



UAEM

Universidad Autónoma
del Estado de México



Algoritmos Genéticos

Centro Universitario UAEM Valle de México

Unidad 2 Algoritmo Genético

Maestría en Ciencias de la Computación (MASCO)

Ph. D. Victor Manuel Landassuri Moreno

vmlandassurim@uaemex.mx

landassuri@gmail.com



2.1 Algoritmo Genético simple

Pseudocódigo generar de un AE o AG

```
Funcion Alg_Gen()
{
    Tpoblacion pob;           // población
    Tparametros param        // tamaño de pob

    obtener_parametros(param);
    pob = poblacion_inicial();
    evaluación(pob, tam_pob, pos_mejor, sumAdaptacion);

    // bucle de evolución
    mientras no se alcanza condición de terminación hacer{
        selección(pob, parametros);
        reproducción(pob, parametros);
        evaluación(pob, parametros, pos_mejor, sumAdaptación);
    }
    devolver pob[pos_mejor];
}
```

Descripción del algoritmo

- La ejecución del algoritmo requiere de diversos parámetros iniciales:
 - Tamaño de la población. Define el comportamiento promedio
 - % de cruzamiento
 - % de mutación
 - No. de individuos de elite
 - Entre otros.
- Dados estos, se genera la población inicial de individuos
 - Cada uno es un potencial candidato a la solución

Descripción del algoritmo

- Posteriormente se evalúa la adaptabilidad de cada y se entra en el bucle principal del algoritmo, donde se seleccionarán, cruzarán y mutarán a los individuos para generar una nueva población.

Características

- *Estocásticos.* Dos ejecuciones distintas pueden dar dos soluciones distintas.
- *De búsqueda múltiple*
- *La optimización es función de la representación de los datos: una buena codificación & una mala*
- *Son intrínsecamente paralelos: independientemente que se programen en paralelo o no, se buscan varias soluciones en paralelo*

2.2 Representación

Representación Binaria

- Sea **b** una cadena binaria
- **b** representa puntos **x** e el espacio de búsqueda del problema.
- Desde el aspecto biológico
 - **b** es el genotipo
 - **x** es el fenotipo
- Un **gen** es la codificación de una característica determinada de un individuo, donde puede ser cada posición de la cadena binaria.
- **Alelo** son los distintos valores que puede tomar un gen
- **Locus** se refiere a una posición determinada de la cadena

Codificación

- La codificación consiste en pasar del problema (fenotipo) a la representación binaria (genotipo)
- Esta tiene que ser capaz de poder representar todos los puntos del espacio de problema.
- Cada posición de la cadena, debe tener un significado para el problema
- Se debe escoger una codificación que permita llegar de forma eficiente al fenotipo (decodificación)
 - Ya que será una operación muy frecuente

2.3 Generación de la población inicial

Población inicial

- Los individuos de la población inicial suelen ser ceros y unos generados de forma aleatoria.
- Para el buen funcionamiento del algoritmo se debe de tener una buena diversidad al inicio, o variedad en los individuos generados
 - Para poder explorar todas las zonas del espacio de búsqueda

2.4 Grado de Adaptación de los individuos

Adaptación

- La evolución de la población, depende de la calidad relativa de los individuos
 - Los cuales compiten por aumentar su presencia en la población
 - Participando en las operaciones de reproducción
- Es común que los problemas se expresen como la optimización de una función matemática explícita
 - Ahí, la función de adaptación coincide con a la función a optimizar.

Adaptación

- Sin embargo, a veces se hacen transformaciones a la función a optimizar, o función de evaluación
 - $g(x)$
- Para transformarla en una función de adaptación adecuada
 - $f(x)$
- Se denomina adaptación bruta de un individuo x a $g(x)$ y adaptación a $f(x)$

2.5 Condición de terminación

Terminación

- Es necesario establecer el criterio de paro del algoritmo:
 - Por generaciones (iteraciones del bucle principal)
 - Alcanzar una solución determinada
 - La mayor parte de la población a convergido a una forma similar
 - Careciendo de suficiente diversidad para que tenga sentido la evolución
 - Por tiempo
 - La combinación de algunas de ellas, etc.

2.6 El proceso de Selección

Selección

- La población se somete a un proceso de selección, el cual puede ser:
 - Selección proporcional o ruleta
 - Muestreo estocástico universal
 - Selección por torneo
 - Muestreo por restos
 - Elitismo
 - Etc.
- A continuación se explican brevemente estos:

Selección proporcional o ruleta

- La idea es que el individuo más apto, tenga más probabilidades de pasar sus genes a la siguiente generación.
- En otras palabras: generar exponencialmente más copias del mejor individuo con respecto de peor.
- Se genera una probabilidad p_i del individuo i , la cual es relativa a su adaptabilidad:

$$p_i = \frac{f_i}{\sum_j f_j}.$$

Selección proporcional o ruleta

- Después se genera un número aleatorio (con distribución uniforme) en el intervalo abierto (0,1)
 - Se define la puntuación acumulada
$$q_0 = 0$$
$$q_i = p_1 + \dots + p_i$$
 - Se genera un número aleatorio a entre (0,1)
 - Selección del individuo que cumpla
$$q_{i-1} < a < q_i$$

Esto es repetido para cada individuo que quiera ser seleccionado

Muestreo estocástico universal

- El método de la ruleta no puede ser del todo adecuado debido a la discrepancia de los datos, los cuales no son generados con un espacio igual
- El Muestreo estocástico universal, genera números igualmente espaciados, y no resolviendo por completo el problema de la ruleta, pero siendo este mucho mejor
 - Aquí el procedimiento es similar al de la ruleta
 - Se genera solo un número aleatorio
 - Después se generan k números espaciados de igual forma

Muestreo estocástico universal

- Así se tiene:

$$a_j = \frac{a + j - 1}{k}$$

- para toda $j = 1, \dots, k$

- Ejemplo:

$a = 0.4$ (aleatorio)

$k = 5$ (para generar individuos)

Muestreo estocástico universal

$$a_1 = (0.4 + (1-1)) / 5 = 0.08$$

$$a_2 = (0.4 + (2-1)) / 5 = 0.28$$

$$a_3 = (0.4 + (3-1)) / 5 = 0.48$$

$$a_4 = (0.4 + (4-1)) / 5 = 0.68$$

$$a_5 = (0.4 + (5-1)) / 5 = 0.88$$

Posteriormente el procedimiento es similar al de la ruleta

Selección por torneo

- Determinista: Aquí se selecciona a p individuos y se toma al mejor de ellos para ser seleccionado, esto se repite para todos los individuos necesarios
- p puede estar entre 2 o 3 individuos
- Probabilístico: Aquí es similar al anterior, solo que además se genera otro número aleatorio, si el número es menor a cierto umbral, se selecciona al mejor de ellos, si no, se escoge al peor

Otros métodos de selección

- Existen otros métodos de selección, como:
 - Ranking.
 - Por restos.
 - O alguna otra.
- Un buen ejercicio, sería plantear un ejemplo dado y aplicar varias formas de hacer la selección para apreciar como cambian entre ellas.

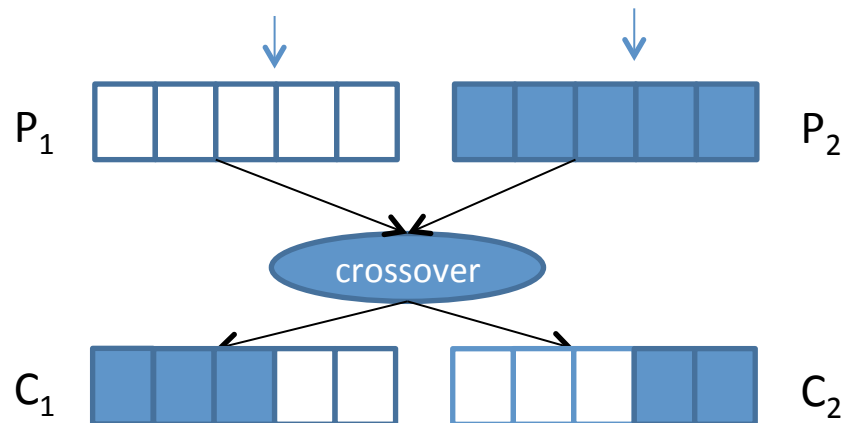
2.7 Reproducción

Reproducción

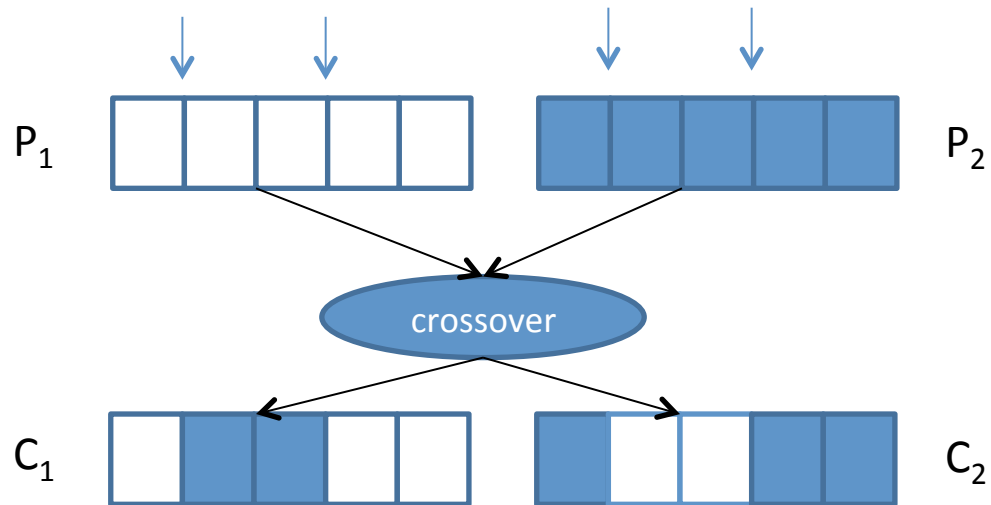
- Una vez seleccionando los individuos (padres) se procede a generar una nueva descendencia a partir de ellos (población temporal), la cual remplazara a algunos, o todos los padres de la actual generación
- Aquí existen dos operadores principales
 - Cruzamiento
 - Mutación
- Puede haber mucha variaciones de cada uno, dependiendo el problema

Cruzamiento en un punto (monopunto)

- Generar un número aleatorio para seleccionar una posición de la cadena de ambos padres y cruzar en dicho punto sus genes



Cruzamiento en dos puntos

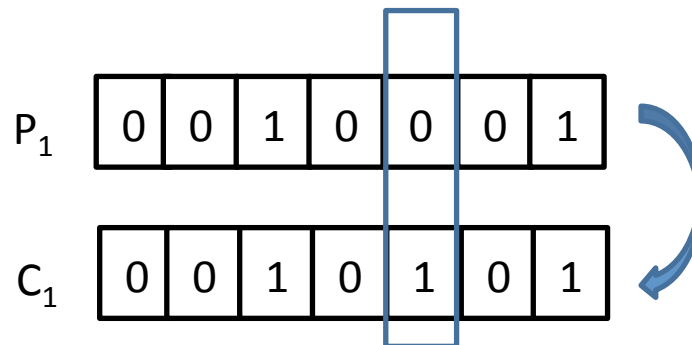


Cruzamiento

- Así, se puede ver que el cruzamiento sirve para intercambiar el material genético de los padres
 - Para generar a los hijos
- Esta es una operación fundamental en los AGs
 - Existen otros AEs que prescinden de ella

Mutación aleatoria

- La forma mas simple, es generar un número aleatorio para seleccionar una posición del genotipo, posteriormente intercambiarla, si es 1 poner 0, y viceversa



Mutación aleatoria

- La mutación tiene usualmente una probabilidad mas pequeña de ser usada, que el cruzamiento
 - Sin embargo, no hay regla y puede variar dependiendo del problema
- La Mutación sirve para aumentar la diversidad de los individuos
- Usualmente podría generar peores individuos, rompiendo la correlación de los genes evolucionados
 - Pero es funcional para escapar de posibles mínimos locales

2.8 Remplazo

Estrategia de remplazo

- **Generacional**

- Se genera una nueva población de individuos a partir de padres e hijos y se remplaza la anterior, teniendo la siguiente población $P(i+1)$

- **Steady-state** (continua)

- Se generan pocos individuos, en comparación con el tamaño total de la población, y mediante alguna técnica de inserción, se introducen en la población, haciendo $P(i+1)$

Otras formas de remplazo

- **Pueden existir diversas formas:**
- Remplazo de padres: Los hijos sustituyen a los padres
- Aleatorio: se remplazan aleatoriamente
- De los peor adaptados
- De los de una adaptabilidad similar
- Etc.

Otras referencias sugeridas

- Au, W.-H.; Chan, K.C.C.; Xin Yao, "A novel evolutionary data mining algorithm with applications to churn prediction," *IEEE Transactions on Evolutionary Computation*, vol.7, no.6, pp.532,545, Dec. 2003.
- 1997. *Handbook of Evolutionary Computation* (1st ed.). Thomas Back, David B. Fogel, and Zbigniew Michalewicz (Eds.). IOP Publ. Ltd., Bristol, UK.
- Mitchell Melanie. An introduction to Genetic Algorithms, A bradford book The MIT Press, 1999.
- Marcos Gestal Pose. Introducción a los algoritmos Genéticos, Universidad de Coruña, Depto. Tecnologías de la Información y las Comunicaciones.