Theoretical Analysis of Schema Processing

S = **1**0*

Representatives of schema S (also called "instances")

11**1**00**0**0

11**1**11**0**1

10**1**01**0**1

01**1**01**0**1

00**1**00**0**1

How many representatives are there?

$$2^{L-o(S)}$$

How many **different schemas** can the string 1111101 represent?

$$2^L$$

A population of N strings thus represents anywhere from 2^L to $N \times 2^L$ distinct schemas

(when all N strings happen to be exactly the same)

*****01

10*

****

1111101

11111**

*1111*1

Some Empirical Results

Population of 100 random genomes

Genome length	***Minimum # of schemas represented***	***Actual # of schemas represented***	***Maximum # of schemas represented***
3	8	27	800
4	16	81	1600
5	32	243	3200
6	64	688	6400
7	128	2021	12800
8	256	5391	25600
9	512	14571	51200
10	1024	36033	102400
11	2048	87314	204800
12	4096	201886	409600
13	8192	463363	819200
14	16384	1049650	1638400
15	32768	2306752	3276800
16	65536	4965857	6553600
17	131072	10324744	13107200
18	262144	21702967	26214400
19	524288	45460355	52428800
20	1048576	93676499	104857600

Probability and Expected Values

A **probability** is a number between 0 and 1

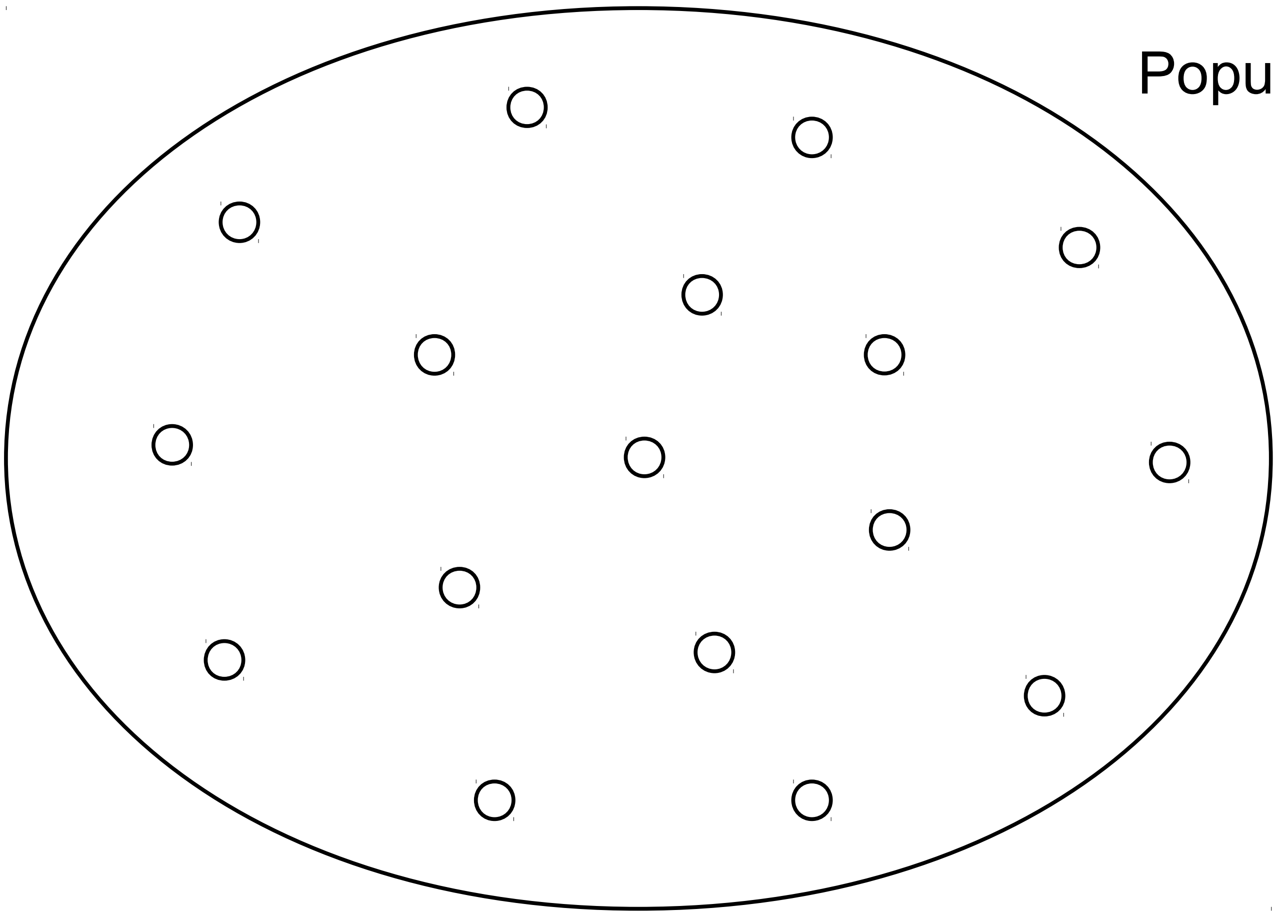
Think of it as a **proportion**

If each individual in a group of size N satisfies some condition with probability p, then the total # of individuals in the group that will end up satisfying the condition is **expected** to be about

$$p \times N$$

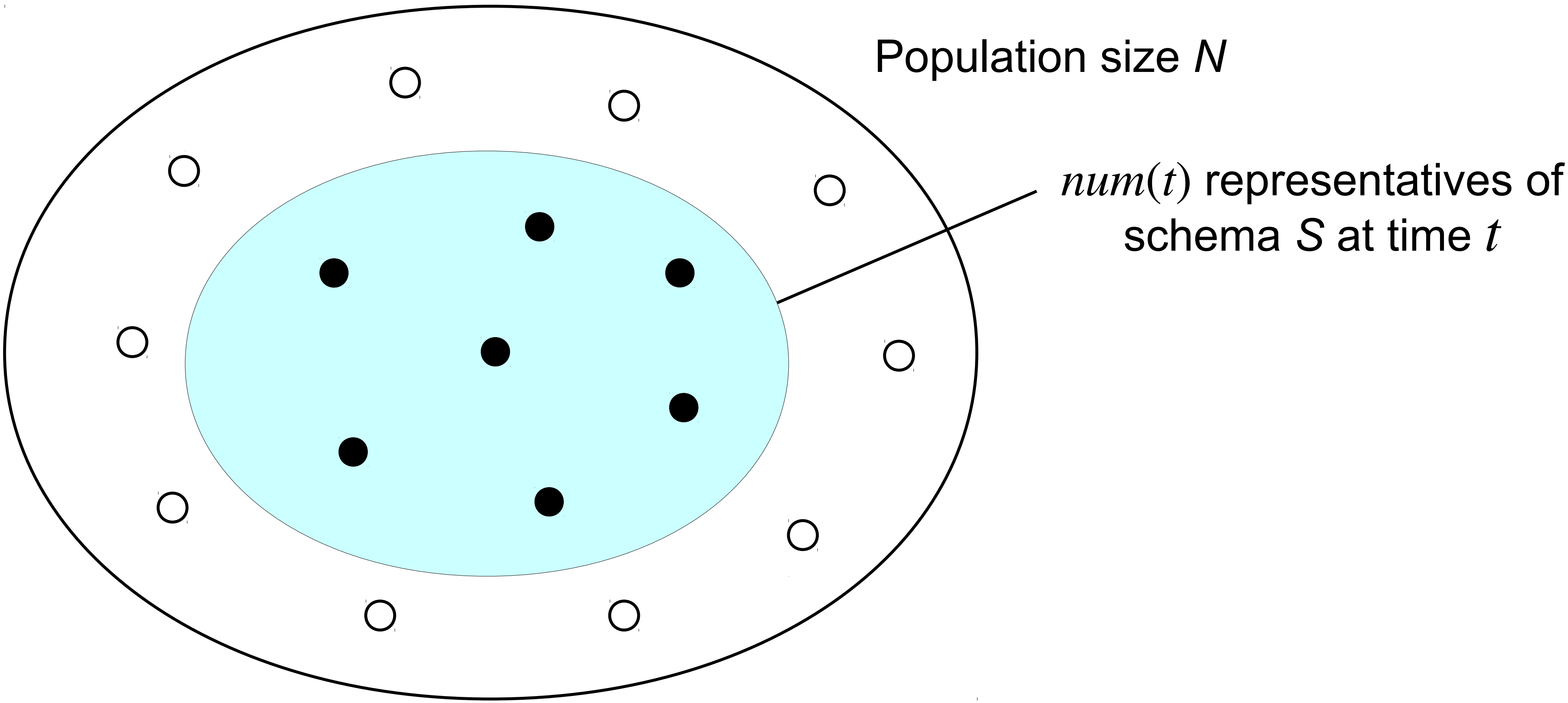
Example: If the probability that a person was born in California is **0.15**, then for a group of 1,000,000 random people, we expect approximately 0.15 $\times$ 1,000,000 = **150,000** of them to have been born in California.

Effect of Selection Operator

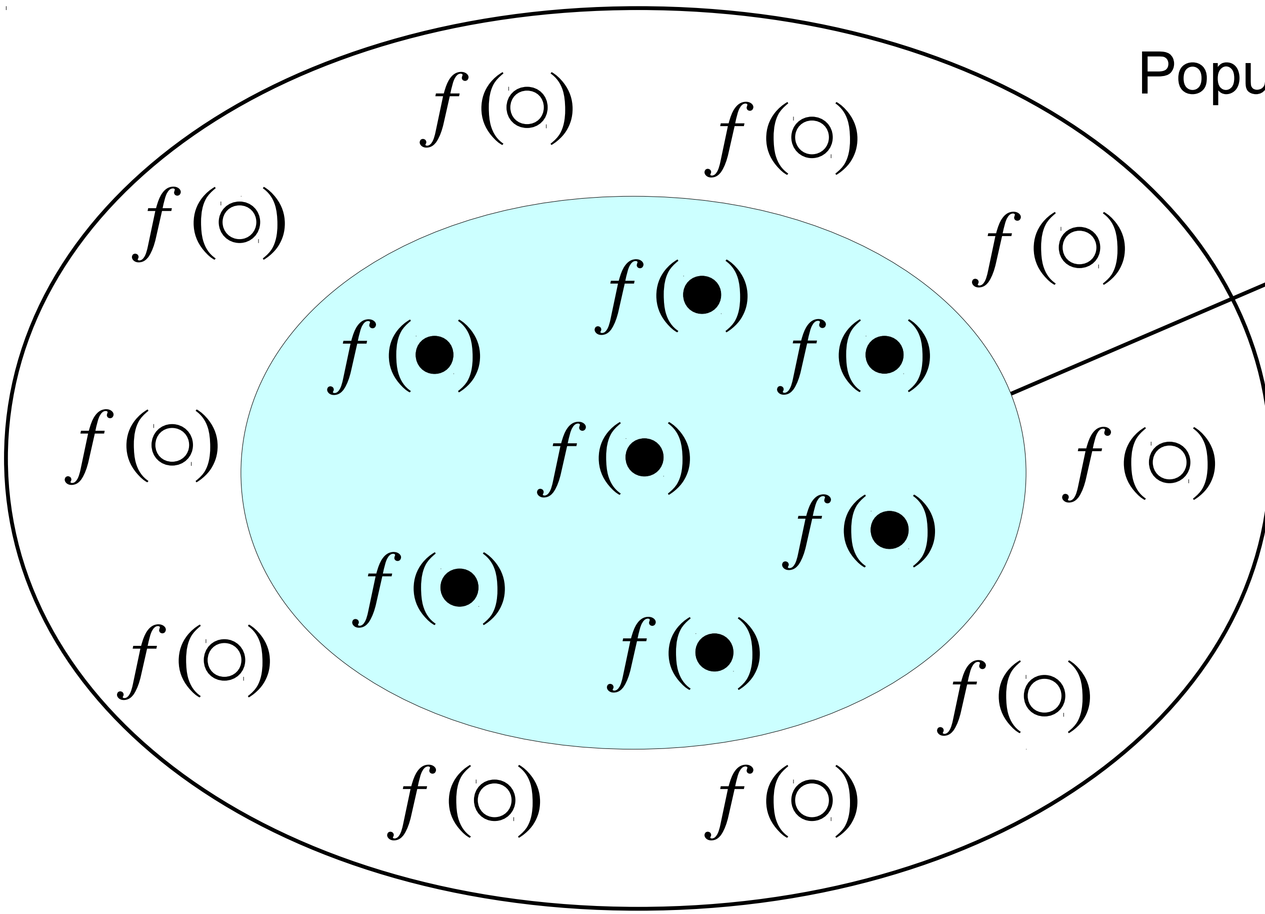


Population size N

Effect of Selection Operator



Effect of Selection Operator



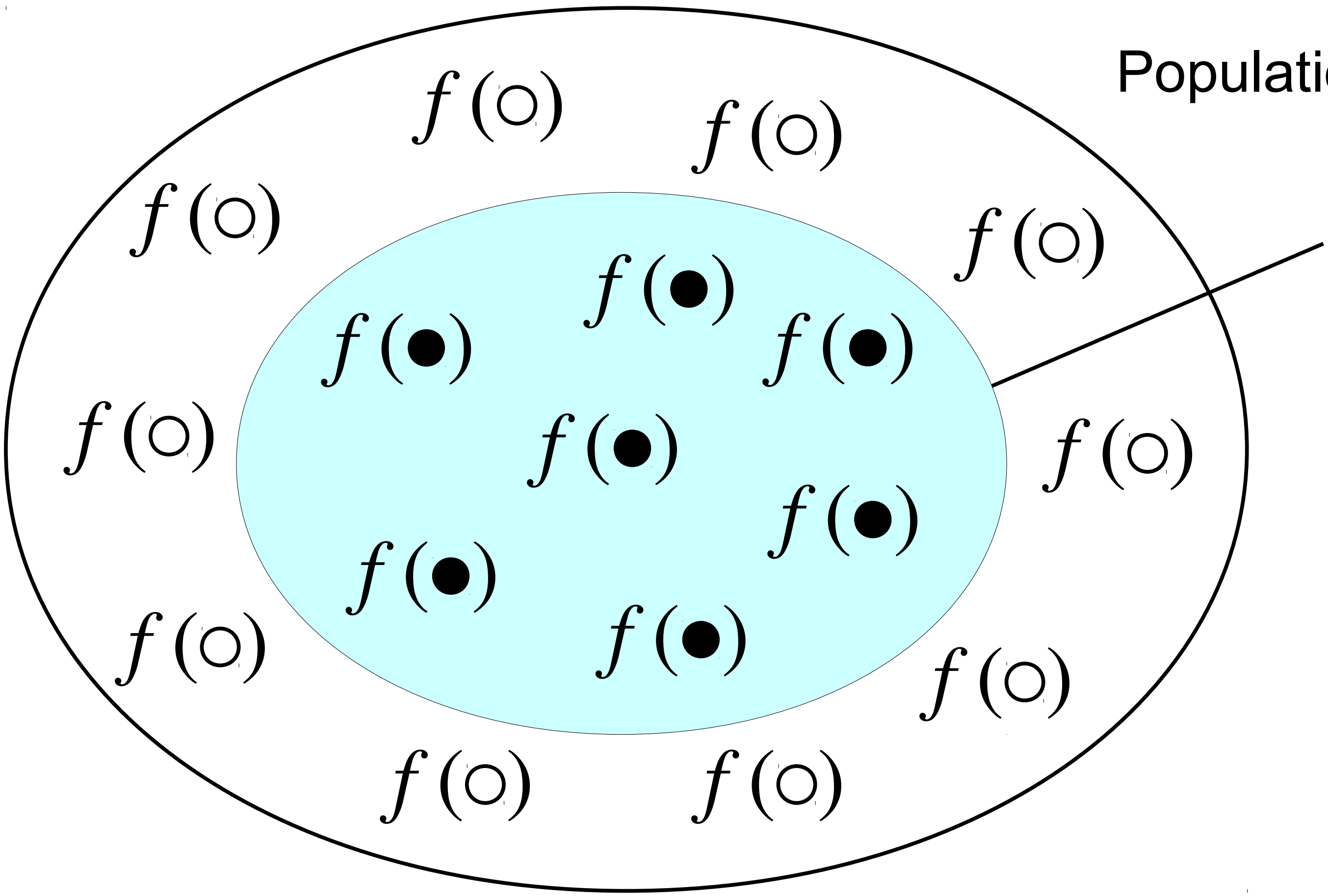
Population size *N*

$num(t)$ representatives of schema *S* at time t

estimated average fitness of schema *S* at time t

$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

Effect of Selection Operator



Population size *N*

$num(t)$ representatives of schema *S* at time t

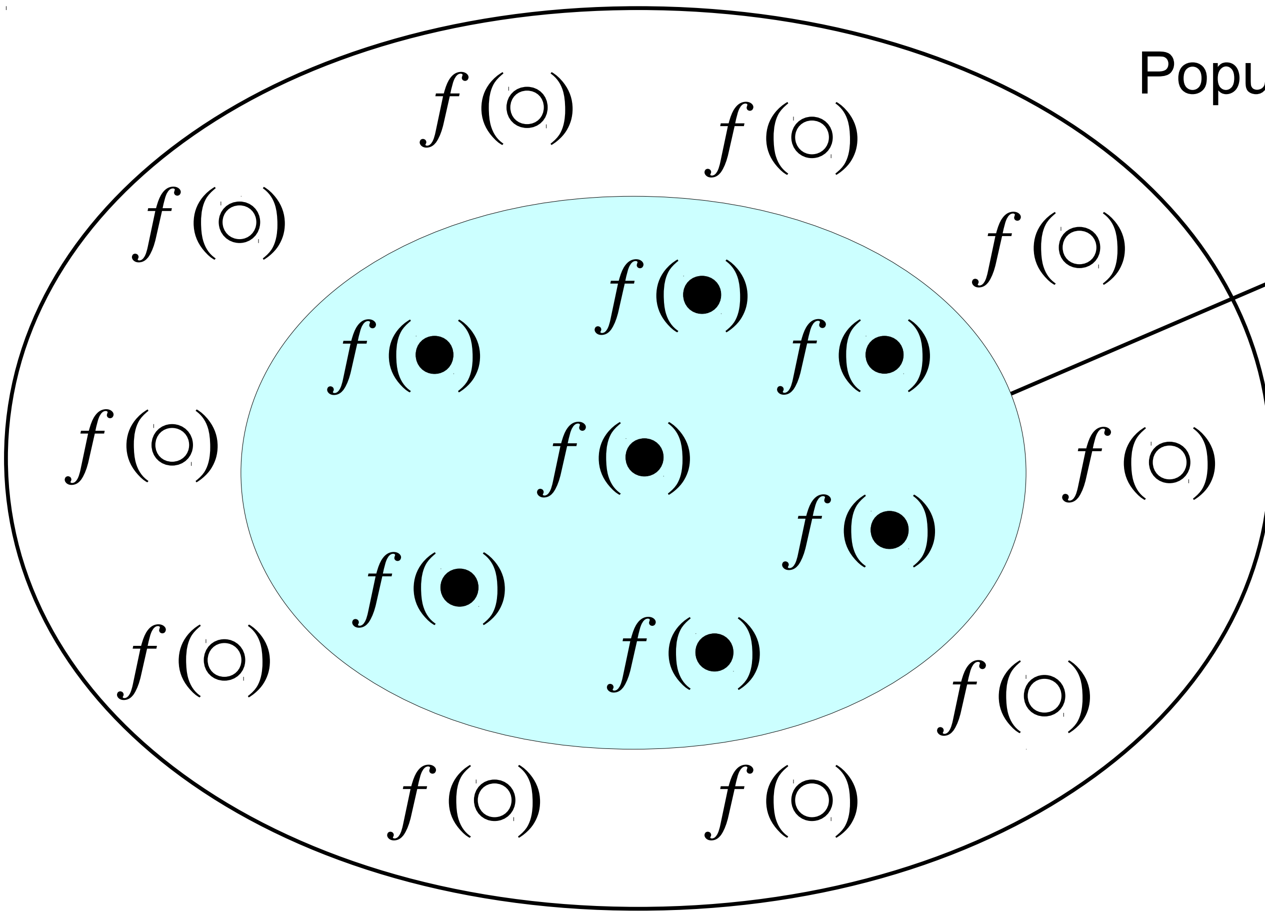
estimated average fitness of schema *S* at time t

$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

Effect of Selection Operator



Population size *N*

$num(t)$ representatives of schema *S* at time t

estimated average fitness of schema *S* at time t

$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

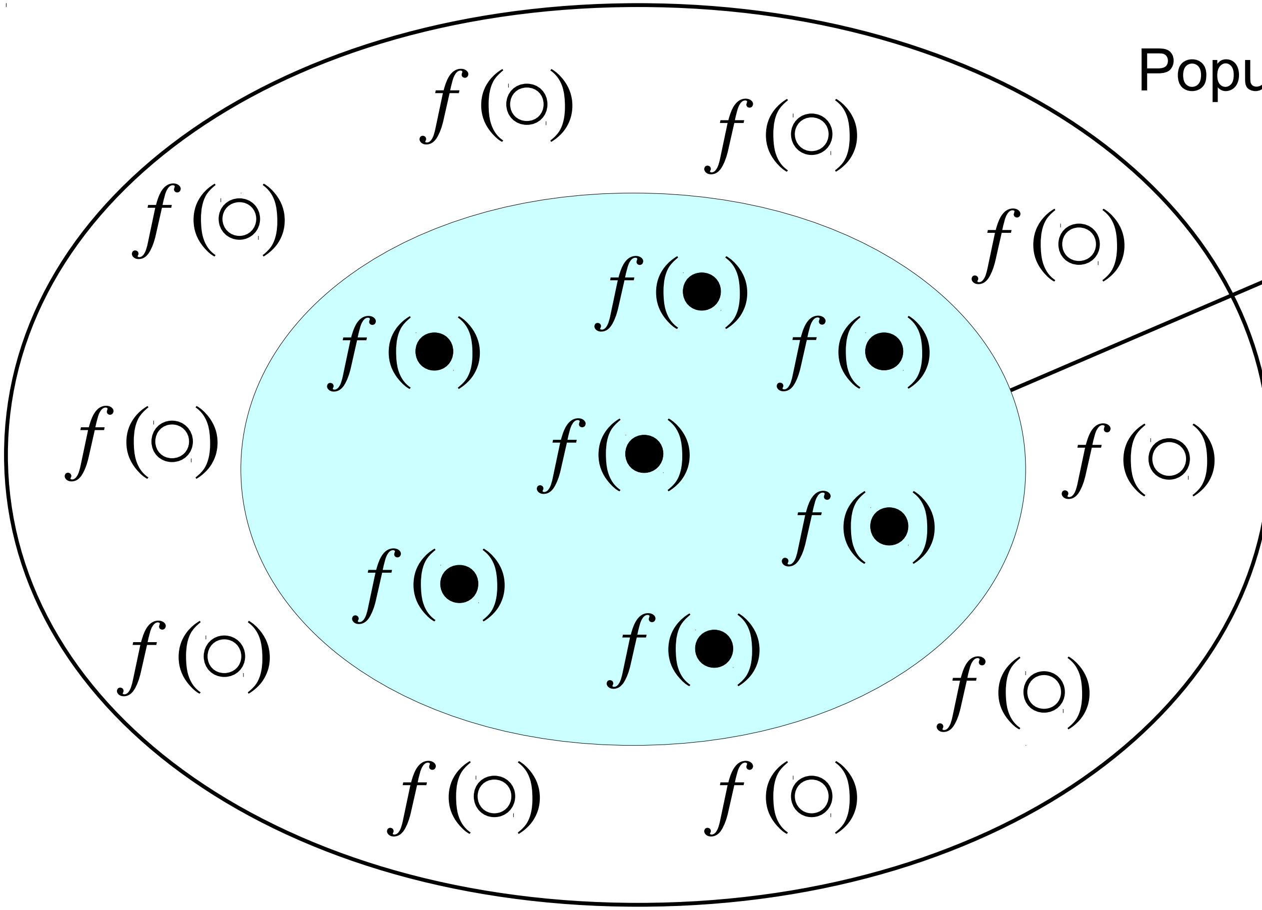
probability of picking a representative of *S* during selection

$$\frac{\sum f(\bullet)}{\sum f}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

Effect of Selection Operator



Population size N

$num(t)$ representatives of schema S at time t

estimated average fitness of schema S at time t

$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

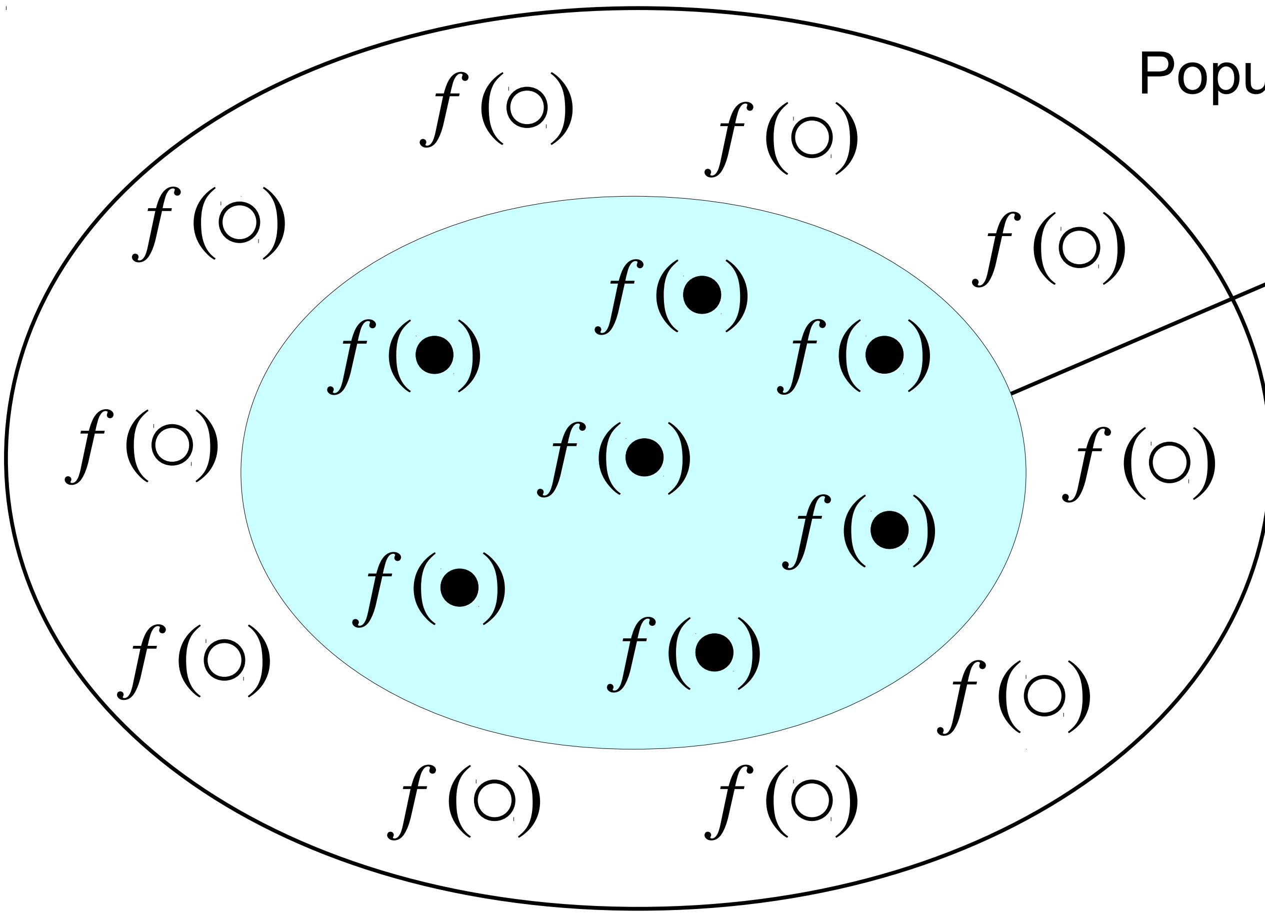
expected number of representatives of S in the next generation

$$\frac{\sum f(\bullet)}{\sum f} \times N$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

Effect of Selection Operator



Population size N

$num(t)$ representatives of schema S at time t

estimated average fitness of schema S at time t

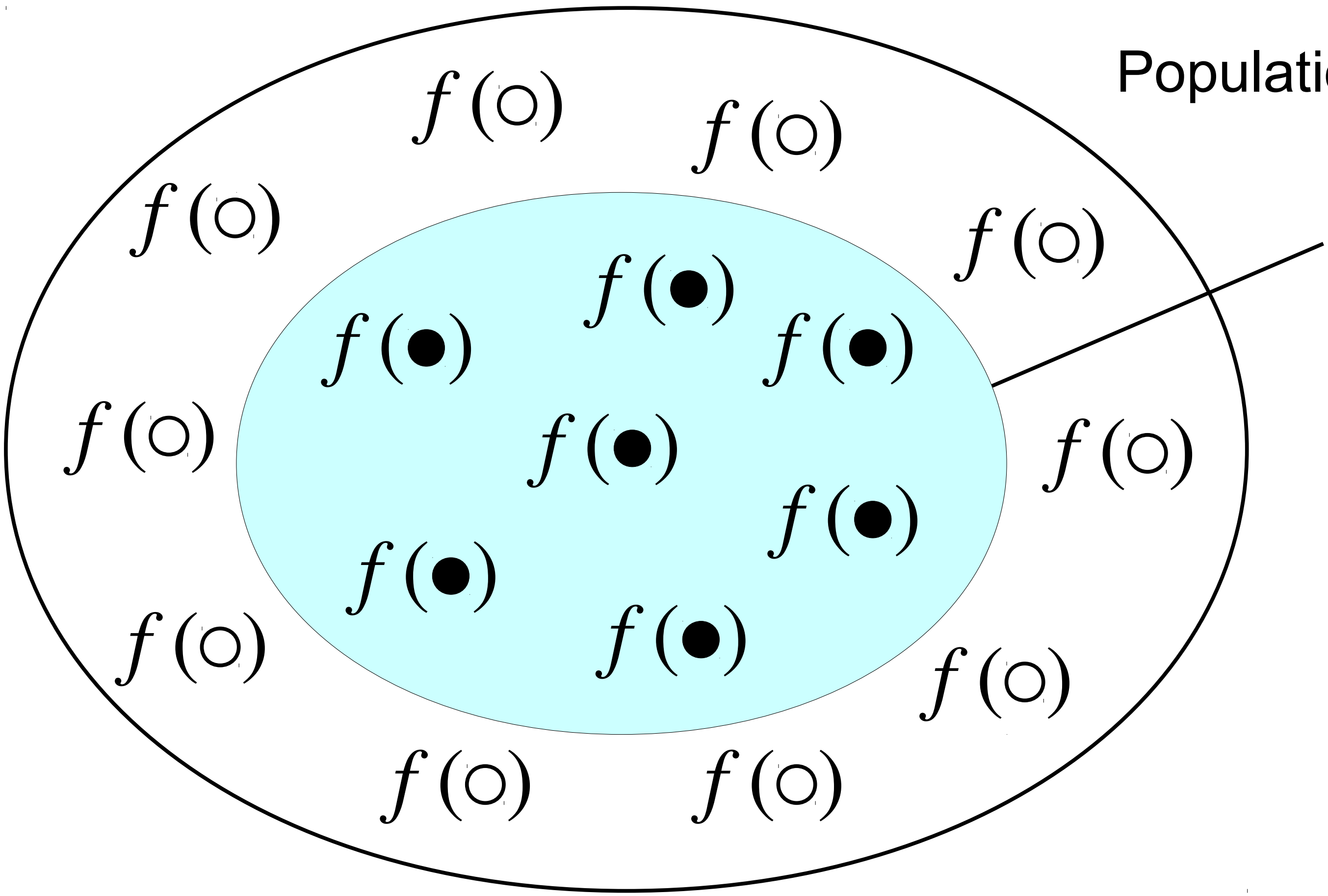
$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

$$num(t+1) = \frac{\sum f(\bullet)}{\sum f} \times N$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

Effect of Selection Operator



Population size *N*

$num(t)$ representatives of schema *S* at time t

estimated average fitness of schema *S* at time t

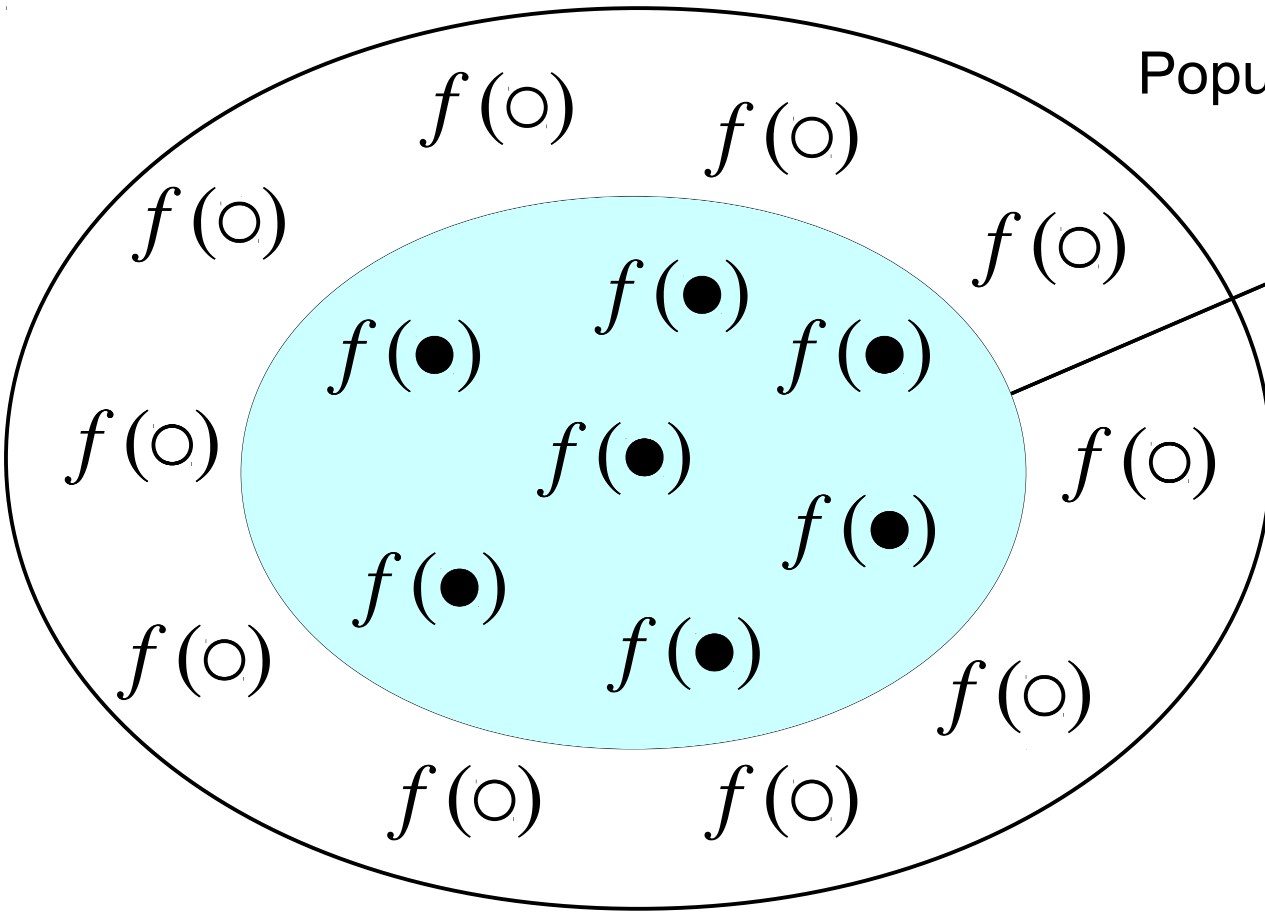
$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

$$num(t+1) \times \sum f = \sum f(\bullet) \times N$$

Effect of Selection Operator



Population size *N*

$num(t)$ representatives of schema *S* at time t

estimated average fitness of schema *S* at time t

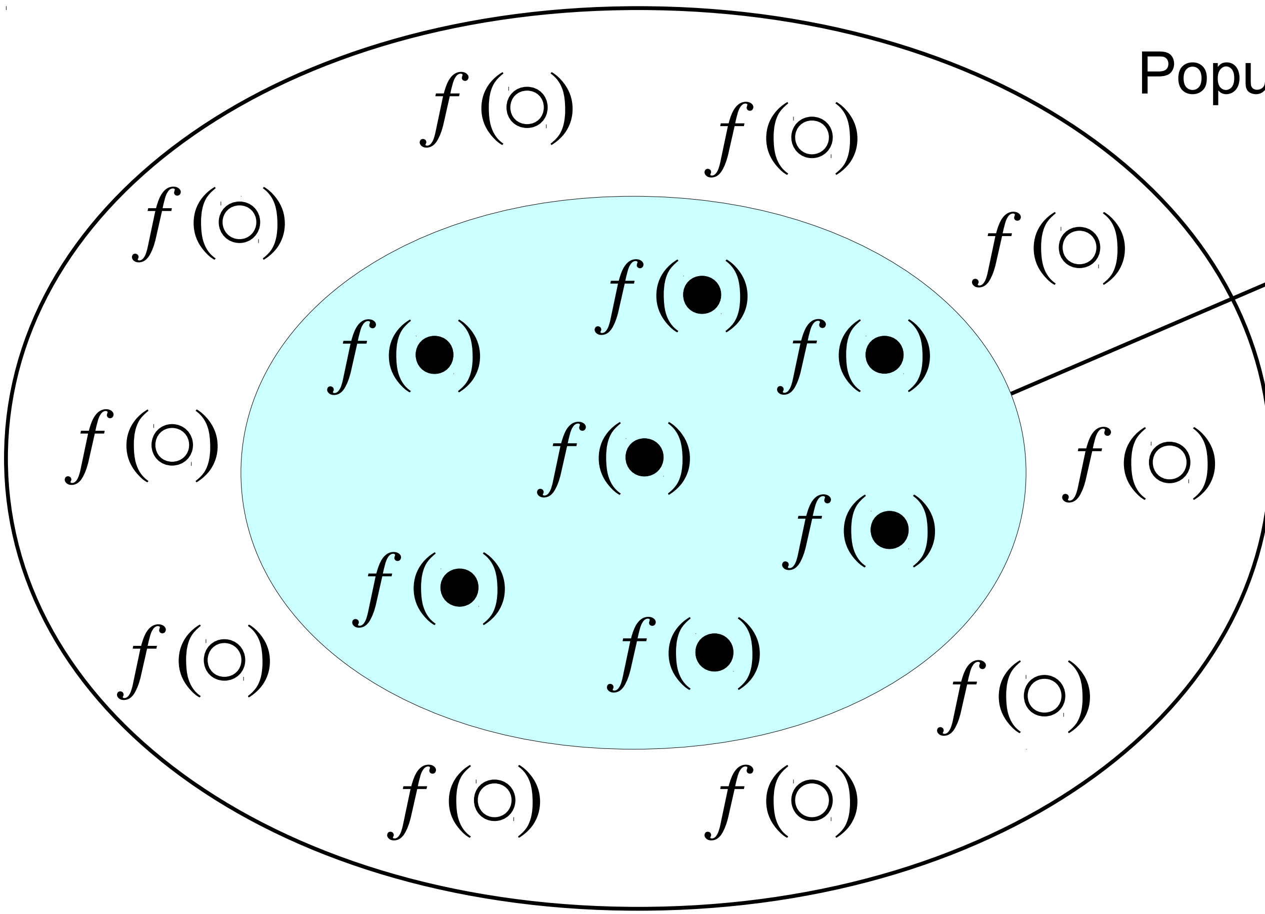
$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

$$num(t+1) \times \frac{\sum f}{N} = \sum f(\bullet)$$

Effect of Selection Operator



Population size N

$num(t)$ representatives of schema S at time t

estimated average fitness of schema S at time t

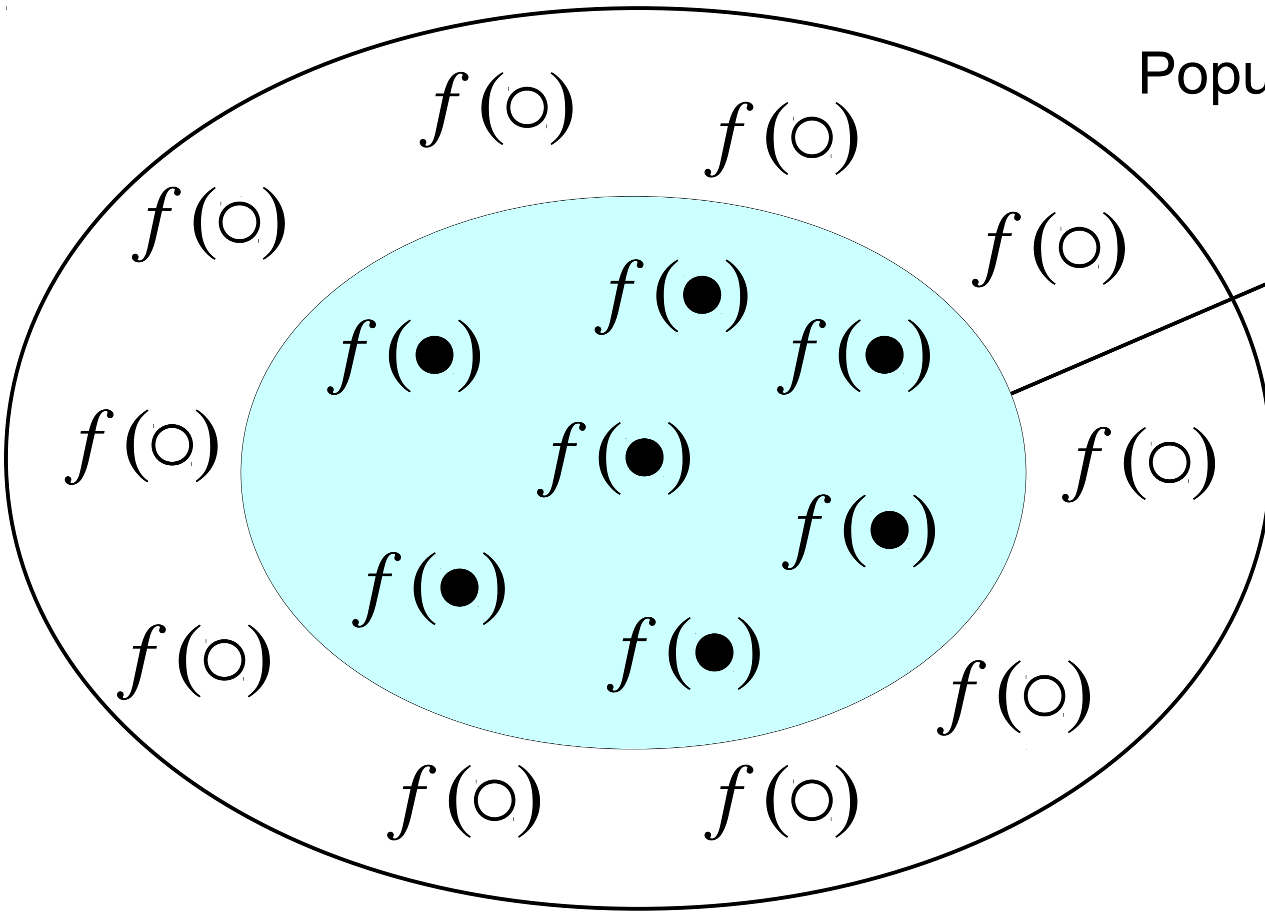
$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

$$num(t+1) \times \frac{\sum f}{N} = \sum f(\bullet)$$

Effect of Selection Operator



Population size N

$num(t)$ representatives of schema S at time t

estimated average fitness of schema S at time t

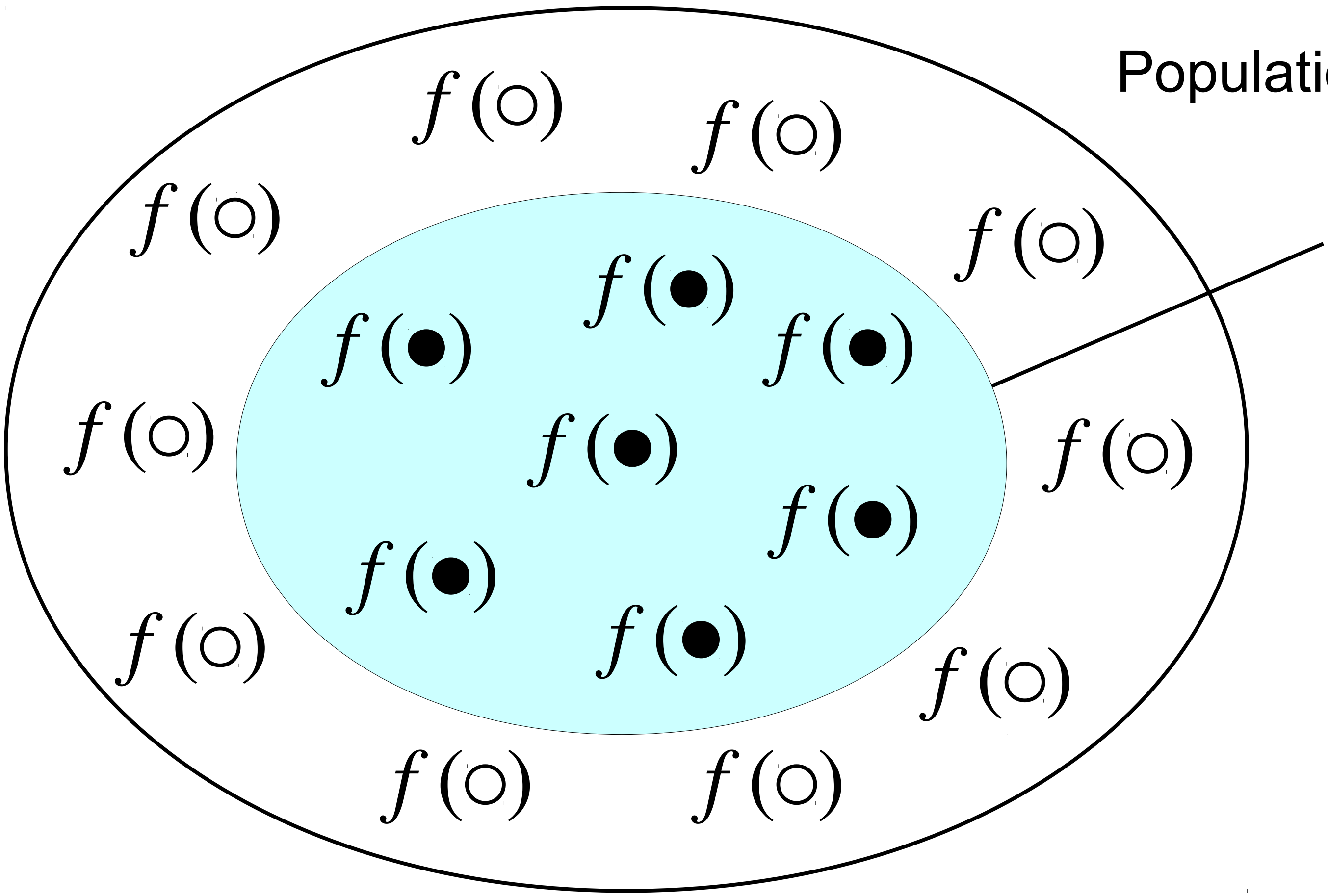
$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

$$num(t+1) \times avg = \sum f(\bullet)$$

Effect of Selection Operator



Population size *N*

$num(t)$ representatives of schema *S* at time *t*

estimated average fitness of schema *S* at time *t*

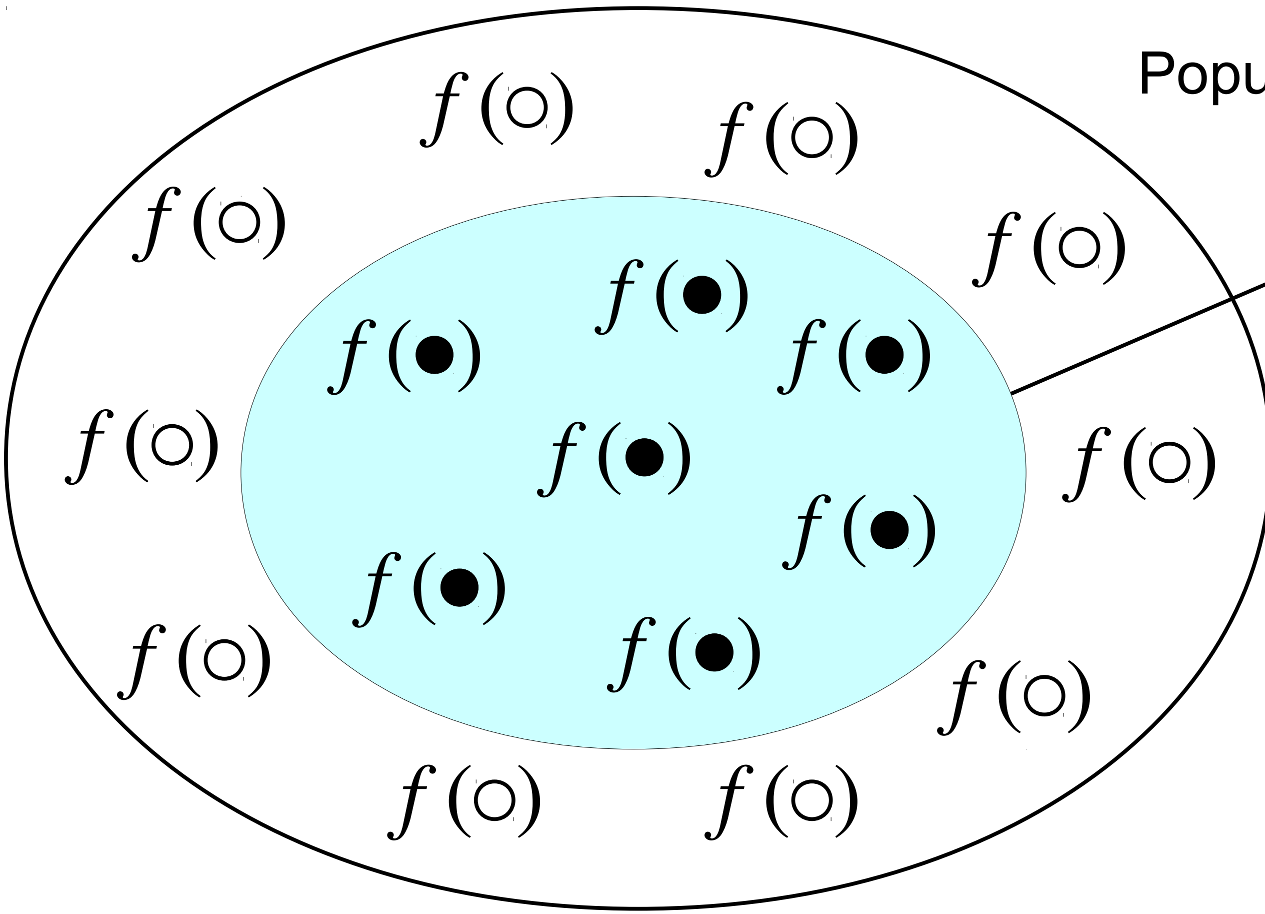
$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

$$\frac{num(t+1) \times avg}{num(t)} = \frac{\sum f(\bullet)}{num(t)}$$

Effect of Selection Operator



Population size *N*

$num(t)$ representatives of schema *S* at time *t*

estimated average fitness of schema *S* at time *t*

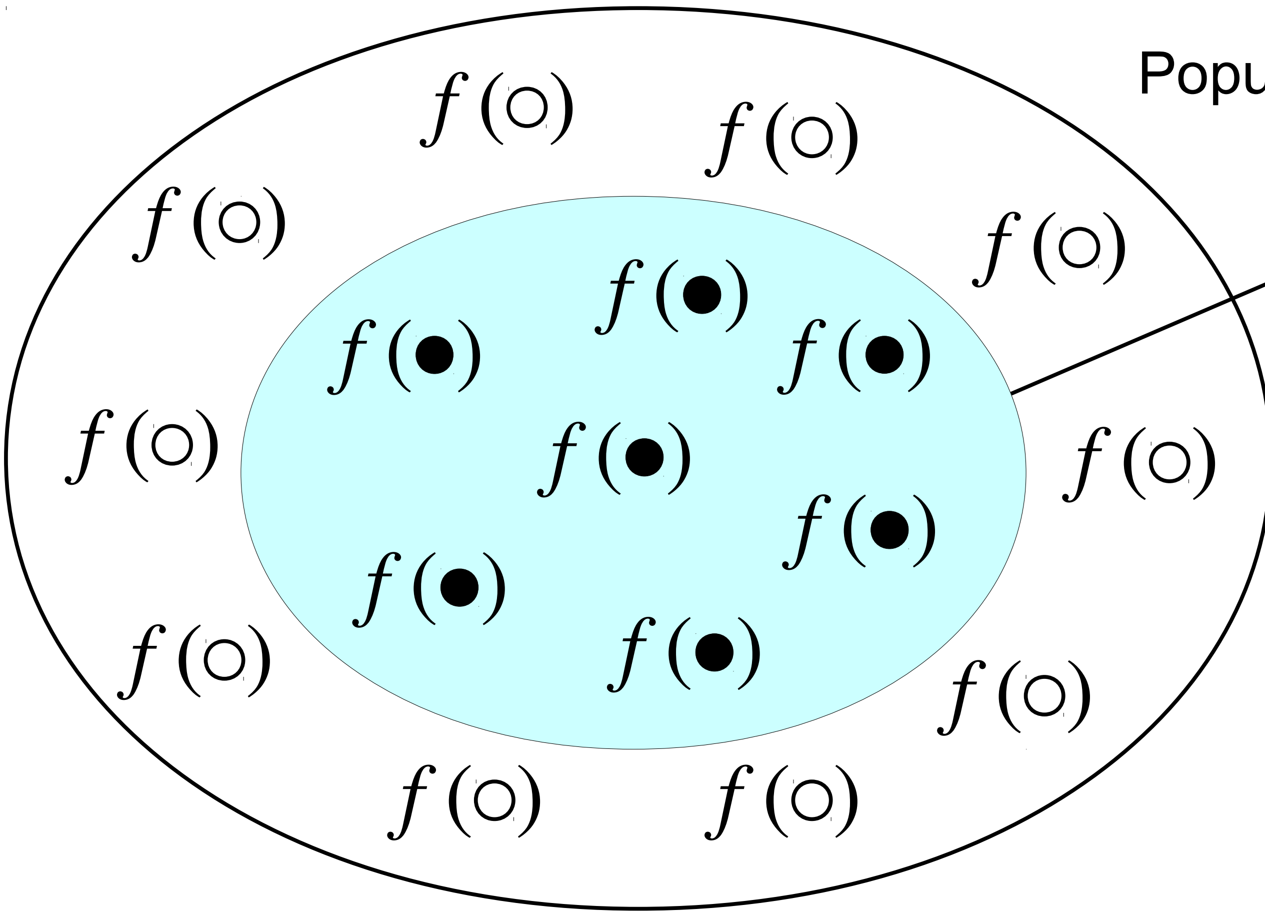
$$est(S) = \boxed{\frac{\sum f(\bullet)}{num(t)}}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

$$\frac{num(t+1) \times avg}{num(t)} = \boxed{\frac{\sum f(\bullet)}{num(t)}}$$

Effect of Selection Operator



Population size N

$num(t)$ representatives of schema S at time t

estimated average fitness of schema S at time t

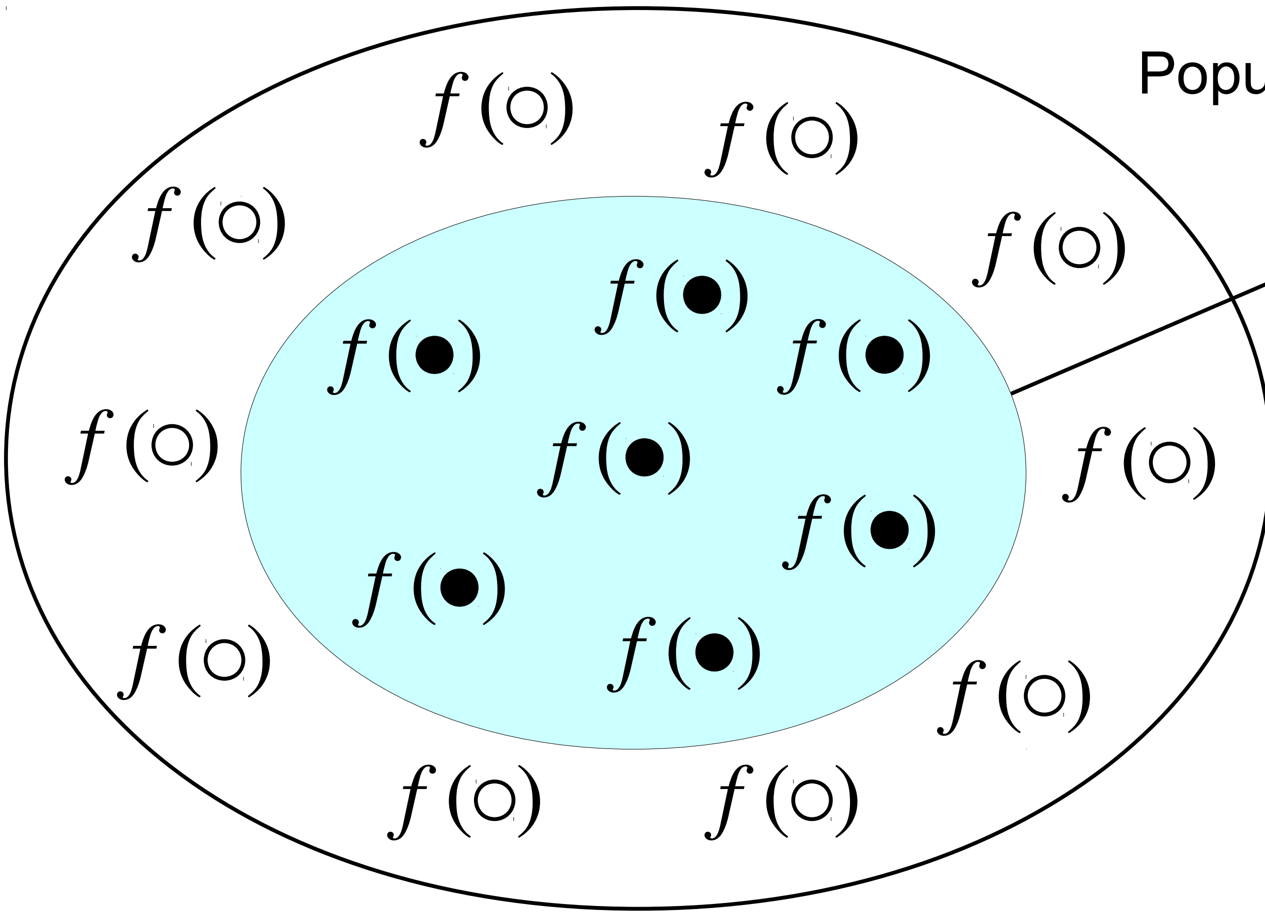
$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

$$\frac{num(t+1) \times avg}{num(t)} = est(S)$$

Effect of Selection Operator



Population size *N*

$num(t)$ representatives of schema *S* at time *t*

estimated average fitness of schema *S* at time *t*

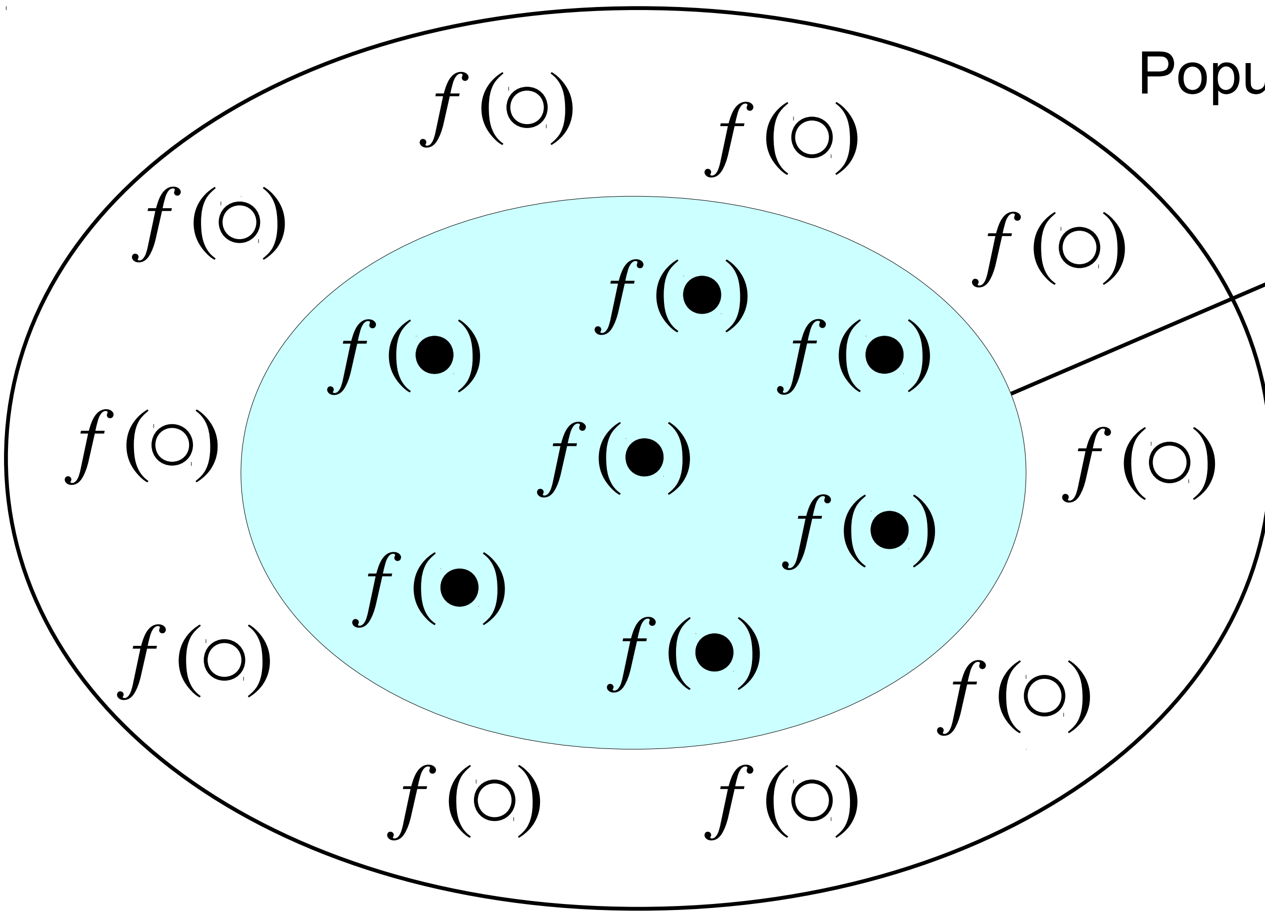
$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

$$num(t+1) \times avg = num(t) \times est(S)$$

Effect of Selection Operator



Population size *N*

$num(t)$ representatives of schema *S* at time *t*

estimated average fitness of schema *S* at time *t*

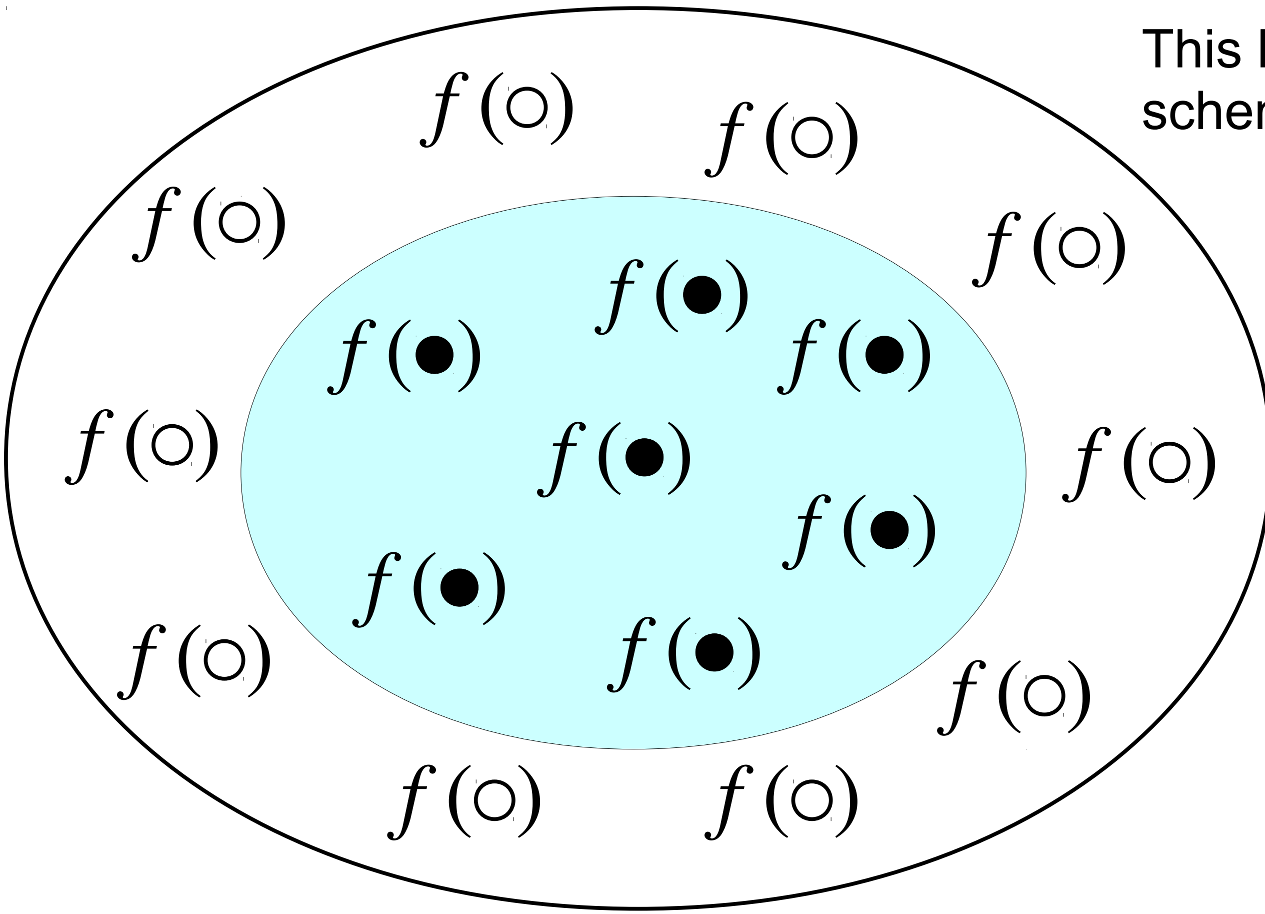
$$est(S) = \frac{\sum f(\bullet)}{num(t)}$$

average fitness of population

$$avg = \frac{\sum f}{N}$$

$$num(t+1) = num(t) \times \frac{est(S)}{avg}$$

Effect of Selection Operator



This holds simultaneously **for all** $O(2^L)$ schemas represented in the population!

“Implicit parallelism”

The fittest schemas receive more and more representatives in the population over time, even though the GA never keeps track of schemas explicitly, and does not calculate the exact average fitness of any schema.

$$num(t+1) \;=\; num(t) \times \frac{est(S)}{avg}$$

$$\frac{est(S)}{avg} > 1$$

# of representatives of *S* **increases exponentially**

$$\frac{est(S)}{avg} < 1$$

# of representatives of *S* **decreases exponentially**

Effect of Crossover Operator

S = * * 1 * * 0 *

Effect of Crossover Operator

$$S = * \quad * \quad \overbrace{1 \; \underbrace{\;\vdots\; * \;\vdots\; * \;\vdots\;}_{\text{“danger zone” for } S} \; 0}^{d(S)} \quad *$$

Probability of **picking a crossover point** in the danger zone?

$$= \frac{d(S)}{L-1}$$

Probability of **crossover actually occurring** in the danger zone?

$$= p_c \times \left(\frac{d(S)}{L-1} \right)$$

Effect of Crossover Operator

$$S = *\ \ *\ \ \overbrace{1 \,\vdots\, * \,\vdots\, * \,\vdots\, 0}^{d(S)}\ \ *$$

“danger zone” for *S*

Probability of *S* **surviving** under crossover?

$$= 1 - p_c \times \left(\frac{d(S)}{L-1}\right)$$

Probability of **crossover actually occurring** in the danger zone?

$$= p_c \times \left(\frac{d(S)}{L-1}\right)$$

Effect of Crossover Operator

$$S = * \quad * \quad \overbrace{1 \,\vdots\, * \,\vdots\, * \,\vdots\, 0}^{d(S)} \quad *$$

"danger zone" for *S*

Probability of *S* **surviving** under crossover?

$$= 1 - p_c \times \left(\frac{d(S)}{L-1}\right)$$

$$\begin{array}{l} 1\;1 \,\vdots\, \mathbf{1}\;0\;1\;\mathbf{0}\;0 \\ 0\;0 \,\vdots\, 0\;0\;0\;1\;0 \end{array} \longrightarrow \begin{array}{l} 0\;0 \,\vdots\, \mathbf{1}\;0\;1\;\mathbf{0}\;0 \\ 1\;1 \,\vdots\, 0\;0\;0\;1\;0 \end{array}$$

S survives
(outside danger zone)

Effect of Crossover Operator

$d(S)$

$$S = * \; * \; \underbrace{1 \,\vdots\, * \,\vdots\, * \,\vdots\, 0}_{\text{"danger zone" for } S} \; *$$

Probability of S **surviving** under crossover?

$$= 1 - p_c \times \left(\frac{d(S)}{L-1} \right)$$

1 1 **1** 0 ⋮ 1 **0** 0	$\rightarrow$	0 0 0 0 ⋮ 1 0 0	S is destroyed
0 0 0 0 ⋮ 0 1 0		1 1 1 0 ⋮ 0 1 0	

Effect of Crossover Operator

$$S = \; * \;\; * \;\; \overbrace{\mathbf{1} \,\vdots\, * \,\vdots\, * \,\vdots\, \mathbf{0}}^{d(S)} \;\; *$$

"danger zone" for *S*

Probability of *S* **surviving** under crossover?

$$= \; 1 - p_c \times \left(\frac{d(S)}{L-1} \right)$$

S may survive anyway, even if crossover occurs in the danger zone

1 1 **1** 0 ⋮ 1 **0** 0	$\longrightarrow$	1 1 **1** 1 ⋮ 1 **0** 0
1 1 1 1 ⋮ 0 1 0		1 1 1 0 ⋮ 0 1 0

S survives
(got lucky)

Effect of Crossover Operator

$$\overset{\longleftarrow\ d(S)\ \longrightarrow}{}$$

$$S = *\ *\ \mathbf{1}\ \vdots\ *\ \vdots\ *\ \vdots\ \mathbf{0}\ *$$

“danger zone” for *S*

Probability of *S* **surviving** under crossover?

$$\boxed{\geq 1 - p_c \times \left(\frac{d(S)}{L-1}\right)}$$

S may survive anyway, even if crossover occurs in the danger zone

1 1 **1** 0 ⋮ 1 **0** 0 → 1 1 **1** 1 ⋮ 1 **0** 0

1 1 1 1 ⋮ 0 1 0 → 1 1 1 0 ⋮ 0 1 0

S survives
(got lucky)

Effect of Mutation Operator

S = **`* * 1 * * 0 *`**

Effect of Mutation Operator

$$S = * \; * \; 1 \; * \; * \; 0 \; *$$

Number of defined bits ("order") $o(S) = 2$

Probability of single bit **mutation** $= p_m$

Probability of single bit **survival** $= (1 - p_m)$

Probability that **all defined bits survive** $= (1 - p_m)^{o(S)}$

Probability of *S* **surviving** under mutation $= (1 - p_m)^{o(S)}$

Putting It All Together

Expected number of representatives of schema S **selected** at time $t+1$

$$num(t+1) = num(t) \times \frac{est(S)}{avg}$$

Probability of survival under **crossover**

$$\geq 1 - p_c \times \left(\frac{d(S)}{L-1}\right)$$

Probability of survival under **mutation**

$$= (1 - p_m)^{o(S)}$$

Putting It All Together

Expected number
of representatives
of schema *S* **selected**
at time $t+1$

$$\boxed{num(t+1) = num(t) \times \frac{est(S)}{avg}}$$

$est(S)$ is the estimated average fitness of schema *S* based on its representatives in the current population

avg is the observed average fitness of all genomes in the current population

"Schemas with average fitness greater than the population average are likely to appear more in the next generation"

Putting It All Together

Expected number of representatives of schema S **selected** at time $t+1$

$$num(t+1) = num(t) \times \frac{est(S)}{avg}$$

Probability of survival under **crossover**

$$\geq 1 - p_c \times \left(\frac{d(S)}{L-1}\right)$$

"Schemas with shorter defining lengths are more likely to survive"

Putting It All Together

Expected number of representatives of schema S **selected** at time $t+1$

$$num(t+1) = num(t) \times \frac{est(S)}{avg}$$

Probability of survival under **crossover**

$$\geq 1 - p_c \times \left(\frac{d(S)}{L-1}\right)$$

Probability of survival under **mutation**

$$= (1 - p_m)^{o(S)}$$

"Lower-order schemas are more likely to survive"

Schema Theorem (Holland, 1975)

Expected number of representatives of schema S at time $t+1$:

$$num(t+1) \geq num(t) \underbrace{\frac{est(S)}{avg}}_{\text{selection}} \underbrace{\left[1 - p_c\left(\frac{d(S)}{L-1}\right)\right]}_{\text{crossover}} \underbrace{(1 - p_m)^{o(S)}}_{\text{mutation}}$$

“Schemas with above-average fitness, short defining length, and lower order are more likely to survive”

Building Block Hypothesis: GAs work by combining short, low-order schemas of above-average fitness into higher-order schemas with even higher fitness