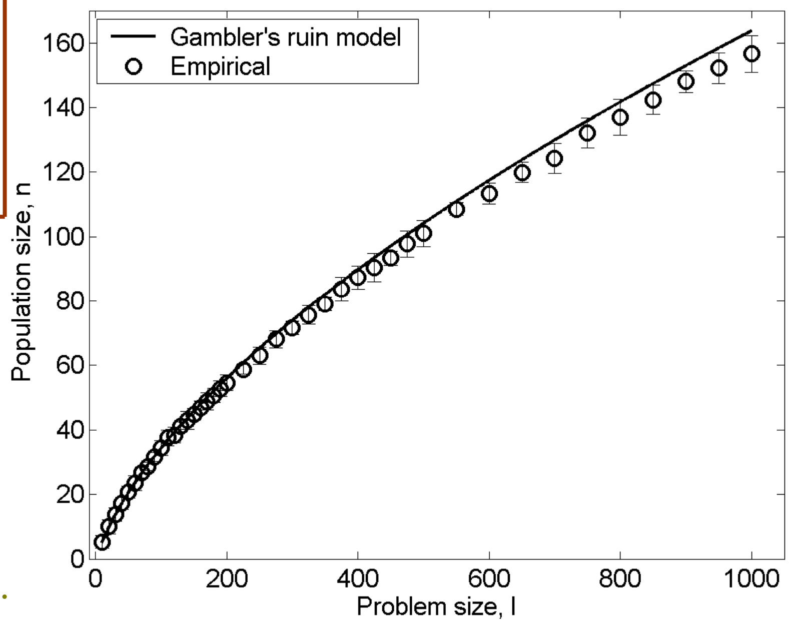
$$n = -\frac{\sqrt{\pi}}{2} \cdot \log(\alpha) \cdot \underbrace{\left(\frac{\sigma_{bb}}{d}\right)} \cdot 2^k \cdot \sqrt{m}$$

Error tolerance

Signal-to-Noise ratio

Competing sub-components

Components (# BBs)



Harik, Cantú-Paz, Goldberg, & Miller, *ECJ 7*(3) 1999.

Population Sizing For Sub-Structure Identification

❖ Estimation of distribution algorithms [Pelikan, Goldberg, and Cantú-Paz, 2001; Sastry & Goldberg, 2001; Pelikan, Sastry, and Goldberg, 2002]

Signal between competing BBs

Failure probability

Fitness variance of a BB

Sub-Component size (BB size)

$$n = -c_n \cdot \log(\alpha) \cdot \left(\frac{\sigma_{bb}^2}{d^2}\right) \cdot 2^k \cdot m$$

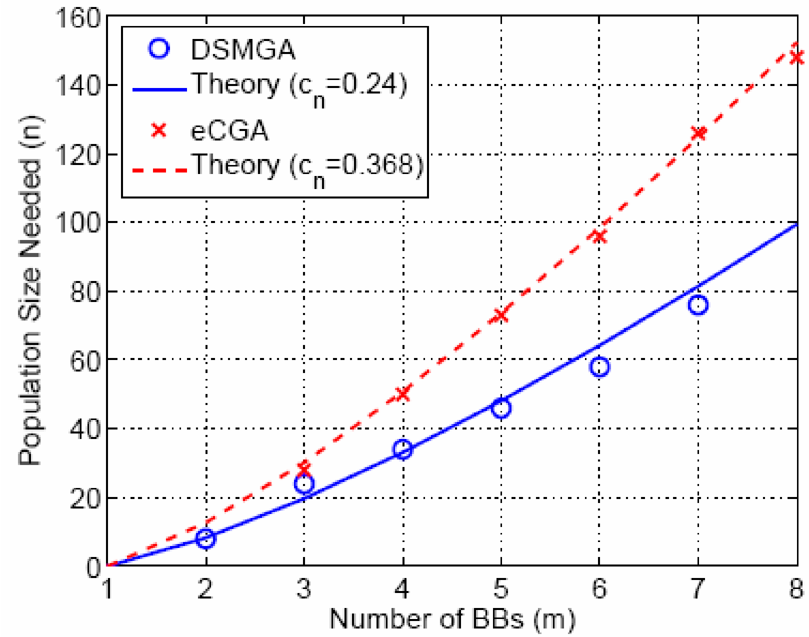
Error tolerance

Signal-to-Noise ratio

Competing sub-components

Components (# BBs)

Bounds DSMGA/EDA population sizing



Purpose: Online Population-Sizing Framework

- Develop an online population-sizing framework using sub-structural information
- Sub-structure identification
  - What are the components?
- Sub-structure quality estimation
  - What are the relative quality of sub-components?
- Sub-structure-based population sizing estimate
  - Ensure adequate supply of good sub-structures?
  - Ensure growth of good sub-structures?

Sub-Structure Identification: Use Linkage-Learning Techniques

- Sub-structure identification
  - What are the components?
- Use linkage-learning techniques
  - Demonstrate using DSMGA [Yu *et al*, 2003]
  - Can also use extended compact GA [Harik, 1999]
- Brief description of DSMGA next…

Design structure matrix (DSM)

- GM powertrain development teams [McCord & Eppinger, 1993]

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
Engine Block	A	A	3	3	3	1	3	1	3	3	3	1	1	1						3		3	2
Cylinder Heads	B	3	B	3	2		1		3	1	3	3								2		1	3
Camshaft/Valve Train	C	3	3	C	1		1			2	1	1								1		1	2
Pistons	D	3	3	2	D	2	2	1		2	2	1					2			1			3
Connecting Rods	E	2	1		3	E	3			2	1												2
Crankshaft	F	3	1	1	3	3	F	3	2	3	1									2	1		2
Flywheel	G	1					3	G		1												2	2
Accessory Drive	H	3	3	1			2		H	1	3	3	2	2	2	3	3	3	1	1	1	1	2
Lubrication	I	3	1	2	2	1	2	1	1	I	1		1							1		2	3
Water Pump/Cooling	J	3	3	2	2				3	2	J	2		1	1		1	1				1	2
Intake Manifold	K	2	3	1					3	1	3	K								2	2		3
Exhaust	L	1							1	1	1		L	3	1	3	2			2	2	1	2
E.G.R.	M	1							1		1		3	M		1	1	1		1	3	1	2
Air Cleaner	N								3				1		N	2	1	3					
A.I.R.	O	1							3		1		3	1	2	O		1			2	1	2
Fuel System	P								2		1			1	1	1	P	1	1			2	2
Throttle Body	Q								2		2			2	3	1	2	Q	3	1	3		2
EVAP	R																2	3	R		2	1	
Ignition	S	3	3	3	1		3	3	2	1	1	2	3	1			3	1		S	3	3	3
E.C.M.	T	1	1	1			2	1	1	1	3	2	1	2		2	3	2	3	3	T	3	2
Electrical System	U	3	1	2	1		1	2	1	2	1	1	3	1		1	2		1	3	3	U	3
Engine Assembly	V	3	3	2	3	2	3	2	2	3	2	3	2	2		2	3	2	1	3	2	3	V

0: No interaction

1: low interaction

2: med interaction

3: high interaction

DSM clustering

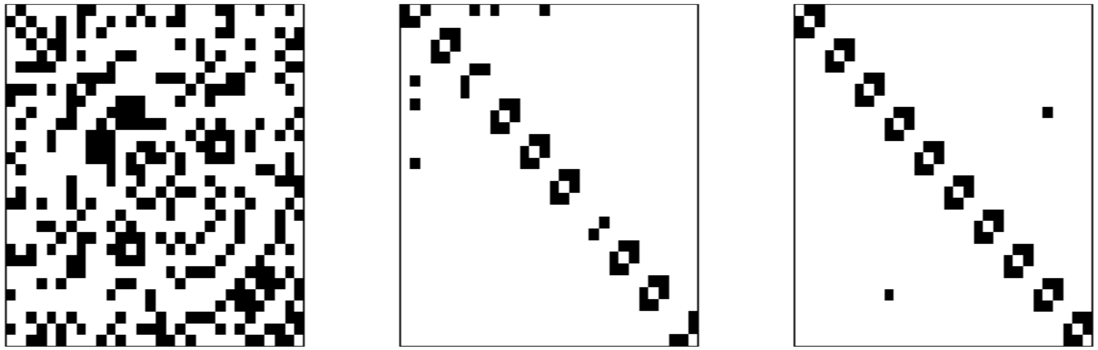
		P	N	Q	R	B	C	J	K	A	D	I	F	G	E	O	L	M	H	S	T	U	V
Fuel System	P	P	1	1	1			1								1		1	2			2	2
Air Cleaner	N	1	N	3												2	1		3				
Throttle Body	Q	2	3	Q	3			2								1		2	2	1	3		2
EVAP	R	2		3	R																2	1	
Cylinder Heads	B					B	3	3	3	3	2	1	1						3	2		1	3
Camshaft/Valve Train	C					3	C	1	1	3	1	2	1							1		1	2
Water Pump/Cooling	J	1	1	1		3	2	J	2	3	2	2						1	3			1	2
Intake Manifold	K					3	1	3	K	2		1							3	2	2		3
Engine Block	A					3	3	3	1	A	3	3	3	1	1		1	1	3	3		3	2
Pistons	D	2				3	2	2	1	3	D	2	2	1	2					1			3
Lubrication	I					1	2	1		3	2	I	2	1	1		1		1	1		2	3
Crankshaft	F					1	1	1		3	3	3	F	3	3				2	2	1		2
Flywheel	G									1		1	3	G								2	2
Connecting Rods	E					1		1		2	3	2	3		E								2
A.I.R.	O		2	1				1		1						O	3	1	3		2	1	2
Exhaust	L	2	1					1		1		1				3	L	3	1	2	2	1	2
E.G.R.	M	1		1				1		1						1	3	M	1	1	3	1	2
Accessory Drive	H	3	2	3	1	3	1	3	3	3		1	2			3	2	2	H	1	1	1	2
Ignition	S	3		1		3	3	1	2	3	1	1	3	3			3	1	2	S	3	3	3
E.C.M.	T	3		2	3	1	1	3	2	1		1	2	1		2	1	2	1	3	T	3	2
Electrical System	U	2			1	1	2	1	1	3	1	2	1	2		1	3	1	1	3	3	U	3
Engine Assembly	V	3		2	1	3	2	2	3	3	3	3	3	2	2	2	2	2	2	3	2	3	V

Reordered

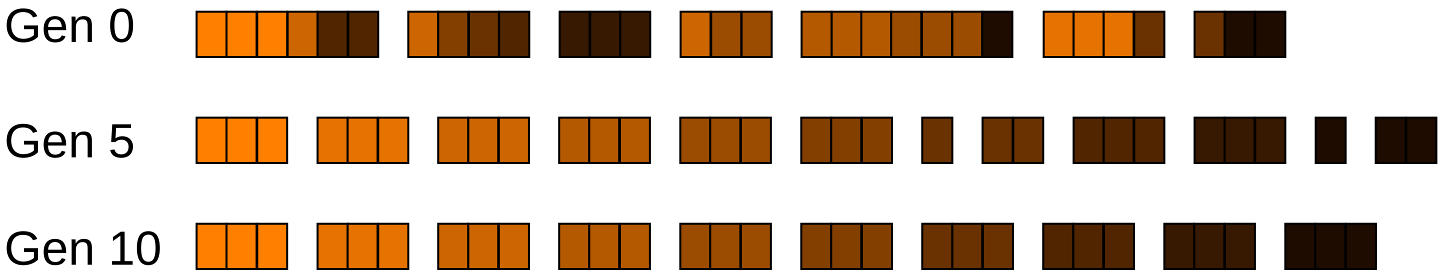
Clusters

Bus modular

DSMGA



Generation 0 Generation 5 Generation 10



10 3-bit Trap Function

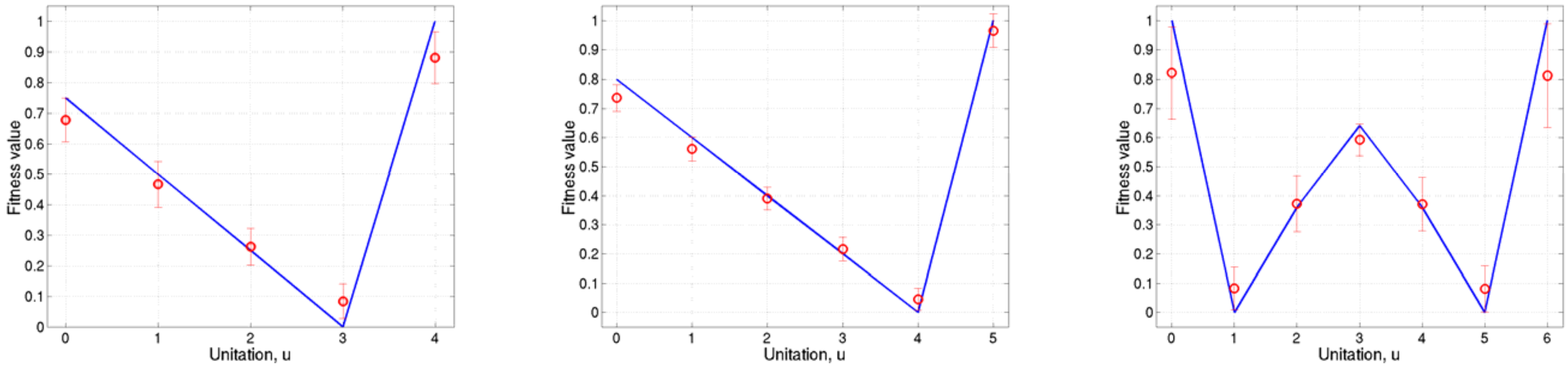
Online Adjustment of Population Size

- Sub-structure identification
  - Use linkage-learning techniques
    - E.g., DSMGA, eCGA, BOA
- Sub-structure quality estimation
  - What are the relative quality of sub-components?
- Use fitness-estimation technique used in developing probabilistic endogenous fitness model
  - Pelikan *et al* 2004, Sastry *et al* 2004

Sub-Structure Quality Estimation

- Identify key sub-structures of the search problem
- Estimate the fitness of sub-structure instances

Fitness of substructure instance	=	Average fitness of all evaluated individuals with the schema instance	−	Average fitness of all evaluated individuals in the population



- Used to estimate fitness of individuals [Sastry, Pelikan, & Goldberg, 2004. Pelikan, Sastry & Goldberg, 2004]

Online Adjustment of Population Size

- Sub-structure identification
  - Use linkage-learning techniques
- Sub-structure quality estimation
  - Schema fitness estimation
- Population-size estimation based on sub-structure info
  - Bounded by sub-structure identification
  - Estimate all parameters on-line using sub-structural measurements

Online Population Size Estimation

$$n = -c_n \cdot \log(\alpha) \cdot \underbrace{\left(\frac{\sigma_{bb}^2}{d^2}\right)} \cdot \sum_{i=1}^{m} 2^{k_i}$$

Signal between competing BBs

Failure probability

Fitness variance of a BB

Components (# BBs)

Sub-Component size (BB size)

Error tolerance

Signal-to-Noise ratio

Competing sub-components

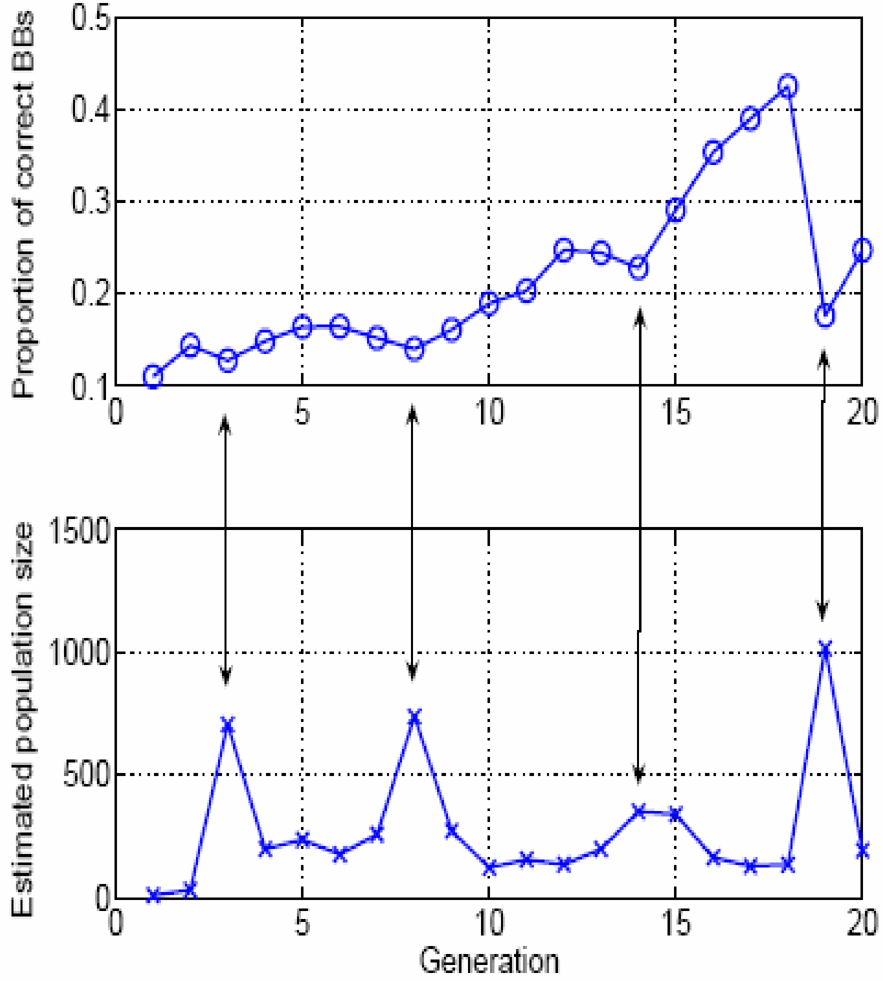
- Sub-structure models gives *m* and k_i
- H_{BB_i}: Schema of the i^th sub-structure

$$\sigma_{BB_i}^2 = \mathrm{Var}_i[f(H_{BB_i})]$$

$$d_i = \max_i[f(H_{BB_i})] - \mathrm{second\_max}_i[f(H_{BB_i})]$$

How To Inject New Individuals?

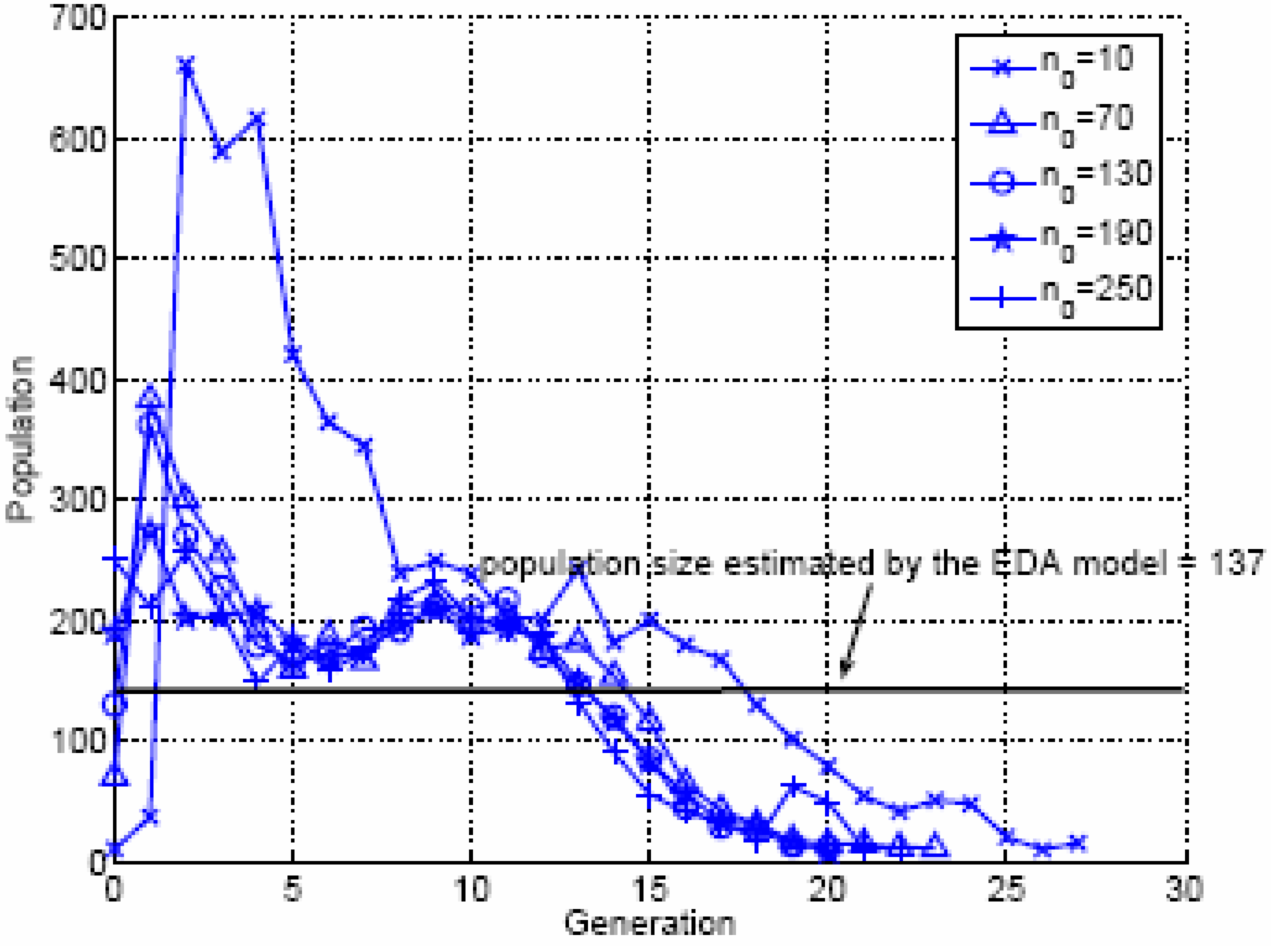
- Aim: Increase the success rate of small initial population size
- Inject new individuals to ensure
  - BB supply
  - Diversity
- Issues to be addressed
  - Average fitness decreases
  - Convergence difficulty



Injection scheme:
Initialize Early and Recombine Later

- Observations
  - Injection of new individuals → BB supply → early stage of GA run
  - Recombination → BB mixing → later stage of GA run
- Injection scheme, when $n_{t+1} > n_t$
  - Early stage: Generate n_t children by recombination, and inject (n_{t+1} - n_t) new individuals.
  - Later stage: Simply generate n_{t+1} children.
- When to switch?
  - The population-size estimation indicates large population is not needed.

Population-size adjustment on-the-fly

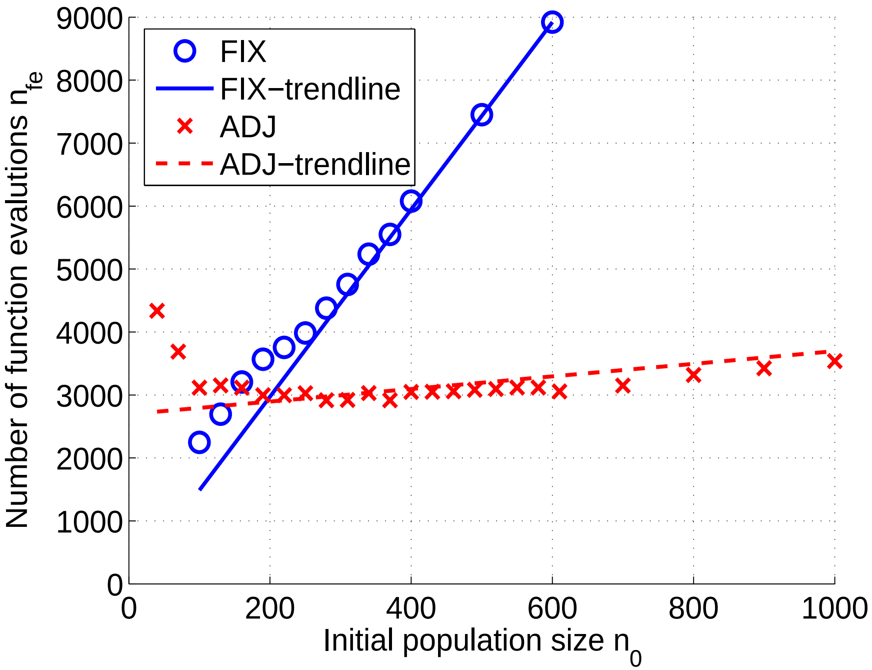


Similar behavior for different initial population sizes (except at extremely small initial population size).

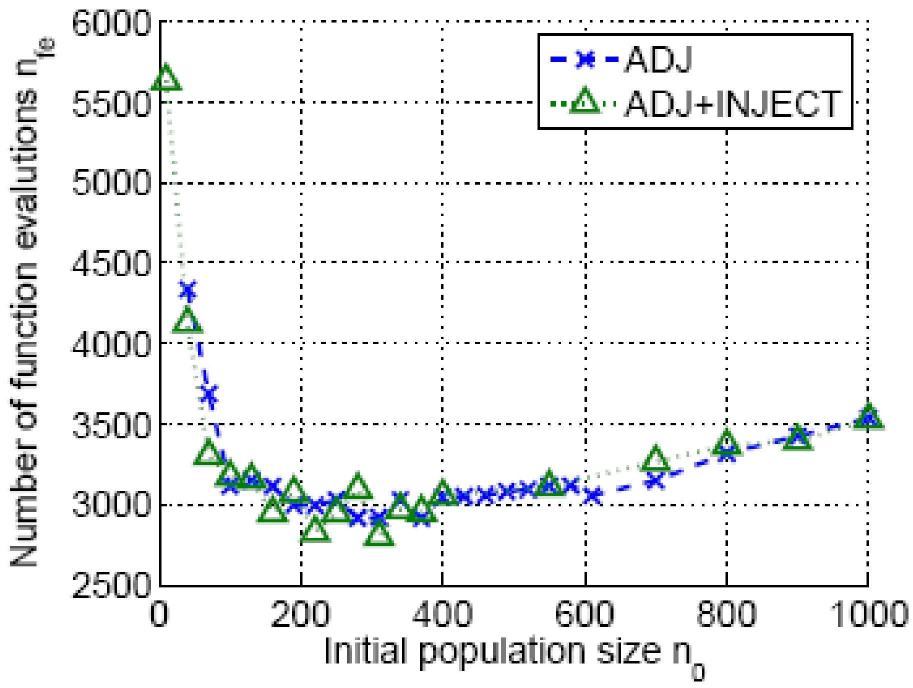
Efficiency:
Adaptation Doesn’t Require Significantly Extra Evaluations

- Large n_0 does not cost too much.

Generate new individuals by recombination



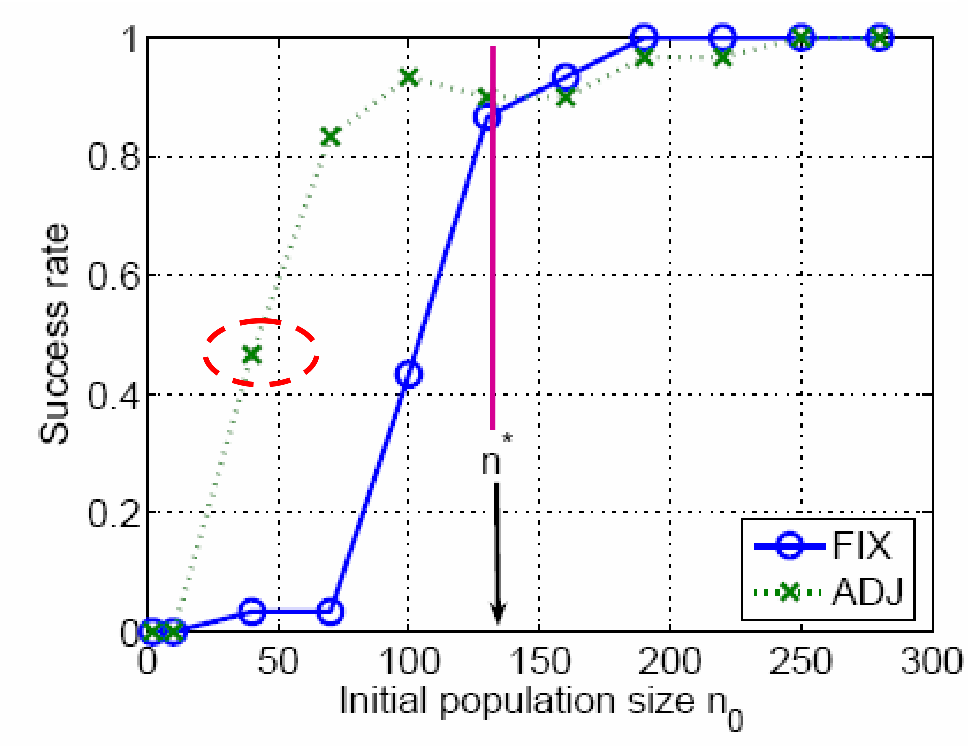
Initialize early recombine later

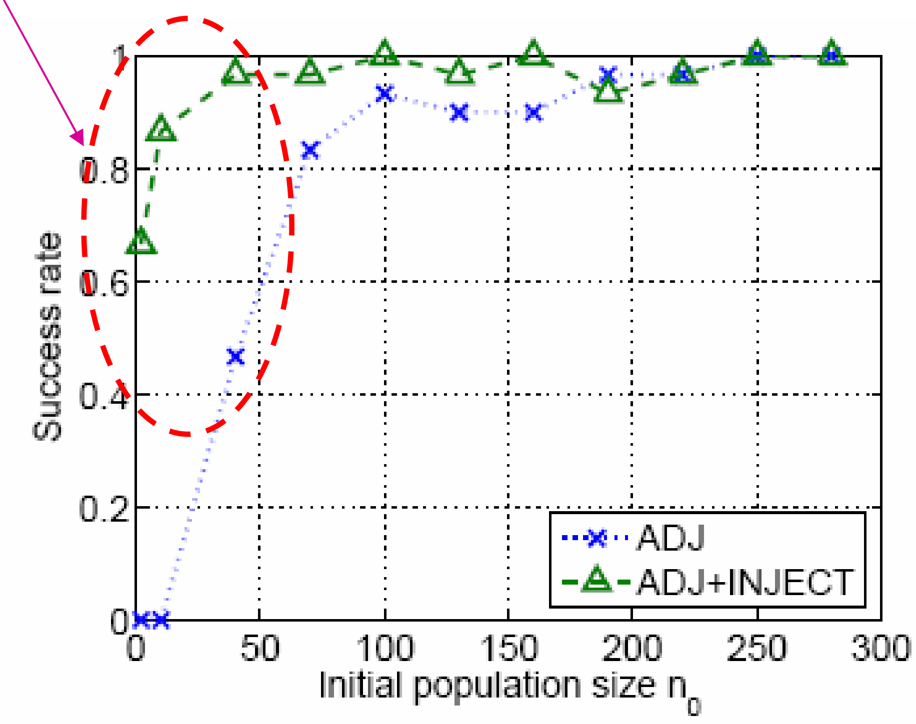


Robustness:
Small Initial Population Sizes Yield Success

- Small n_0 does not easily cause GA failure.

Recovered from small initial population sizes.





What's next?

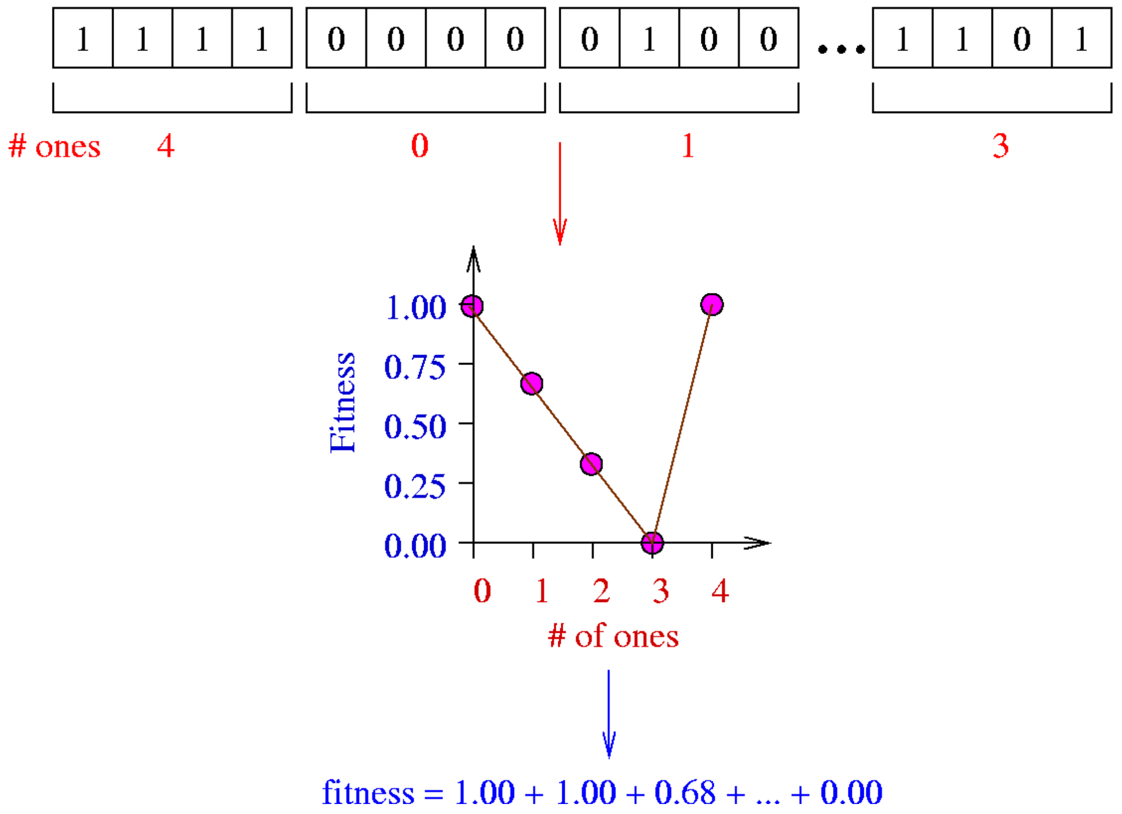
- Other linkage learning GEAs
- Update rule
  - $n_{t+1} = \alpha n_t + (1-\alpha)\hat{n}_{t+1}$
  - Pros: The change of the estimated population size are smoother. A more noise-robust scheme.
  - Cons: Another parameter to tune.
- Optimal switch time
  - When do we have high enough confidence to switch from the individual injection scheme to the regular recombination?
- Real-world problems

Initial guess

- Not seriously sensitive
- Prior knowledge (k=4~8, an antenna problem)
- Doubling scheme
  - Start from small initial population
  - Keep doubling it until the parameter (*m,k,…*) does not vary much
  - Use the final result as the initial guess of population size
  - Worst case: 4 times of what is needed, but only for the first generation.

Test Problem: Modified m-k Trap Functions

- Linkage learning is critical to GA success
- Massively multimodal: 2^m distinct optima



[Ackley, 1987; Goldberg, 1987; Deb & Goldberg, 1993]