



ALGORITMOS GENETICOS

Dr. Flabio Gutiérrez Segura (flabiogs@yahoo.es)

ALGORITMOS GENETICOS

NOCIONES BASICAS

- **DEFINICION DE ALGORITMOS GENETICOS**
- **ESTRUCTURA DE UN ALGORITMO GENETICO**
- **UN EJEMPLO INTRODUCTORIO**
- **CONCEPTOS BIOLOGICOS Y NOMENCLATURA**
- **MAPPING ENTRE FUNCION OBJETIVO → FUNCION DE FITNESS**
- **OTRO EJEMPLO**
- **OPERADORES GENETICOS**
- **REPRESENTACION**

Algoritmos Genéticos (AGs)

Procedimientos de búsqueda basados en los principios de la evolución y la selección natural. Pueden ser utilizados en problemas de optimización donde la búsqueda de soluciones óptimas se lleva a cabo en un espacio de soluciones codificadas mediante cadenas de longitud finita. Esta codificación es lo que se conoce como “representación genética” de cada posible solución.

Los orígenes de los AG pueden situarse a comienzos de los años 50, con los trabajos sobre simulaciones por computador de sistemas biológicos. el concepto de algoritmo genético tal como se conoce hoy en día es debido a los trabajos de Joh Holland de la Univesidad de Michigan a comienzos de de los 70.

Los operadores básicos que componen todo algoritmo genetico son:
Reproducción, cruce y mutación

Algoritmos Genéticos (AGs)

Campos de Aplicación

Aplicaciones Industriales

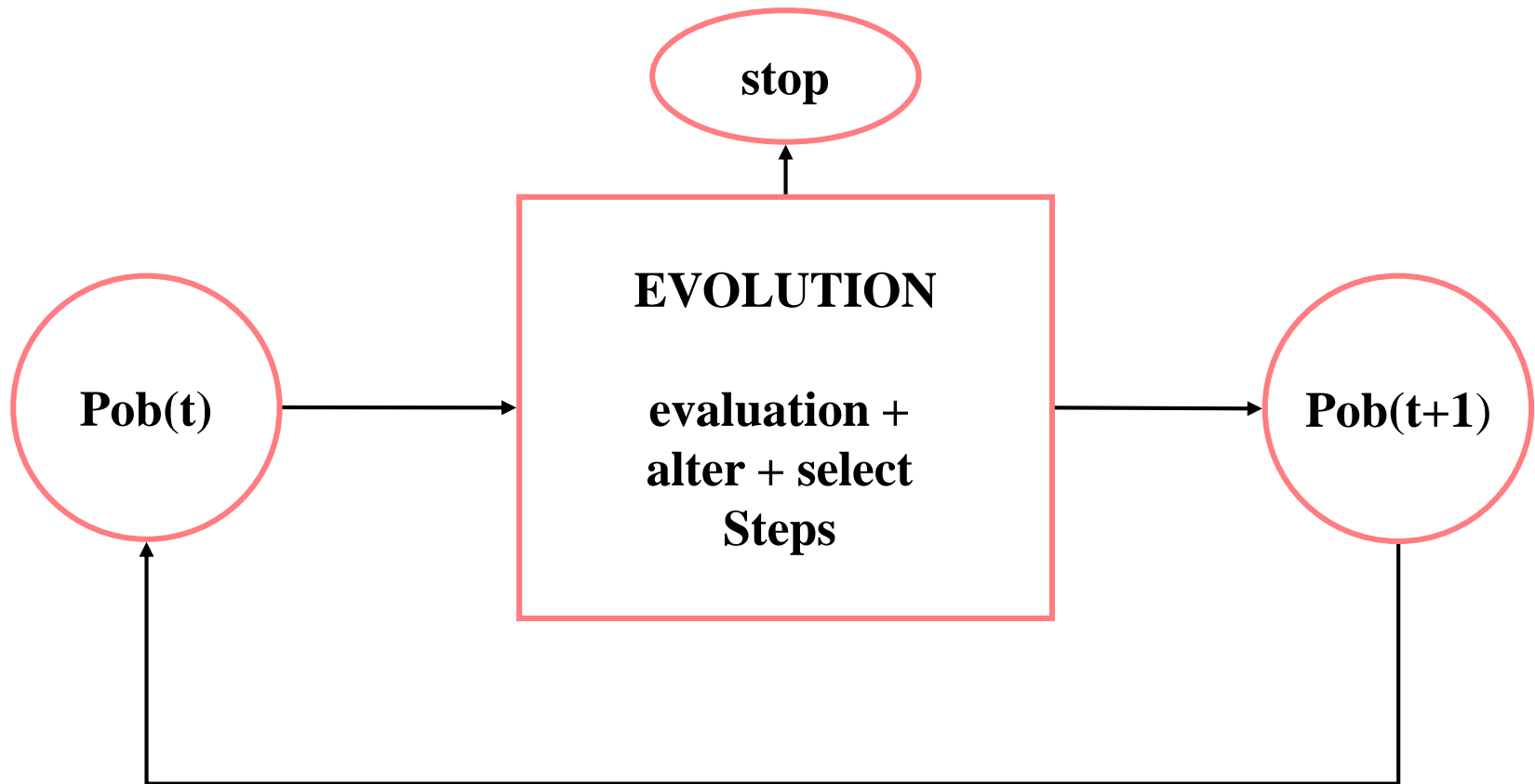
Gestión Empresarial

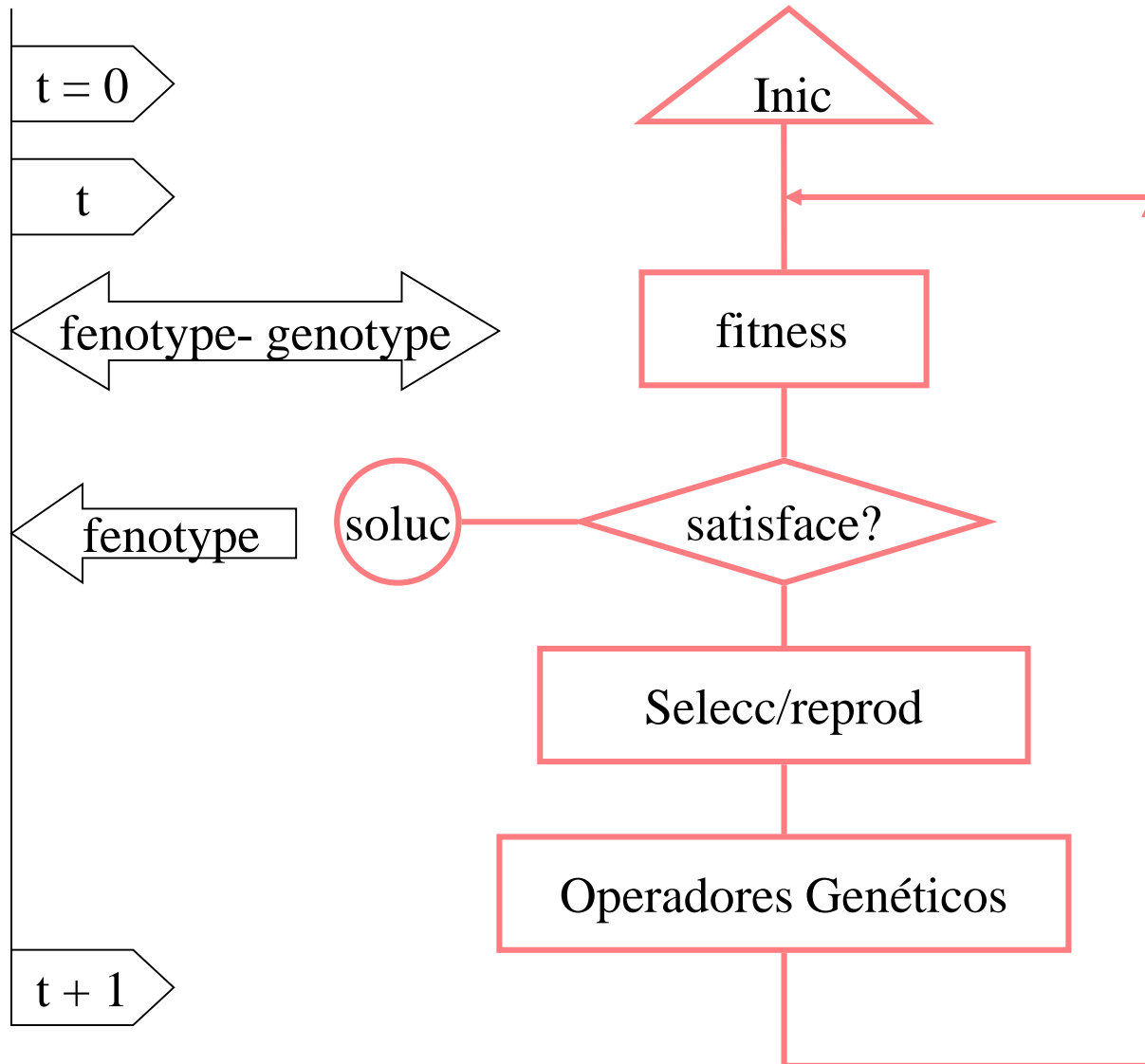
Finanzas

Logística

Investigación de Operaciones

• ESTRUCTURA DE UN ALGORITMO GENETICO





UN EJEMPLO

Sea el problema de maximizar $y=f(x)=x^2$ en $D=[0,15] \in \mathbb{Z}^+$.

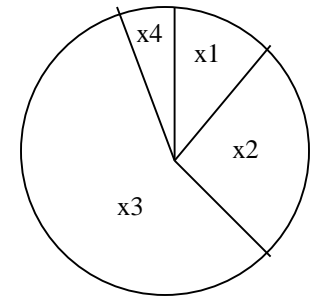
Representación

Optamos por una representación binaria. 4 dígitos son suficientes para representar un entero de 0 a 15 [(0000) a (1111)].

Generación de la población inicial

decidimos $\text{Pop_size} = 4$ y aleatoriamente resulta:

$$P(0) = \{ 0101, 1000, 1100, 0011 \}$$

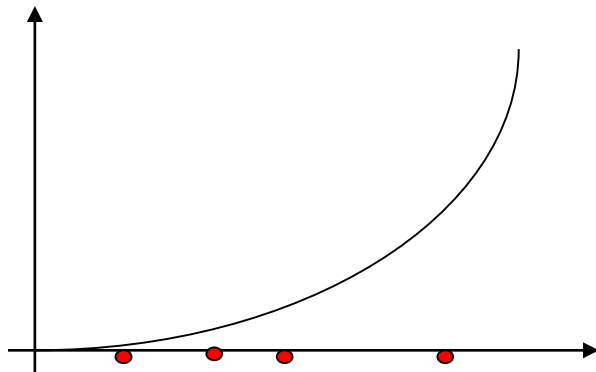


Evaluar

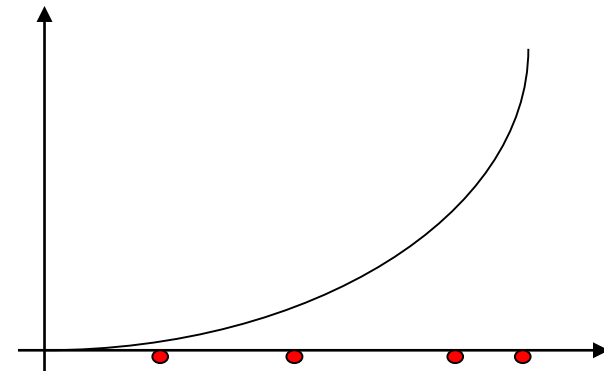
Individuo	Representación (genotipo)	Valor de x (fenotipo)	Fitness $f(x) = x^2$	$\frac{f_i}{\sum_i f_i}$	$\frac{\text{Nro esperado de hijos}}{n} = \frac{f_i}{\sum_i f_i}$	# copias
x_1^0	0 1 1 0	6	36	0.142	0.57	1
x_2^0	1 0 0 0	8	64	0.253	1.01	1
x_3^0	1 1 0 0	12	144	0.569	2.28	2
x_4^0	0 0 1 1	3	9	0.036	0.14	0
Totales			253	1.000		

Generamos el conjunto de apareamiento y aplicamos los operadores genéticos con $P_c = 1$ y $P_m = 0.001$

Individuo	conjunto de apareamiento luego de la selección	Pareja aleatoriamente seleccionada	Nueva Población	valor de x	Fitness $f(x) = x^2$	Psel $\frac{f_i}{\sum_i f_i}$
x_1^1	0 1 1 1	2	0 1 0 0	4	16	0.04
x_2^1	1 1 0 0	1	1 1 1 1	15	225	0.50
x_3^1	1 0 0 0	4	1 0 0 0	8	64	0.14
x_4^1	1 1 0 0	3	1 1 0 0	12	144	0.32
Totales					448	1.00



P(0) con $\bar{f} = 63,25$



P(1) con $\bar{f} = 112$

A través de las generaciones el AE tiende a converger (mayor fitness medio poblacional y agrupamiento cercano al óptimo).

CONCEPTOS Y NOMENCLATURA

El ejemplo anterior nos sirve para introducir algunos conceptos;

Cromosoma o *individuo*, es una solución potencial del problema.

Gen, elemento constitutivo de un cromosoma, identifica una posición del mismo. su valor se denomina *alelo*.

Los individuos pueden verse como el *genotipo* (su programa genético) y el *fenotipo* (rasgos de comportamiento expresado).

Paradigma *Neo Darwiniano* : La historia de la vida se basa en procesos físicos operando en poblaciones y especies. La evolución se transforma entonces en el resultado inevitable de la interacción entre estos procesos físicos:

- *reproducción*: transferencia del programa genético de un individuo a su progenie.
- *mutación*: errores de replicación durante la transferencia de información
- *competición*: consecuencia de la expansión de poblaciones en un espacio de recursos finitos
- *selección*: el resultado de la replicación competitiva a medida que la especie colma el espacio disponible

Vista de esta forma la evolución es un proceso de optimización (no necesariamente conducente a la perfección).

La selección, conduce a los fenotipos tan cerca del óptimo como sea posible. Sin embargo el ambiente cambia continuamente. Las poblaciones siguen estos cambios constantemente, evolucionando hacia nuevos óptimos.

Sin embargo la selección nunca deja de operar independientemente de la posición de la población en la topografía.

MAPPING ENTRE LA FUNCIÓN OBJETIVO Y LA FUNCIÓN DE FITNESS

Función objetivo $f(x)$, es la que deseamos optimizar

Función de fitness $g(x)$ es el valor de mérito de un individuo en la población (siempre con valor positivo o nulo)

Los problemas de optimización se clasifican como de maximización o minimización.

Para no obtener valores de fitness negativos debemos hacer una transformación $g(x)=T(f(x))$. Consideremos el dominio $D = [a,b]$.

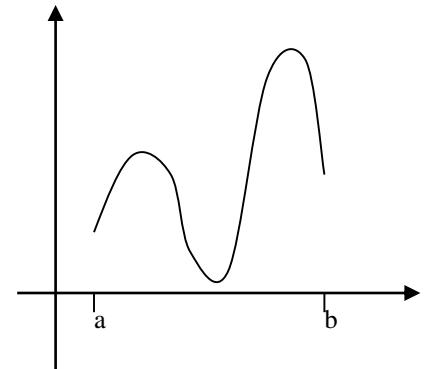
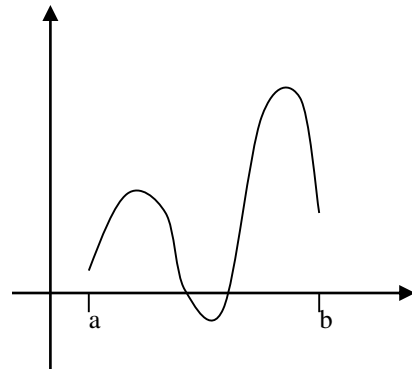
Maximización

Caso 1) Si $f(x) \geq 0 \quad \forall x \in D \Rightarrow g(x) = f(x)$

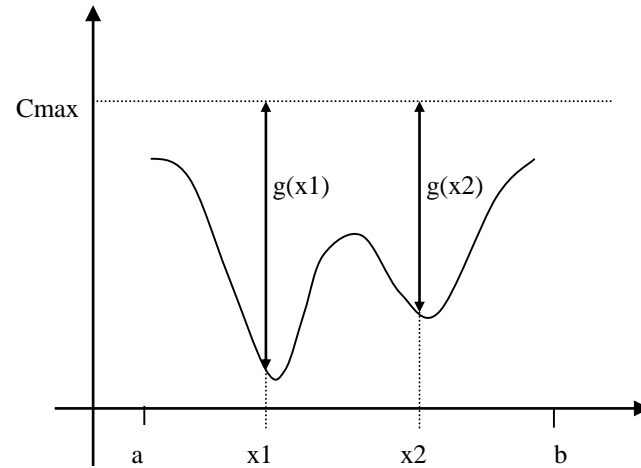
Caso 2) Si $f(x) < 0$ para algún $x \in D \Rightarrow g(x) = f(x) + C_{\min}$

tal que $g(x) \geq 0 \quad \forall x \in D$

$C_{\min} \geq |\min f(x)|$



Minimización



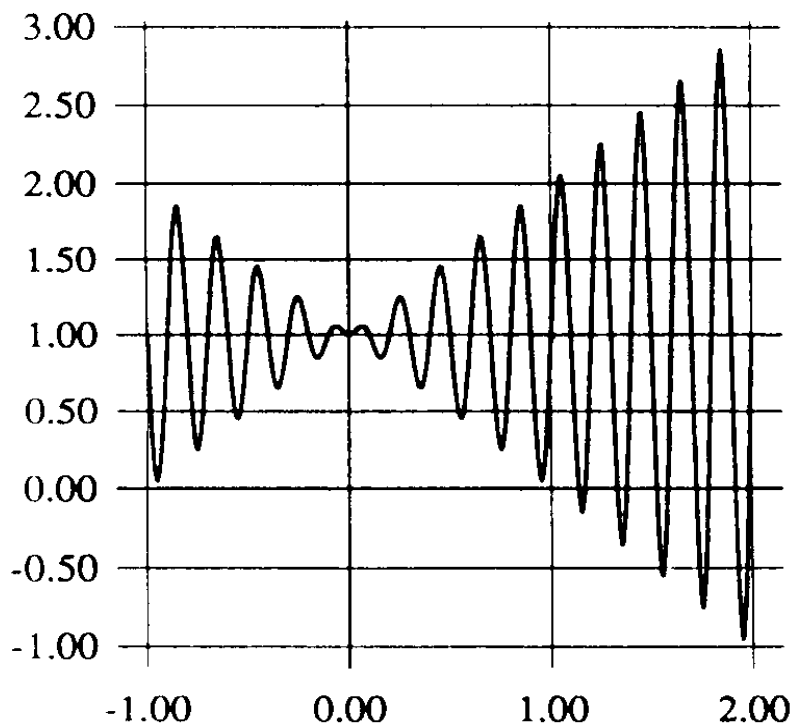
Para calcular el fitness de cada individuo, estimamos C_{max} tal que $C_{max} \geq \max f(x)$ en $D = [a, b]$ y definimos:

$$g(x) = C_{max} - f(x)$$

OTRO EJEMPLO

$$f(x) = x \cdot \sin(10\pi \cdot x) + 1.0$$

Debemos hallar x_o tal que $f(x_o) \geq f(x) \quad \forall x \in D$
con $D = [-1, 2]$



Los ceros para la derivada de esta función

$$f'(x) = \sin(10 \cdot \pi \cdot x) + 10\pi x \cdot \cos(10 \cdot \pi \cdot x) = 0$$

se encuentran en puntos de la forma

$$x_0 = 0,$$

$$x_i = \frac{2i \pm 1}{20} + \varepsilon_i, \text{ para } i = \pm 1, \pm 2, \dots \text{ donde los}$$

valores de ε_i son reales cercanos a cero.

Los máximos locales se alcanzan cuando i es impar y los mínimos cuando i es par. Como el dominio del problema es D , la función alcanza su máximo en

$$x_{19} = \frac{37}{20} + \varepsilon_{19} = 1.85 + \varepsilon_{19}$$

Son ligeramente mayores que $f(1.85) = 2.85$.

Representación

String binario representa un valor de la variable real x , con aproximación al 6to dígito decimal. El dominio de x tiene longitud 3 y los requerimientos de precisión indican que D debe dividirse en al menos 3×10^6 intervalos iguales.

Entonces 22 bits son suficientes pues,

$$2097152 = 2^{21} < 3000000 < 2^{22} = 4194304.$$

Mapping del string binario en un real $x \in D$

- Convertir el string binario a base 10, obteniendo x'

- A partir de x' obtener x por medio de la expresión , $x = a + x' \frac{b - a}{2^{22} - 1}$

donde $a = -1$ y $b = 2$

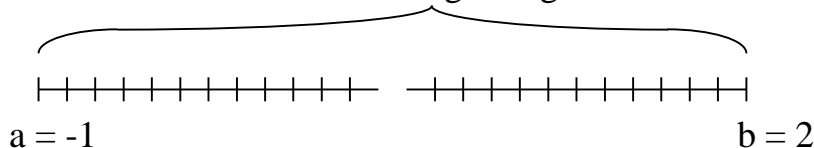
Por ejemplo el string (1000101110110101000111) representa el número

$x = 0.637197$ porque

$$x' = (1000101110110101000111)_2 = 2288967 \text{ y}$$

$$x = -1.0 + 2288967 (3/4194303) = 0.637197$$

4.194.303 intervalitos de longitud igual a 0.000007152559...



Población Inicial

El proceso de inicialización creará una población de cromosomas donde cada cromosoma es un string binario de 22 bits y cada bit es generado aleatoriamente.

Función de evaluación

La función *eval*, para un string binario s es equivalente a la función f : $eval(s) = f(x)$, donde cada cromosoma s representa un valor real x .

Por ejemplo dados tres cromosomas:

$$\begin{array}{lll} s_1 = (1000101110110101000111) & \text{con} & x_1 = 0.637197, \\ s_2 = (00000011100000000010000) & \text{con} & x_2 = -0.958973, \\ s_3 = (1110000000111111000101) & \text{con} & x_3 = 1.627888, \end{array}$$

y sus fitness dados por

$$\begin{array}{l} eval(s_1) = f(x_1) = 1.586345, \\ eval(s_2) = f(x_2) = 0.078878 \\ eval(s_3) = f(x_3) = 0.250650 \end{array}$$

Operadores Genéticos

mutación

Supongamos que el quinto gen de s_3 sufre una mutación entonces obtenemos:

$s'_3 = (1110100000111111000101)$, que representa a $x'_3 = 1.721638$ con el correspondiente valor $f(x'_3) = -0.082257$. Una caída en el valor del cromosoma (Fitness negativo ?)

Pero si el décimo gen de s_3 sufre una mutación entonces obtenemos :

$s'_3 = (1110000000111111000101)$, que representa a $x''_3 = 1.630818$ con el correspondiente valor $f(x''_3) = 2.50650$. Una mejora en el valor del cromosoma.

Crossover

Sean los cromosomas s_1 y s_3 y el punto de corte luego del 5to gen:

$$s_2 = (00000|01110000000010000)$$

$$s_3 = (11100|00000111111000101)$$

resultando los siguientes vástagos:

$$s'_2 = (00000|00000111111000101)$$

$$s'_3 = (11100|01110000000010000)$$

$$\text{eval}(s'_2) = f(x'_2) = f(-0.998113) = 0.940865$$

$$\text{eval}(s'_3) = f(x'_3) = f(1.666028) = 2.459245$$

El 2do vástago tiene mejor evaluación que sus padres.

Parámetros y Resultados

Habiendo elegido, $\text{pop_size} = 50$, $P_c = 0.25$, $P_m = 0.01$ se obtuvieron los siguientes resultados:

generación #	Eval
1	1.441942
10	2.250363
51	2.738930
137	2.850217
145	2.850227

El mejor cromosoma luego de 150 generaciones fue

$s_{\max} = (1111001101000100000101)$

que corresponde a $x_{\max} = 2.850773$

OPERADORES GENÉTICOS: MUTACIÓN Y CROSSOVER

La *mutación*, reemplaza aleatoriamente el alelo de un gen. Es decir, se crean nuevas soluciones tentativas haciendo pequeños cambios aleatorios en la representación de soluciones tentativas previas.

Aunque la mayoría de los AEs, basados en AGs, usan la mutación combinada con el *crossover*, aquella se usa como un operador secundario asegurando que la población consista de un *pool* diverso de alelos que pueden ser explotados por el *crossover*.

La idea para la aplicación del *crossover* radica en que teniendo dos individuos con alto valor de *fitness* deseamos combinar las mejores características de cada uno.

Como no sabemos que características cuentan para una buena *performance* (en caso contrario no estaríamos haciendo una búsqueda), lo mejor que podemos hacer es recombinar esas características aleatoriamente.

REPRESENTACIÓN

Cual será la estructura asociada a un individuo en la población?

String Binario, string de enteros, vectores de componentes reales, permutaciones de enteros, secuencias de vectores de componentes enteras, matrices, otras estructuras ?

La decisión depende del problema a tratar.

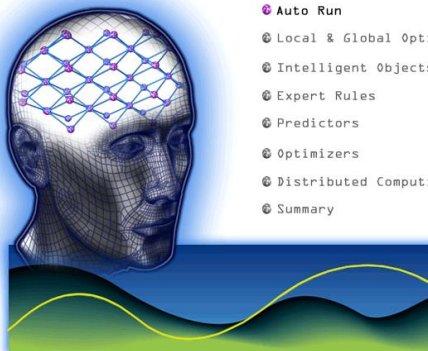
Los operadores genéticos deberán adecuarse a la representación usada, o si se hacen demasiado complejos deberemos revisar la representación.

Discutamos algunos ejemplos en clase....

Pensamiento:

Puede ser que la mente humana sea demasiado compleja para ser comprendida por la misma mente humana. Pero el deseo de intentar lo imposible parece ser una de sus características más persistentes.

(Sloman, 1978)





MUCHAS GRACIAS.....