

Clase 8: Algoritmos Genéticos

Sistemas de Inteligencia Artificial

ITBA - Primer Cuatrimestre de 2016

Introducción

Son una variante de Stochastic Beam Search en donde se combinan dos estados padres.

Analogía entre búsqueda local y evolución por selección natural:

- Los estados corresponden a individuos.
- Una función de evaluación que mide la calidad de los estados.
- Combinando buenos estados se obtienen estados mejores.

Selección natural



Algoritmos Genéticos

Son algoritmos de búsqueda u optimización basados en la mecánica de la selección natural (supervivencia del más apto), creados por John Holland en 1975

Los AG definen una *función de aptitud*.

La robustez de un método de optimización es el producto de

- la eficiencia, y
- la generalidad: se pueden adaptar a una gran variedad de problemas

Algoritmos Genéticos - Componentes

- Población de individuos (*cromosomas, posibles soluciones*)
- Selección de individuos para reproducción según su aptitud (*Fitness*)
- Generar la descendencia (*Cruce o Crossover*)
- Mutación al azar de los nuevos individuos
- Reemplazo de individuos para obtener una nueva generación

¿Cómo se representa un individuo?

La representación es fundamental para el correcto funcionamiento del algoritmo.

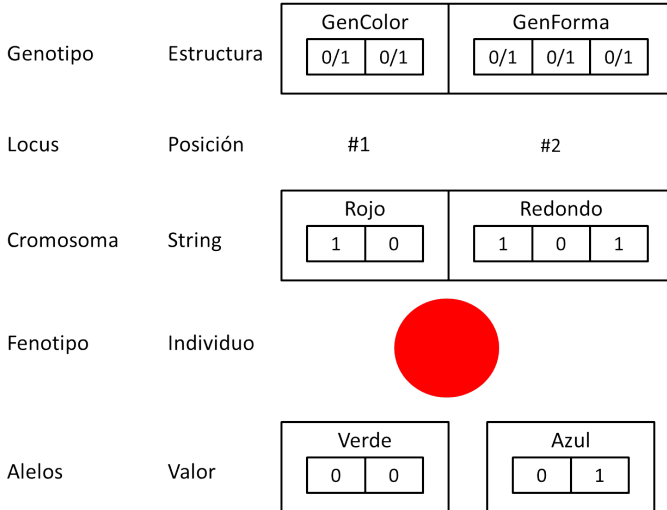
- Cadenas de bits
- Árboles
- Matrices
- Estructuras de datos complejas

¿Por qué funcionan?

Sean las siguientes cadenas binarias y los fitness correspondientes en un determinado problema (calcular el máximo de la función $f(x) = x^2$)

<i>Cadena</i>	<i>Fitness</i>
01101	169
11000	576
01000	64
10011	361

Nomenclatura



Aptitud relativa o normalizada

Dada una población x_1, x_2, \dots, x_n con aptitudes $f(1), f(2), \dots, f(n)$, respectivamente

La aptitud relativa de un individuo se define como:

$$p_i = \frac{f(i)}{\sum_{i=1}^n f(i)}$$

Convergencia Prematura

Convergencia de la población antes del tiempo necesario para llegar al óptimo global.

Esto provoca que el resultado del proceso de búsqueda arroje como solución un resultado subóptimo (generalmente un óptimo local).

Difícil de detectar pues muchas veces no se conoce a priori el valor del óptimo global.

Métodos de selección

Muestreo directo - Elite: Se seleccionan los individuos con mayor aptitud.

Muestreo aleatorio: Se seleccionan los individuos asignándoles, a cada uno de ellos, la misma probabilidad de ser elegidos.

Muestreo estocástico: Se seleccionan los individuos combinando su aptitud y el azar.

Selección por Elitismo

Kenneth De Jong (1975)

Retiene los mejores individuos de generación en generación.

Sólo Elitismo puede llevar a convergencia prematura.

Muchos investigadores mostraron que mejora significativamente la performance.

Métodos de selección estocásticos

- Ruleta
- Universal
- Boltzmann
- Torneos
- Ranking

Selección por Ruleta

Se calculan las aptitudes acumuladas:

$$q_0 = 0$$
$$q_i = \sum_{j=1}^i p_j$$

Se generan k números aleatorios $r_j \leftarrow URand[0, 1)$

y se seleccionan los k individuos x_i que cumplen:

$$q_{i-1} < r_j < q_i$$

Selección por Ruleta - Ejemplo

	<i>Aptitudes</i>		
	<i>Individual</i>	<i>Relativa</i>	<i>Acumulada</i>
x_1	3	0.086	0.086
x_2	6	0.171	0.257
x_3	11	0.314	0.571
x_4	14	0.400	0.971
x_5	1	0.029	1
	35		

Para $k = 3$

$$r_1 = 0.276$$

$$r_2 = 0.095$$

$$r_3 = 0.893$$

Selección por Ruleta - Ejemplo

	<i>Aptitudes</i>		
	<i>Individual</i>	<i>Relativa</i>	<i>Acumulada</i>
x_1	3	0.086	0.086
x_2	6	0.171	0.257
x_3	11	0.314	0.571
x_4	14	0.400	0.971
x_5	1	0.029	1
	35		

Para $k = 3$

$$r_1 = 0.276$$

$$r_2 = 0.095$$

$$r_3 = 0.893$$

Selección por Ruleta - Ejemplo

	<i>Aptitudes</i>		
	<i>Individual</i>	<i>Relativa</i>	<i>Acumulada</i>
x_1	3	0.086	0.086
x_2	6	0.171	0.257
x_3	11	0.314	0.571
x_4	14	0.400	0.971
x_5	1	0.029	1
	35		

Para $k = 3$

$$r_1 = 0.276$$

$$r_2 = 0.095$$

$$r_3 = 0.893$$

Selección por Ruleta - Ejemplo

	<i>Aptitudes</i>		
	<i>Individual</i>	<i>Relativa</i>	<i>Acumulada</i>
x_1	3	0.086	0.086
x_2	6	0.171	0.257
x_3	11	0.314	0.571
x_4	14	0.400	0.971
x_5	1	0.029	1
	35		

Para $k = 3$

$$r_1 = 0.276$$

$$r_2 = 0.095$$

$$r_3 = 0.893$$

Selección por Ruleta - Problema

Un problema de la selección de ruleta se presenta cuando existe una pequeña fracción de la población (en el límite, sólo un individuo) que posee una medida de desempeño excesivamente superior al resto.

Esto provoca pérdida de diversidad y puede conducir a convergencia prematura pues la mayor parte de los individuos seleccionados será una copia de los pocos predominantes.

Selección Universal

A partir de un solo número $r \leftarrow URand[0, 1)$ se calculan:

$$r_j = \frac{r + j - 1}{k} \quad j \in [1, k]$$

Se elijen los individuos que cumplen:

$$q_{i-1} < r_j < q_i$$

Selección Universal - Ejemplo

	<i>Aptitudes</i>		
	<i>Individual</i>	<i>Relativa</i>	<i>Acumulada</i>
x_1	3	0.086	0.086
x_2	6	0.171	0.257
x_3	11	0.314	0.571
x_4	14	0.400	0.971
x_5	1	0.029	1
	35		

Para $k = 3$ y $r = 0.276$

$$r_1 = 0.092$$

$$r_2 = 0.425$$

$$r_3 = 0.759$$

Selección Universal - Ejemplo

	<i>Aptitudes</i>		
	<i>Individual</i>	<i>Relativa</i>	<i>Acumulada</i>
x_1	3	0.086	0.086
x_2	6	0.171	0.257
x_3	11	0.314	0.571
x_4	14	0.400	0.971
x_5	1	0.029	1
	35		

Para $k = 3$ y $r = 0.276$

$$r_1 = 0.092$$

$$r_2 = 0.425$$

$$r_3 = 0.759$$

Selección Universal - Ejemplo

	<i>Aptitudes</i>		
	<i>Individual</i>	<i>Relativa</i>	<i>Acumulada</i>
x_1	3	0.086	0.086
x_2	6	0.171	0.257
x_3	11	0.314	0.571
x_4	14	0.400	0.971
x_5	1	0.029	1
	35		

Para $k = 3$ y $r = 0.276$

$$r_1 = 0.092$$

$$r_2 = 0.425$$

$$r_3 = 0.759$$

Selección Universal - Ejemplo

	<i>Aptitudes</i>		
	<i>Individual</i>	<i>Relativa</i>	<i>Acumulada</i>
x_1	3	0.086	0.086
x_2	6	0.171	0.257
x_3	11	0.314	0.571
x_4	14	0.400	0.971
x_5	1	0.029	1
	35		

Para $k = 3$ y $r = 0.276$

$$r_1 = 0.092$$

$$r_2 = 0.425$$

$$r_3 = 0.759$$

Selección de Boltzmann

Basada en Simulated Annealing

Usa una función de variación de “temperatura” que controla la presión de la selección

Presión de selección: condiciones a las que se somete una población. Cuanto mayor es la *presión de selección* mayor es la probabilidad de que sólo las más aptas logren sobrevivir.

Se usa un valor alto de temperatura al principio, lo cual hace que la *presión de selección* sea baja.

Con el paso de las generaciones, la temperatura disminuye, lo que aumenta la *presión de selección*.

Selección de Boltzmann

En general, se usa la siguiente expresión para calcular el valor esperado de un individuo:

$$ExpVal(i, t) = \frac{\exp^{f(i)/T}}{\langle \exp^{f(i)/T} \rangle^t}$$

donde:

- T es la temperatura, y
- $\langle \rangle^t$ denota el promedio de la población en la generación t

Selección por Torneos

Propuesta por Wetzel y estudiada por Brindle (1981) en su tesis doctoral

La idea básica del método es seleccionar en base a comparaciones directas de los individuos.

Hay 2 versiones:

1. Determinística
2. Probabilística

Selección por Torneos Determinística

Esta selección se efectúa mediante un torneo entre un pequeño subconjunto de individuos elegidos al azar desde la población.

Se elijen m individuos al azar y de ellos el mejor. Este proceso se repite k veces.

El parámetro m es un número entero (usualmente, 2 ó 3).

Selección por Torneos Determinística - Ejemplo

	<i>Aptitud</i>
x_1	3
x_2	4
x_3	15
x_4	25
x_5	1
x_6	2
	50

Para $k = 3$ y $m = 2$

$$x_3, x_4 \rightarrow x_4$$

Selección por Torneos Determinística - Ejemplo

	<i>Aptitud</i>
x_1	3
x_2	4
x_3	15
x_4	25
x_5	1
x_6	2
	50

Para $k = 3$ y $m = 2$

$$x_3, x_4 \rightarrow x_4$$

Selección por Torneos Determinística - Ejemplo

	<i>Aptitud</i>
x_1	3
x_2	4
x_3	15
x_4	25
x_5	1
x_6	2
	50

Para $k = 3$ y $m = 2$

$$x_3, x_4 \rightarrow x_4$$
$$x_1, x_3 \rightarrow x_3$$

Selección por Torneos Determinística - Ejemplo

	<i>Aptitud</i>
x_1	3
x_2	4
x_3	15
x_4	25
x_5	1
x_6	2
	50

Para $k = 3$ y $m = 2$

$$x_3, x_4 \rightarrow x_4$$

$$x_1, x_3 \rightarrow x_3$$

Selección por Torneos Determinística - Ejemplo

	<i>Aptitud</i>
x_1	3
x_2	4
x_3	15
x_4	25
x_5	1
x_6	2
	50

Para $k = 3$ y $m = 2$

$$x_3, x_4 \rightarrow x_4$$

$$x_1, x_3 \rightarrow x_3$$

$$x_2, x_3 \rightarrow x_3$$

	<i>Aptitud</i>
x_1	3
x_2	4
x_3	15
x_4	25
x_5	1
x_6	2
	50

$$x_2, x_3 \rightarrow x_3$$

Selección por Torneos Probabilística

Se eligen 2 individuos al azar.

Se toma un número al azar $r \in [0, 1]$.

Si $r < 0.75$ se selecciona el individuo más apto (de mayor aptitud), sino se selecciona al menos apto.

Ambos individuos se devuelven a la población original y podrían ser elegidos nuevamente.

Selección por Torneos - Beneficios

- Velocidad de aplicación (dado que no es necesario evaluar ni comparar la totalidad de la población)
- Capacidad de prevenir, en cierto grado, la convergencia prematura.
- Técnica eficiente y fácil de implementar.
- No requiere escalamiento de la función de aptitud (usa comparaciones directas).

Selección por Torneos - Desventaja

- Puede introducir una presión de selección muy alta porque a los individuos menos aptos no se les da oportunidad de sobrevivir.

Selección por Ranking

Se ordenan los cromosomas según el fitness y se les asigna una medida de desempeño inversamente proporcional a su fitness original.

La probabilidad de ser elegidos es proporcional a su posición en la lista (no importa si la diferencia de fitness entre un individuo y el siguiente es grande o chica).

Convergencia prematura

Una de las causas principales es la incorrecta parametrización de los operadores del algoritmo de búsqueda.

Por ejemplo:

- Una presión de selección muy elevada
- Una probabilidad de mutación muy baja
- Un tamaño de población insuficiente

Crossover - Cruza

Una vez seleccionada una pareja de individuos, se producirá entrecruzamiento con una probabilidad p_c

Tipos de cruce:

- Cruce de Un Punto
- Cruce de Dos Puntos
- Anular
- Cruce Uniforme Parametrizado

Crossover - Cruce de Un Punto

Se elige un locus al azar y se intercambian los alelos, a partir de ese locus.

Individuos seleccionados:

x_2	1	1	0	0	0	0	1	1	1
x_4	1	0	0	1	1	1	0	1	0

Punto de Cruce: 4

Hijos obtenidos:

h_1	1	1	0	1	1	1	0	1	0
h_2	1	0	0	0	0	0	1	1	1

Crossover - Cruce de Dos Puntos

Se eligen dos locus al azar, r_1 y r_2 con $r_1 < r_2$ y se intercambian los alelos correspondientes en ese intervalo.

Individuos seleccionados:

x_2	1	1	0	1	1	0	0	1	1
x_4	0	1	1	1	0	0	1	0	1

Puntos de Cruce: $r_1 = 3$ y $r_2 = 7$

Hijos obtenidos:

h_1	1	1	1	1	0	0	0	1	1
h_2	0	1	0	1	1	0	1	0	1

Crossover - Cruce Anular

Se elige un locus y luego un segmento, hacia la derecha, de longitud $l \in [1, L/2]$, donde L es la longitud del cromosoma.

Individuos seleccionados:

x_2	1	1	0	1	1	0	0	1	1
x_4	0	1	1	1	0	0	1	0	1

Puntos de Cruce: $r_1 = 5$ y $l = 2$

Hijos obtenidos:

h_1	1	1	0	1	0	0	1	1	1
h_2	0	1	1	1	1	0	0	0	1

Crossover - Cruce Anular

Se elige un locus y luego un segmento, hacia la derecha, de longitud $l \in [1, L/2]$, donde L es la longitud del cromosoma.

Individuos seleccionados:

x_2	1	1	0	1	1	0	0	1	1
x_4	0	1	1	1	0	0	1	0	1

Puntos de Cruce: $r_1 = 8$ y $l = 4$

Hijos obtenidos:

h_1	0	1	1	1	1	0	0	0	1
h_2	1	1	0	1	0	0	1	1	1

Crossover - Cruce Uniforme

Se produce el cruce de un alelo en cada locus con probabilidad:
 p (En general, $p = 0.5$)

Individuos seleccionados:

x_2	1	1	0	1	1	0	0	1	1
x_4	0	1	1	1	0	0	1	0	1

Hijos:

h_1	1	1	0	1	0	0	0	0	1
h_2	0	1	1	1	1	0	1	1	1

Mutación

Se define mutación como una variación de la información contenida en el código genético.

Se realiza la mutación con una probabilidad p_m

Motivación:

- Desbloqueo del algoritmo.
- Salir de poblaciones degeneradas.
- Incrementar el número de saltos evolutivos.
- Enriquecer la diversidad genética.

Mutación

Existen varias técnicas distintas de mutación. Algunas de éstas son:

Mutación de bit

Mutación multibit

Mutación de gen

Mutación multigen

Mutación de intercambio

Mutación de barajado

Método de Reemplazo 1

Población en la generación t con N individuos.

repeat

 | Seleccionar 2 padres

 | Recombinar

 | Mutar

until *Generar N hijos;*

Método de Reemplazo 2

Población original N individuos.

Seleccionar k padres

Recombinar

Mutar y generar k hijos

Seleccionar $N - k$ individuos de la población en la generación t que pasan sin modificación a la generación $t + 1$

Método de Reemplazo 3

Población original N individuos.

Seleccionar k padres

Recombinar

Mutar y generar k hijos

Seleccionar $N - k$ individuos de la población en la generación t que pasan sin modificación a la generación $t + 1$

Del total de $N + k$ individuos de la población en la generación t , seleccionar N que pasan a la siguiente generación

Brecha Generacional (Generation Gap)

Determina cuántos padres son copiados directamente desde la población anterior a la nueva

$$G \in [0, 1]$$

Si $G = 1$, toda la población original es reemplazada

Si $G = 0$, ningún individuo de la población original es reemplazado

Si $G \in (0, 1)$, se seleccionan $(1 - G) \times N$ individuos de la población anterior con cualquiera de los criterios de selección.

Los $G \times N$ no elegidos son reemplazados por los descendientes.
(Para el método de reemplazo 2, $G = k/N$).

Reemplazo - Selección Mixta

Por ejemplo: Elitismo + Ruleta

De la fracción G que se va a reemplazar, G/a con $a > 1$, se genera con Elitismo y $G(a - 1)/a$ con Ruleta

Parámetros

Dependen del problema a resolver.

- $N \approx [20, 200]$
- $p_m \approx [0.01, 0.001]$
- $p_c \approx [0.6, 0.95]$
- $G \approx [0.6, 1]$

Ejemplo

Un proceso controlado por cinco interruptores on-off tiene un retorno que varía en función de la configuración de los interruptores

Se desea maximizar el retorno

Se utiliza como codificación la representación binaria de cada una de los cinco interruptores

Por ejemplo: $x = 10010$ representa el primer y el cuarto interruptor prendido y el resto apagado

El proceso tiene un retorno Ret dado por la función: $Ret = x^2$, donde x es la representación binaria de un número entero

Ejemplo: Población inicial generada al azar

$$\begin{array}{rcl} x_1 & = & 0 \quad 1 \quad 1 \quad 0 \quad 1 \\ x_2 & = & 1 \quad 1 \quad 0 \quad 0 \quad 0 \\ x_3 & = & 0 \quad 1 \quad 0 \quad 0 \quad 0 \\ x_4 & = & 0 \quad 1 \quad 0 \quad 1 \quad 1 \end{array}$$

Ejemplo: Aptitud Inicial

Individuo	String	Aptitud	Aptitud relativa	Aptitud acumulada
x_1	0 1 1 0 1	$f(13) = 169$	0.14	0.14
x_2	1 1 0 0 0	$f(24) = 576$	0.49	0.63
x_3	0 1 0 0 0	$f(8) = 64$	0.06	0.69
x_4	1 0 0 1 1	$f(19) = 361$	0.31	1.00
		Total = 1170	Total = 1	
		Avg = 292.5		

Ejemplo: Selección universal

$$k = 2$$

$$r = 0.43 \Rightarrow \begin{cases} r_1 = 0.23 \\ r_2 = 0.73 \end{cases}$$

Individuo	Aptitud Acumulada
x_1	0.14
x_2	0.63
x_3	0.69
x_4	1.00

Progenitores: x_2, x_4

Ejemplo: Cruce

Punto de cruce:

$$\lfloor URand[0, 1) * (strLength - 1) \rfloor = \lfloor 0.31 \times (5 - 1) \rfloor = 1$$

Progenitores:

x_2 1 1 0 0 0

x_4 1 0 0 1 1

Hijos:

xd_1 1 1 || 0 1 1

xd_2 1 0 || 0 0 0

Ejemplo: Mutación

$$p_m = 0.003$$

Número aleatorio para cada gen de ambos descendientes:

r_i	=	0.7489	0.1665	0.7073	0.9669	0.8250
		0.2252	0.3613	0.0023	0.2856	0.6124

Hijos:

$$\begin{array}{rcl} xd_1 & = & 1\ 1\ 0\ 1\ 1 \\ xd_2 & = & 1\ 0\ 0\ 0\ 0 \end{array}$$

Hijos Mutados:

$$\begin{array}{rcl} xd_1 & = & 1\ 1\ 0\ 1\ 1 \\ xd_2 & = & 1\ 0\ \textcolor{red}{1}\ 0\ 0 \end{array}$$

Ejemplo: Reemplazo

Individuo	String	Aptitud	Aptitud relativa	Aptitud acumulada
x_1	0 1 1 0 1	$f(13) = 169$	0.074	0.074
x_2	1 1 0 0 0	$f(24) = 576$	0.250	0.324
x_3	0 1 0 0 0	$f(8) = 64$	0.028	0.352
x_4	1 0 0 1 1	$f(19) = 361$	0.157	0.509
xd_1	1 1 0 1 1	$f(27) = 729$	0.317	0.826
xd_2	1 0 1 0 0	$f(20) = 400$	0.174	1.000
		Total = 2299	Total = 1	

Ejemplo: Reemplazo por Ruleta

 $k = 4$
$$r_i = \begin{array}{cccc} 0.7219 & 0.8926 & 0.2044 & 0.6511 \end{array}$$

Se eligen los individuos: xd_1, xd_2, x_3
y xd_1

Individuo	Aptitud Acumulada
x_1	0.074
x_2	0.324
x_3	0.352
x_4	0.509
xd_1	0.826
xd_2	1.000

Población final

Individuo	String	Aptitud
x_1	1 1 0 1 1	$f(27) = 729$
x_2	1 0 1 0 0	$f(20) = 400$
x_3	0 1 0 0 0	$f(8) = 64$
x_4	1 1 0 1 1	$f(27) = 729$
		Total = 1922
		Avg = 480.5

Pseudocódigo

Generar una Población inicial

Evaluar el fitness de todos los individuos

while *No se cumpla la condición de corte* **do**

 Seleccionar individuos para su reproducción

 Recombinarlos

 Mutar algunos individuos

 Evaluar el fitness de los individuos obtenidos

 Generar una nueva población

end

Criterios de Corte

- Máximo número de generaciones.
- Se alcanza la solución óptima o se llega a un fitness inferior a una cota.
- Una parte relevante de la población no cambia de generación en generación (Estructura).
- El mejor fitness de la población no progresa con las generaciones (Contenido).

Teorema de los esquemas

Aquellos esquemas con un fitness medio superior a la media de la población, de longitud pequeña y con un orden bajo, aumentarán su presencia de manera exponencial en las sucesivas generaciones

Propiedades de los esquemas

Un esquema con k símbolos “*” representa 2^k cadenas binarias

Una cadena binaria de longitud Len encaja en 2^{Len} esquemas distintos

Existen 3^{Len} posibles esquemas

Una población de N individuos, codificados como cadenas binarias, contiene entre 2^{Len} y $N \times 2^{Len}$ esquemas distintos

Orden del esquema S , $o(S)$

Número de posiciones fijas que contiene dicho esquema

$S_1 = (*01*)$	$o(S_1) = 2$
$S_2 = (**1)$	$o(S_2) = 1$
$S_3 = (110*)$	$o(S_3) = 3$
$S_4 = (**001*110)$	$o(S_4) = 6$
$S_5 = (****00**0*)$	$o(S_5) = 3$
$S_6 = (11101**001)$	$o(S_6) = 8$

Se utiliza para calcular la probabilidad de supervivencia del esquema con relación al operador mutación

Longitud característica de un esquema, $\delta(S)$

Distancia en dígitos entre las posiciones fijas extremas

$S_1 = (*01*)$	$\delta(S_1) = 3 - 2 = 1$
$S_2 = (**1)$	$\delta(S_2) = 4 - 4 = 0$
$S_3 = (110*)$	$\delta(S_3) = 3 - 1 = 2$
$S_4 = (**001*110)$	$\delta(S_4) = 10 - 4 = 6$
$S_5 = (****00**0*)$	$\delta(S_5) = 9 - 5 = 4$
$S_6 = (11101**001)$	$\delta(S_6) = 10 - 1 = 9$

Se utiliza para calcular la probabilidad de supervivencia del esquema con relación al operador de cruce

Hipótesis de Bloques Constructivos

Esquemas cortos, de bajo orden y con un fitness alto serán elegidos, recombinados y re-elegidos para ir formando cromosomas con mayor fitness. Llamamos a estos esquemas ***Bloques Constructivos***

Al usarlos se reduce la complejidad del problema

En lugar de probar todas las combinaciones posibles de cromosomas, se van construyendo cada vez mejores cromosomas usando las mejores soluciones parciales obtenidas