On The Supply of Building Blocks

David E. Goldberg
Kumara Sastry
Thomas Latoza
Illinois Genetic Algorithms Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801
http://www-illigal.ge.uiuc.edu

Genetic and Evolutionary Computation Conference
(GECCO-2001)
July 7-11, 2001
San Francisco, CA

Overview

- GA design decomposition
- Building block supply
- Objective
- Supply of a single BB
- Supply for partition success
- Population-sizing for BB supply
- Summary

GA Design Decomposition

- Goldberg, Deb, & Clark, 1992
  1. Know what GAs are processing—building blocks (BBs)
  2. Know thy BB challengers—BB-wise difficult problems.
  3. Ensure an adequate supply of raw BBs
  4. Ensure increased market share for superior BBs
  5. Know BB takeover and convergence time
  6. Make decisions well among competing BBs
  7. Mix BBs well

Building Block Supply

- Previous work:
  - Holland, 1975; Goldberg, 1989; Reeves, 1993.
- Spatial & temporal approach
- Selectorecombinative GAs
- BB growth usually supersedes supply
- Supply sometimes governs population size
- Complex models exist (Harik et al, 1997)
- Better models are needed

Objective

- Develop facetwise models
  - Supply of a single BB
  - Supply of all schemata in a partition
- Population sizing for BB supply
- Verify models with empirical results

Supply Of A Single Building Block

- consider a string:
  - Of length ℓ
  - Containing alphabets of cardinality χ
  - Has schemata of order k
- n randomly generated strings
- At least one string has the desired schema

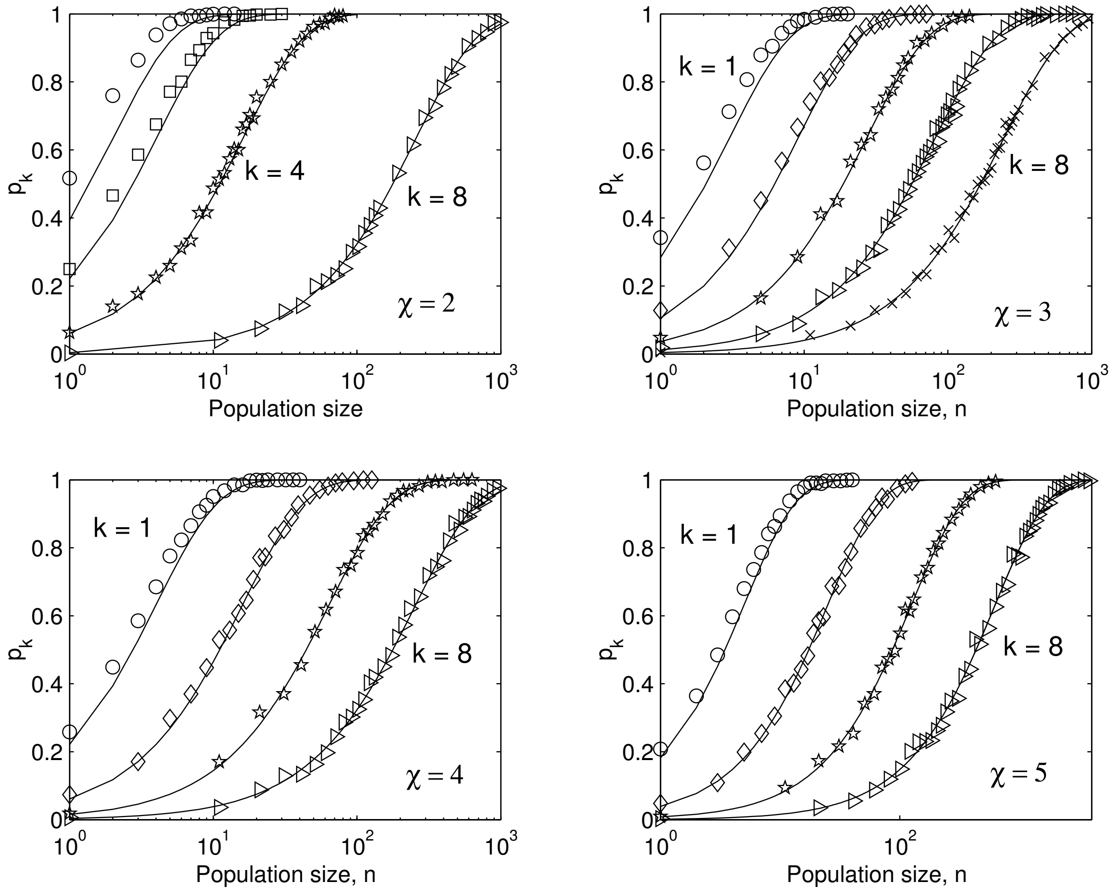
Single Building Block Supply

- Total number of schemata: χ^k
- Desired schema is not in a string: $1 - \frac{1}{\chi^k}$
- It is not in any of n strings: $\left[1 - \frac{1}{\chi^k}\right]^n$
- At least one string has the desired schema

$$p_k = 1 - \left[1 - \left(\frac{1}{\chi^k}\right)\right]^n \approx 1 - \exp\left(-\frac{n}{\chi^k}\right)$$

- $n > \chi^k$: $p_k \to 1$

Results: Supply Of A Single BB



Supply For Partition Success

- Earlier we considered only one schema
- At least one copy of all schemata
  - Don't know which schema is good
  - Want a good share of all of them
- Exact model is complex
- Approximations
  - $n > n_k$
  - BB successes are independent

Exact Model

- Consider not having at least one schema
- Illustration through a test case
  - Unit length BB, $k = 1$
  - Tertiary alphabet, $\chi = 3$
- Generalize the result

Illinois Genetic Algorithms Laboratory
Department of General Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801. USA.
http://www-illigal.ge.uiuc.edu

Partition Success: $k = 1, \chi = 3$

- h_1: **0**, h_2: **1**, h_3: **2**
- Absence of at least one schema
  - (all h_1) OR (all h_2) OR (all h_3)
  - (h_1 or h_2) OR (h_1 or h_3) OR (h_2 or h_3)
- Total ways of absence: $3 \cdot 2^n - 3$
- Total possible ways: $\chi^{kn} = 3^n$
- Partition success

$$p_s = 1 - \frac{2^n - 1}{3^{n-1}}$$

Exact Model

- $\chi = 2$: $p_s = 1 - \frac{1}{2^{n-1}}$
- $\chi = 3$: $p_s = 1 - \frac{2^n - 1}{3^{n-1}}$
- $\chi = 4$: $p_s = 1 - \frac{3^n - 3 \cdot 2^{n-1} + 1}{4^{n-1}}$
- Generalize the result

$$p_s = 1 - \frac{1}{\chi^{nk}} \left[\sum_{i=1}^{\chi^k - 1} (-1)^{i-1} \binom{\chi^k}{i} \left(\chi^k - 1\right)^n \right]$$

Approximation Of Exact Model

- Consider the exact model

$$p_s = 1 - \frac{1}{\chi^{nk}}\left[\sum_{i=1}^{\chi^k-1}(-1)^{i-1}\left(\begin{array}{c}\chi^k\\ i\end{array}\right)\left(\chi^k-1\right)^n\right]$$

- Assume $n > \chi^k$: $(\chi^k - 1)^n > (\chi^k - 2)^n > \cdots > 1$
- Consider only the first term

$$\begin{aligned} p_s &= 1 - \chi^k\left[1 - \left(\frac{1}{\chi^k}\right)^n\right] \\ p_s &\approx 1 - \chi^k \exp\left(-\frac{n}{\chi^k}\right) \end{aligned}$$

Approximate Model

- Schema partition success is independent
- Recall, single BB success probability:

$$p_k = 1 - \chi^k \exp\left(-\frac{n}{\chi^k}\right)$$

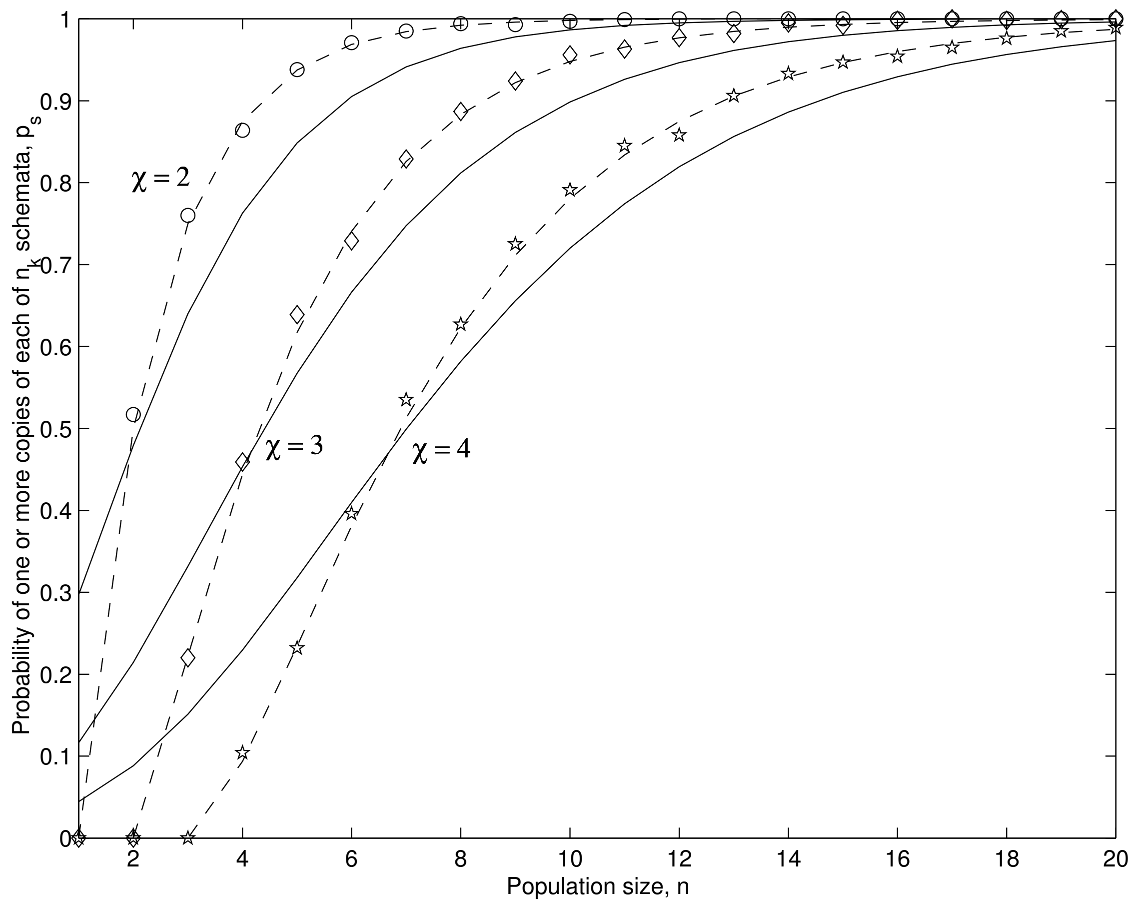
- Using independence assumption

$$p_s = \left[1 - \chi^k \exp\left(-\frac{n}{\chi^k}\right)\right]^{\chi^k}$$

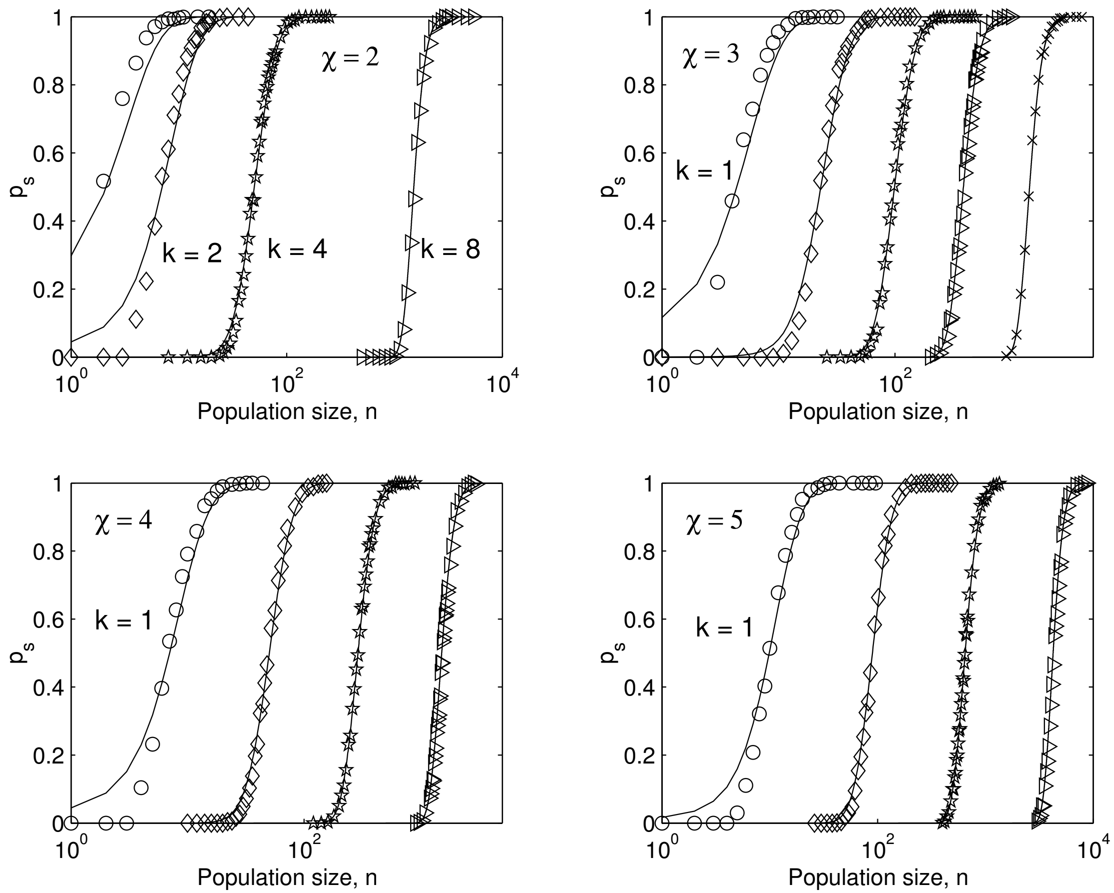
- $(1 - r/n)^n \approx e^{-r}$. $n > \chi^k$, $(1 - x)^n \approx 1 - nx$

$$p_s \approx 1 - \chi^k \exp\left(-\frac{n}{\chi^k}\right)$$

Exact Model vs Approximate Model



Results: Partition Success



Population Sizing For BB Supply

- Require a population size
  - All schemata are present
  - Success probability asymptotically reaches 1
  - Tolerate small failure probability α
  - Partition success, $p_s = 1 - \chi^k \exp\left(-\frac{n}{\chi^k}\right)$
  - Equate $p_s = 1 - \alpha$

$$n = \chi^k \left(k \log \chi - \log \alpha\right)$$

Population-Sizing Model

- m building blocks of length k
- At most one BB need not be present
- Failure probability, $\alpha = \frac{1}{m}$

$$n = \chi^k \left(k \log \chi - \log \alpha\right)$$

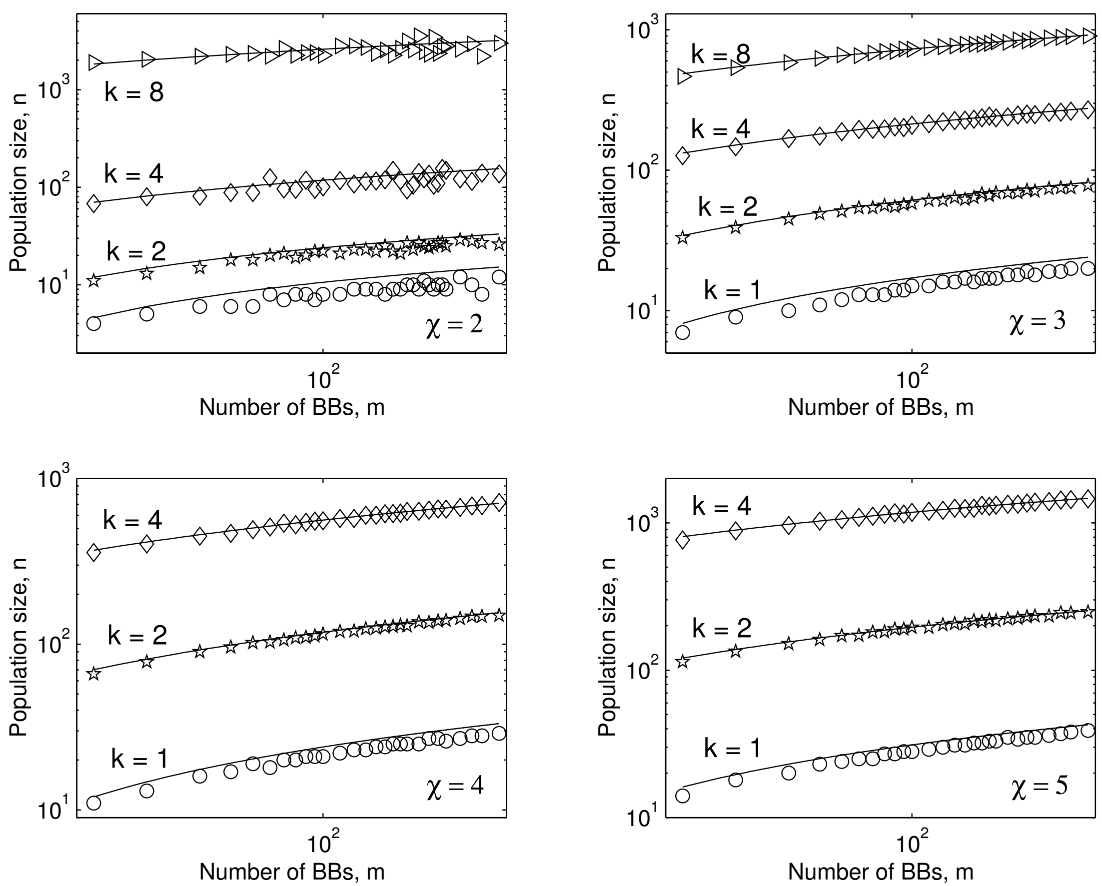
- Large problem: $m \gg \chi^k$

$$n = \mathcal{O}\left(\chi^k \log m\right)$$

- Complex problem: $\chi^k \gg m$

$$n = \mathcal{O}\left(\chi^k \log m\right)$$

Results: Population Sizing



Summary

- Facetwise models
  - Supply of a single schema
  - Supply of all raw BBs
  - Population sizing for BB supply
- Two asymptotic cases
  - Large problem: $\mathcal{O}\left(\chi^k \log m\right)$
  - Complex problem: $n = \mathcal{O}\left(\chi^k \log m\right)$

Acknowledgments

- Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant F49620-00-1-0163.
- National Science Foundation under grant DMI-9908252.
- U. S. Army Research Laboratory under the Federated Laboratory Program, Cooperative Agreement DAAL01-96-2-0003.