

CSE/ECE 848

Introduction to

Evolutionary Computation

Module 2, Lecture 8, Part 2
The Schema Theorem

Erik D. Goodman, Executive Director
BEACON Center for the Study of Evolution in
Action
Professor, ECE, ME, and CSE

The Fundamental Theorem of Genetic Algorithms -- “The” Schema Theorem

Holland published in 1975; had taught it much earlier (by 1968, for example, when I started Ph.D. at UM)

It provides ***lower bound*** on change in sampling rate of a single schema from generation t to $t+1$.

We’ ll derive it in several steps, starting from the change caused by selection alone:

$$M(H, t + \textit{intermed}) = M(H, t) \frac{f(H, t)}{\bar{f}}$$

Schema Theorem Derivation (cont.)

Now we want to add the effect of crossover:

A fraction p_c of pop undergoes crossover, so:

$$M(H, t+1) = (1 - p_c)M(H, t) \frac{f(H, t)}{\bar{f}} + p_c [M(H, t) \frac{f(H, t)}{\bar{f}} (1 - \text{losses}) + \text{gains}]$$

Will make a conservative assumption that crossover within the defining length of H is always disruptive to H , and will ignore gains (we're after a LOWER bound -- won't be as tight, but simpler). Then:

$$M(H, t+1) \geq (1 - p_c)M(H, t) \frac{f(H, t)}{\bar{f}} + p_c [M(H, t) \frac{f(H, t)}{\bar{f}} (1 - \text{disruptions})]$$

Reminder about Linkage and Defining Length

- Linkage -- “coadapted alleles”
- Argument that probability of disruption of schema H of length $\Delta(H)$ by one-point crossover is:

$$\Delta(H)/(L-1)$$

(= possible places to *disrupt* schema / all possible places to do one-pt crossover)

- Example:

11***** and 1*****1

$1/(L-1)$

1

Schema Theorem Derivation (cont.)

Whitley considers one *non*-disruption case that Holland didn't consider originally:

If cross H with an instance of itself, anywhere, get no disruption.

Chance of doing that, drawing second parent at random, is $P(H,t) = M(H,t)/\text{popsize}$: so prob. of disruption by x-over is:

$$\frac{\Delta(H)}{L-1} (1 - P(H,t))$$

Then can simplify the inequality, dividing by popsize and rearranging re p_c :

$$P(H,t+1) \geq P(H,t) \frac{f(H,t)}{\bar{f}} [1 - p_c \frac{\Delta(H)}{L-1} (1 - P(H,t))]$$

This version ignores mutation and assumes second parent is chosen at random. But it's usable, already!

Schema Theorem Derivation (cont.)

Now, let's recognize that we'll choose the second parent for crossover based on fitness, too:

$$P(H, t + 1) \geq P(H, t) \frac{f(H, t)}{\bar{f}} \left[1 - p_c \frac{\Delta(H)}{L-1} \left(1 - P(H, t) \frac{f(H, t)}{\bar{f}} \right) \right]$$

Now, let's add the effects of mutation. What is the probability that a mutation affects schema H?

(Assuming mutation always flips a bit or changes allele):

Each fixed bit of schema ($o(H)$ of them) changes with probability p_m , so they ALL stay UNCHANGED with probability:

$$(1 - p_m)^{o(H)}$$

Schema Theorem Derivation (cont.)

Now we have a more comprehensive schema theorem:

$$P(H, t+1) \geq P(H, t) \frac{f(H, t)}{f} \left[1 - p_c \frac{\Delta(H)}{L-1} \left(1 - P(H, t) \frac{f(H, t)}{f} \right) \right] (1 - p_m)^{o(H)}$$

(This is where Whitley stops. This is useful...but...)

Holland earlier generated a simpler, but less accurate bound, neglecting self-crosses and approximating the mutation loss factor by: (if $p_m = .01$, $o(H) = 3$, what's probability of NO loss of H from mutation?)

ABOUT: $(1 - o(H)p_m)$, assuming $p_m \ll 1$.

Schema Theorem Derivation (cont.)

That yields:

$$P(H, t+1) \geq P(H, t) \frac{f(H, t)}{\bar{f}} [1 - p_c \frac{\Delta(H)}{L-1}] [1 - o(H) p_m]$$

But, since $p_m \ll 1$, we can ignore small cross-product terms and get:

$$P(H, t+1) \geq P(H, t) \frac{f(H, t)}{\bar{f}} [1 - p_c \frac{\Delta(H)}{L-1} - o(H) p_m]$$

That is what many people recognize as the “classical” form of the schema theorem.

In Module 2, Lecture 8, Part 3, we’ll talk about what the schema theorem helps us to understand