

Introducción a la Computación Evolutiva

Dr. Carlos A. Coello Coello

Departamento de Computación

CINVESTAV-IPN

Av. IPN No. 2508

Col. San Pedro Zacatenco

México, D.F. 07300

email: ccoello@cs.cinvestav.mx

http: [//delta.cs.cinvestav.mx/~ccoello](http://delta.cs.cinvestav.mx/~ccoello)

Técnicas de Selección

- Selección Proporcional
- Selección mediante torneo
- Selección de Estado Uniforme

Selección Proporcional

- Este nombre describe a un grupo de esquemas de selección originalmente propuestos por Holland que eligen individuos de acuerdo a su contribución de aptitud con respecto al total de la población.

Selección Proporcional

Técnicas:

- La Ruleta
- Sobrante Estocástico
- Universal Estocástica
- Muestreo Determinístico

Selección Proporcional

Aditamentos:

- Escalamiento Sigma
- Jerarquías
- Selección de Boltzmann

La Ruleta

- Técnica propuesta por DeJong (1975), fue el método más comúnmente usado en los orígenes de los AGs.
- El algoritmo es simple, pero ineficiente: $O(n^2)$.
- El individuo menos apto puede ser seleccionado más de una vez.

La Ruleta

Implementación de DeJong:

- Calcular la suma de valores esperados T
- Repetir N veces (N es el tamaño de la población):
 - Generar un número aleatorio r entre 0.0 y T
 - Ciclar a través de los individuos de la población sumando los valores esperados hasta que la suma sea mayor o igual a r
 - El individuo que haga que esta suma exceda el límite es el seleccionado

Sobranate Estocástico

- Propuesta por Booker (1982) y Brindle (1981) para acercarse más a los valores esperados de cada individuo.
- Asignar determinísticamente las partes enteras de los valores esperados para cada individuo.
- Reduce los problemas de la ruleta, pero puede causar convergencia prematura.

Sobranate Estocástico

- Hay 2 variantes que difieren en qué hacer con las partes decimales de los valores esperados:
 - 1) **Sin Reemplazo:** Usar flip con las partes decimales para elegir los padres restantes.
 - 2) **Con Reemplazo:** Construir una ruleta con las partes decimales y usarla para seleccionar los padres faltantes.

Sobranate Estocástico

Complejidad:

- La versión con reemplazo es $O(n^2)$
- La versión sin reemplazo es $O(n)$

La más popular es la versión sin reemplazo, la cual parece ser superior a la ruleta (Booker, 1982).

Universal Estocástica

- Propuesta por Baker (1987).
- Objetivo: minimizar la mala distribución de los individuos en la población en función de sus valores esperados.
- El algoritmo es: $O(n)$

Universal Estocástica

Algoritmo:

```
ptr=Rand( ); /* Regresa un aleatorio en el rango [0,1] */  
for ( $sum = 0, i = 1; i \leq n; i++$ )  
    for ( $sum += Valesp(i, t); sum > ptr; ptr++$ )  
        Seleccionar (i);
```

Universal Estocástica

Problemas:

- Puede ocasionar convergencia prematura.
- Hace que los individuos más aptos se multipliquen muy rápidamente.
- No resuelve el problema más serio de la selección proporcional.

Muestreo Determinístico

Algoritmo:

- Calcular $P_{select} = f_i / \sum f$
- Calcular $Valesp_i = p_{select} * n$
- Asignar determinísticamente la parte entera de $Valesp_i$
- Ordenar la población de acuerdo a las partes decimales (de mayor a menor)
- Obtener los padres faltantes de la parte superior de la lista.

Muestreo Determinístico

- El algoritmo es $O(n)$ para la asignación determinística y es $O(n \log n)$ para la ordenación.
- Padece de los mismos problemas que el sobrante estocástico.

Escalamiento Sigma

- Es una técnica ideada para mapear la aptitud original de un individuo con su valor esperado, de manera que el AG sea menos susceptible a la convergencia prematura.
- La idea es mantener más o menos constante la **presión de selección** a lo largo del proceso evolutivo.

Escalamiento Sigma

- Usando esta técnica, el valor esperado de un individuo está en función de su aptitud, la media de la población y la desviación estándar de la población:

$$Valesp(i, t) = \begin{cases} 1 + \frac{f(i) - \bar{f}(t)}{2\sigma(t)} & \text{si } \sigma(t) \neq 0 \\ 1.0 & \text{si } \sigma(t) = 0 \end{cases} \quad (1)$$

Escalamiento Sigma

- Al inicio de una corrida, el valor alto de la desviación estándar impedirá que los mejores individuos obtengan los segmentos más grandes de la ruleta.
- Hacia el final, la desviación estándar será más baja y los individuos más aptos podrán multiplicarse más fácilmente.

Selección por Jerarquías

- Propuesta por Baker (1985) para evitar la convergencia prematura.
- No requiere escalamiento de las aptitudes.
- Alenta sobremanera la convergencia del AG.

Selección por Jerarquías

Algoritmo:

- Ordenar (o jerarquizar) la población con base en su aptitud, de 1 a N (donde 1 representa al menos apto).
- Elegir Max ($1 \leq Max \leq 2$)
- Calcular $Min = 2 - Max$

Selección por Jerarquías

Algoritmo:

- El valor esperado de cada individuo será:

$$Valesp(i, t) = Min + (Max - Min) \frac{jerarquia(i, t) - 1}{N - 1}$$

- Usar selección proporcional aplicando los valores esperados obtenidos.

Selección por Jerarquías

- Util cuando la función tiene ruido (por ejemplo, cuando hay una variable aleatoria).
- Existen otros métodos de asignación de jerarquías además del lineal (p.ej. exponencial).
- Su complejidad es $O(n \log n)$ + tiempo de selección.
- Diluye la presión de selección, por lo que causa convergencia lenta.

Selección de Boltzmann

- Basada en el recocido simulado: usar una función de variación de “temperatura” que controle la presión de selección.
- 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

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

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

donde: T es la temperatura y $\langle \rangle^t$ denota el promedio de la población en la generación t .

Selección de Boltzmann

- Se ha utilizado más para optimización multimodal y multiobjetivo (formación de nichos).
- Existen pruebas de convergencia de la técnica hacia el óptimo global.
- Tiene el inconveniente de requerir la definición de la función de variación de temperatura.

Selección mediante Torneo

- Los métodos de selección proporcional requieren de 2 pasos por generación del AG:
 - 1) Calcular la aptitud media (y, la desviación estándar si se usa escalamiento sigma).
 - 2) Calcular el valor esperado de cada individuo.

Selección mediante Torneo

- El uso de jerarquías requiere que se ordene toda la población (una operación cuyo costo puede volverse significativo en poblaciones grandes).
- La selección mediante torneo es similar al uso de jerarquías en términos de la presión de selección, pero es computacionalmente más eficiente y más fácil de paralelizarse.

Selección mediante Torneo

- Propuesta por Wetzel y estudiada en la tesis doctoral de Brindle (1981).
- La idea básica del método es seleccionar con base en comparaciones directas de los individuos.
- Hay 2 versiones:
 - 1) Determinística
 - 2) Probabilística

Selección mediante Torneo

Algoritmo (versión determinística):

- Barajar los individuos de la población
- Escoger un número P de individuos (típicamente 2)
- Compararlos con base en su aptitud
- El ganador del “torneo” es el individuo más apto
- Se debe barajar la población un total de P veces para seleccionar N padres.

Selección mediante Torneo

Algoritmo (versión probabilística):

- La versión probabilística es idéntica excepto por el paso en el que se escoge al ganador. En vez de seleccionar siempre al individuo con aptitud más alta, se aplica $\text{flip}(p)$ y si el resultado es cierto, se selecciona al más apto. De lo contrario, se selecciona al menos apto.
- La probabilidad p permanece fija durante todo el proceso evolutivo y se escoge de manera que:

$$0.5 \leq p \leq 1$$

Selección mediante Torneo

Análisis:

- La versión determinística garantiza que el mejor individuo será seleccionado P veces.
- Cada competencia requiere la selección aleatoria de un número constante de individuos de la población. La comparación entre estos individuos puede realizarse en tiempo constante. Se requieren n competencias de este tipo para completar una generación. Por tanto, el algoritmo es $O(n)$.

Selección mediante Torneo

Análisis:

- Técnica eficiente y fácil de implementar.
- No requiere escalamiento de la función de aptitud (usa comparaciones directas).
- Puede introducir una presión de selección muy alta porque a los individuos menos aptos no se les da oportunidad de sobrevivir.

Selección mediante Torneo

- Si se usa $t_{torneo} = 1$, se produce una caminata aleatoria con una presión de selección muy baja.
- Si se usa $t_{torneo} = \infty$, la selección se vuelve completamente determinística.
- Si se usa $t_{torneo} \geq 10$, la selección se considera **dura**.
- Si se usa $2 \leq t_{torneo} \leq 5$, la selección se considera **blanda**.

Selección de Estado Uniforme

- Propuesta por Whitley (1989).
- Se usa en AGs no generacionales, en los cuales sólo unos cuantos individuos son reemplazados en cada generación (los menos aptos).
- Esta técnica suele usarse cuando se evolucionan sistemas basados en reglas (p.ej., sistemas de clasificadores) en los que el aprendizaje es incremental.

Selección de Estado Uniforme

- Esta técnica es útil cuando los miembros de la población resuelven colectivamente (y no de manera individual) un problema.
- Asimismo, los AGs no generacionales se usan cuando es importante “recordar” lo que se ha aprendido antes.

Selección de Estado Uniforme

Algoritmo:

- Seleccionar de G (población original) R individuos ($1 \leq R \leq M$) de entre los más aptos.
Normalmente, $R = 1$, o $R = 2$.
- Efectuar cruce y mutación a los R individuos seleccionados.
Llamaremos H a sus hijos.
- Elegir al mejor individuo en H (o a los S mejores).
- Reemplazar los S peores individuos de G por los S mejores individuos de H .

Selección de Estado Uniforme

Análisis:

- Técnica especializada de selección.
- Su complejidad (en la variante incluida en GENITOR) es $O(n \log n)$
- Los *AGs* no generacionales no son muy comunes en aplicaciones de optimización, aunque sí pueden utilizarse.

Selección más (+)

Es también posible en un algoritmo genético usar una selección más (+) como en las estrategias evolutivas. Esta selección consiste en unir la población de padres con la de hijos y seleccionar la mejor mitad de ellos.

Este tipo de selección resulta particularmente útil para resolver problemas de optimización global.

Brecha Generacional

- Muy ligada a la selección de estado uniforme se encuentra el concepto de **brecha generacional** (*generation gap*).
- Es importante reconocer en primer término que las poblaciones pueden ser “no traslapables” (*nonoverlapping*) o “traslapables” (*overlapping*).

Brecha Generacional

- Una población **no traslapable** es aquella en la que los padres nunca compiten contra sus hijos. Es decir, toda la población de padres es siempre reemplazada por la población de hijos.
- En una población **traslapable**, los padres compiten contra sus hijos.

Brecha Generacional

- Se denomina **brecha generacional** a la cantidad de traslape existente entre padres e hijos. Una brecha generacional grande implica poco (o ningún) traslape poblacional.
- Históricamente, la programación evolutiva y las estrategias evolutivas han usado poblaciones *traslapables*, mientras que los AGs han usado poblaciones *no traslapables*.

Brecha Generacional

- De Jong (1975) parece haber sido el primero en estudiar AGs con poblaciones traslapables.
- De Jong sugirió que las ventajas de las poblaciones traslapables se diluían debido a los efectos negativos del *desvío genético*.
- Más tarde, Grefenstette (1986) confirmaría que una brecha generacional mayor parecía mejorar el desempeño del AG.

Brecha Generacional

- Los primeros experimentos con los *sistemas de clasificadores*, confirmarían, sin embargo, un comportamiento exactamente opuesto (Holland & Reitman, 1978).
- En los *sistemas de clasificadores*, el desempeño del AG parecía degradarse conforme se aumentaba la brecha generacional.
- Algunos investigadores atribuyen los resultados de De Jong y Grefenstette al uso de poblaciones pequeñas.

Brecha Generacional

- Los AGs tradicionales siguen usando, sin embargo, poblaciones no traslapables.
- Los algoritmos evolutivos de estado uniforme son aquellos en los que la población es traslapable.
- Normalmente, sólo uno o dos hijos se producen en cada iteración de un AE de estado uniforme.