

Genetic Programming

Riccardo Poli
Department of Computer Science
University of Essex

June 2004

R. Poli - University of Essex

2



AND

Videos from Koza's GP4 DVD

June 2004

R. Poli - University of Essex

3

Overview: Part I

- Introduction
- Basics
- Examples
- More advanced techniques
- Applications

June 2004

R. Poli - University of Essex

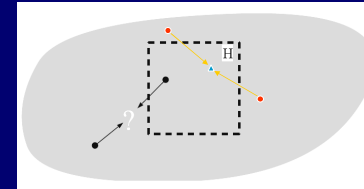
4

Overview: Part II

- Theory in theory
- GA schema theory
- GP schema theory
- Lessons and implications
- Conclusions

- $m(H,t)$ is number of individuals in the schema H at generation t ,
- M is the population size,
- $p(H,t)$ is the selection probability for strings in H at generation t ,
- p_m is the mutation probability,
- $O(H)$ is the schema order, i.e. number of defining bits,
- p_c is the crossover probability,
- $L(H)$ is the defining length, i.e. distance between the furthest defining bits in H ,
- N is the bitstring length.

□ Idea:



□ Features:

- The theorem includes an *expected value*
- It provides a *lower bound*
- So, it is *difficult to make accurate predictions*

- The factor σ differs in the different formulation of the schema theorem:
 - $\sigma = 1 - m(H,t)/M$ in (Holland, 1975),
 - $\sigma = 1$ in (Goldberg, 1989),
 - $\sigma = 1 - p(H,t)$ in (Whitley, 1994).
- In 1997 **Stephens** and collaborators produced an exact formulation for $\alpha(H,t)$: an *“exact” schema theorem*.

How can we get an exact schema theorem?

- Let us assume that only **reproduction** and (one-offspring) **crossover** are performed.
- Because these two **operators are mutually exclusive**, for a generic schema H we have:

$$\alpha(H,t) = \Pr[\text{An individual in } H \text{ is obtained via reproduction}] + \Pr[\text{An offspring matching } H \text{ is produced by crossover}]$$

- Reproduction is performed with probability p_r and crossover with probability p_c (with $p_r + p_c = 1$), so

$$\alpha(H, t) = p_r \times \Pr[\text{An individual in } H \text{ is selected for cloning}] + p_c \times \Pr[\text{The parents and the crossover points are such that the offspring matches } H]$$

where $\Pr[\text{Selecting an individual in } H \text{ for cloning}] = p(H, t)$

- The process of crossover point selection is **independent from the actual primitives** in a parent.
- The probability of choosing a particular crossover point **depends only on the actual size** of the parent.
- E.g., the probability of choosing any crossover point in

1 1 0 1 0 1

is identical to the probability of choosing any crossover point in

0 0 0 1 1 0

$$\Pr[\text{The parents and the crossover points are such that the offspring matches } H] = \sum_{\text{For all crossover points } i} \Pr[\text{The parents are such that if crossed over at position } i \text{ the offspring matches } H] = \sum_{\text{For all crossover points } i} \Pr[\text{Choosing crossover point } i] \times \Pr[\text{Selecting parents such that if crossed over at point } i \text{ produce an offspring in } H]$$

- Let us assume that crossover points are selected with uniform probability:

$$\Pr[\text{Choosing crossover point } i] = \frac{1}{\text{Number of bits} - 1}$$

Pr [Selecting parents such that if crossed over at point i produce an offspring in H]

=Pr [Selecting a first parent such that if crossed over at point i provides the necessary material to create an offspring in H]

×Pr [Selecting a second parent such that if crossed over at point i provides the remaining necessary material]

Stephens and Waelbroeck's Exact GA Schema Theory (1997)

- For a binary GA with one point crossover applied with probability p_{xo} (and assuming $p_m=0$)

$$E[m(H, t + 1)/M] = (1 - p_{xo})p(H, t) + \frac{p_{xo}}{N - 1} \sum_{i=1}^{N-1} p(L(H, i), t)p(R(H, i), t)$$

$\Pr(L(H, i), t) =$

Pr [Selecting a first parent such that if crossed over at point i provides the necessary material to create an offspring in H]

$\Pr(R(H, i), t) =$

Pr [Selecting a second parent such that if crossed over at point i provides the remaining necessary material]

- $L(H, i)$ is obtained by replacing the elements of H to the right of position i with “don't care” symbols
- $R(H, i)$ is obtained by replacing the elements of H to the left of position $i+1$ with “don't care” symbols

- For example, if $H=1*11$, then $L(H, 1)=1****$, $R(H, 1)=**111$, $L(H, 3)=1*1**$, $R(H, 3)=***11$.
- For the schema $*11$, the theorem gives:

$$E[m(*11, t)/M] = (1 - p_{xo})p(*11, t) + \frac{p_{xo}}{2}(p(***, t)p(*11, t) + p(*1*, t)p(**1, t)) = (1 - \frac{p_{xo}}{2})p(*11, t) + \frac{p_{xo}}{2}p(*1*, t)p(**1, t),$$

since $p(***, t)=1$.