
INSTITUTO POLITÉCNICO NACIONAL

Centro de Investigación en Computación



ASIGNATURA:

Metaheurísticas

Actividad #21:

Guía Taller No.10

"Operadores de selección en Algoritmos Genéticos"

PROFESORA:

Dra. Yenny Villuendas Rey

PRESENTA:

Juan René Hernández Sánchez

Adriana Montserrat García Carrillo



Centro de Investigación
en Computación
Instituto Politécnico Nacional

1. Introducción

Los algoritmos genéticos (AG) son algoritmos inspirados en la selección natural que han logrado semejar todo el proceso evolutivo de los vivos para resolver problemas de optimización, búsqueda y aprendizaje en las máquinas

La teoría de los algoritmos genéticos fue desarrollada por John H. Holland, sus colegas y sus estudiantes en la Universidad de Michigan, en un proyecto que buscaba explicar el proceso de adaptación natural de los sistemas y diseñar sistemas artificiales que conservaran los más importantes mecanismos de los sistemas naturales. Casi todos los conceptos de los de los algoritmos genéticos están basados en conceptos de biología y genética.

Las características de los seres vivos se encuentran registradas en los cromosomas y los genes, lo que hace que cada individuo sea diferente dentro de una población. En algoritmos genéticos los cromosomas se representan mediante cadenas (strings) y cada posición de la cadena representa un gen.

El primer paso al desarrollar un algoritmo genético es, entonces, codificar el parámetro deseado como una cadena y usando un alfabeto. La longitud de un cromosoma está determinada por el número de genes que lo conforman. El valor que pueda tomar cada gen se conoce como alelo. Si utiliza el alfabeto binario cada gen sólo tomar dos valores, uno o cero.

El conjunto ordenado de genes de un cromosoma se conoce como genotipo y las características del individuo. Por en el cromosoma el genotipo y representa las características del individuo. Por ejemplo, en el cromosoma 10010, el genotipo es {1, 0, 0, 1, 0}.

De acuerdo con los valores que tengan los genes de una persona podremos decir que es alta o baja, de ojos azules o verdes, de cabello liso o rizado, etc. Esto se denomina fenotipo, y en los algoritmos genéticos corresponde a la decodificación de un cromosoma. El fenotipo para el cromosoma 10010 será su representación en decimal (18).

En los sistemas naturales, un individuo es un ser que caracteriza su propia especie. En los algoritmos un individuo es un solo cromosoma, que en conjunto forman una población de cromosomas.

Cada cromosoma es una solución, buena o mala, del problema. A medida que el algoritmo genético se va desarrollando las soluciones se acercan cada vez más a la óptima. A cada cromosoma se le asigna un valor determinado como una medida de qué tan buena es la solución. Este valor se obtiene evaluando cada cromosoma por medio de una función llamada función objetivo.

Simulando los sistemas de selección natural a los que se encuentran sometidos los organismos vivos, los algoritmos genéticos utilizan tres operadores básicos para solucionar el problema: Selección, Cruzamiento y Mutación.

2. Desarrollo

Asignatura: Metaheurísticas

Actividad No.21

Guía Taller No.10

Título: Operadores de selección en Algoritmos Genéticos

Contenido:

- Métodos heurísticos de solución de problemas.
- Algoritmos Genéticos
- Operadores de selección

Objetivo: Implementar operadores de selección para algoritmos genéticos, en lenguajes de alto nivel, para la solución de problemas de la profesión.

1. Enuncie las ventajas y desventajas de los Algoritmos Genéticos

- **Ventajas:**

- a. Son muy versátiles
- b. Cuando se usan para problemas de optimización resultan menos afectados por los máximos locales (falsas soluciones) que las técnicas tradicionales.
- c. Usan operadores probabilísticos, en vez de los típicos operadores determinísticos de las otras técnicas, lo que les ayuda a distribuir mejor los puntos de búsqueda
- d. Pueden realizar un procesamiento paralelo de las soluciones, contrario a los algoritmos convencionales que trabajan en serie y sólo pueden explorar soluciones en una dirección al mismo tiempo.
- e. Usan la información que les proporciona la función objetivo y no requieren de otros métodos o conocimientos auxiliares (no necesitan conocimientos específicos sobre el problema que intentan resolver). Esta característica hace que se puedan adaptar más fácilmente a los diferentes problemas.

- **Desventajas:**

- a. Tienen muchos parámetros a configurar
- b. Pueden tardar mucho en converger, o no converger en absoluto, dependiendo en cierta medida de los parámetros que se utilicen, el tamaño de la población, el número de generaciones, etc.
- c. Pueden converger prematuramente debido a una serie de problemas de diversa índole. La convergencia prematura y el utilizar el tamaño de poblaciones muy pequeñas podría incluso extinguir a dichas poblaciones o mermar la diversidad de la población.
- d. Se debe especificar un lenguaje robusto para especificar las soluciones candidatas, es decir el lenguaje debe ser capaz de tolerar cambios aleatorios que no produzcan constantemente errores fatales o resultados sin sentido.

- e. Se debe definir la función aptitud (función objetivo de cada individuo) de manera cuidadosa, ya que los AG pueden trabajar con funciones aptitud engañosas provocando que la calidad de la solución sea equivocada.

2. Diga las diferencias entre fenotipo y genotipo.

Diferencias:

a. El genotipo comprende a los genes y el fenotipo a la expresión de los componentes.

El genotipo está almacenado como ADN, en secuencias creadas por cuatro bases nucleicas. La combinación de estas bases, acomodadas en pares complementarios, crea los genes. Pero el fenotipo se expresa en una cantidad enorme de componentes. Pueden ir desde enzimas, color de pelo, conducta reproductiva, tamaño, tiempo de vida, ciclo circadiano, entre muchas otras.

b. El genotipo es único, el fenotipo puede ser igual.

Puede que existan fenotipos iguales para genotipos diferentes. El genotipo comprende a la información genética que representa un organismo, tanto de genes recesivos como dominantes. Los fenotipos muestran solamente los genes heterocigotos dominantes o recesivos.

c. El genotipo es transmitido, el fenotipo es expresado

Los organismos reciben de sus padres la información genética, con secuencias de ambas partes seleccionadas por recombinación genética. Por otro lado, el fenotipo y su expresión es tarea del organismo hijo.

d. El genotipo no es visible, el fenotipo sí

El genotipo contiene la información fundamental empaquetada en forma de ADN. Esta información debe ser expresada para que pueda actuar dentro del tiempo y espacio. El fenotipo es la expresión del genotipo, que, aunque no sea visible como tal, atendiendo a lo que planteamos como caracteres fenotípicos anteriormente, sí es tangible y está descodificado.

e. El genotipo no puede ser modificado por el ambiente, el fenotipo sí

Si bien el genotipo tiene determinadas características que van a ser expresadas, por ejemplo, en rasgos físicos la complexión o el color de piel, el entorno al que se somete el fenotipo puede ser modificado, como por ejemplo si un individuo se encuentra en estado sedentario, su estructura física cambiará.

f. El fenotipo depende del genotipo, pero no viceversa

El fenotipo es expresado por el genotipo, no puede haber características que no hayan sido previamente determinadas por el genotipo. El fenotipo no puede dictar cómo actuarán los genes.

3. Enuncie las semejanzas y diferencias en los modelos generacional y estacionario.

- **Semejanzas**

Modelo generacional: Existe la posibilidad de preservar al mejor individuo obtenido en cada generación, teniendo así una configuración elitista.

Modelo estacionario: Normalmente, los descendientes reemplazan a los peores individuos de la población anterior. En este caso, el modelo sería elitista ya que se preservaría al mejor individuo obtenido en cada generación.

- **Diferencias:**

Modelo generacional: Durante cada iteración se crea una población completa con nuevos individuos. La nueva población reemplaza directamente a la antigua.

Modelo estacionario: Durante cada iteración se reemplaza solamente una parte de la población, por lo que no sólo se deben seleccionar los individuos a ser padres sino también los individuos de la población que serán reemplazados por los descendientes, para así mantener el tamaño de dicha población. Por tanto, en este modelo se introduce una nueva probabilidad, denominada probabilidad de reemplazamiento generacional o de selección que indica la proporción de la población que se reemplaza en cada iteración.

4. Explique el funcionamiento del operador de selección por muestreo aleatorio universal.

El muestreo aleatorio universal es un procedimiento similar al muestreo por ruleta, pero en este caso se genera un solo número aleatorio y a partir de él se obtienen los k individuos que se requieren. Este método fue desarrollado por Baker en 1987 con el fin de corregir algunos de los problemas de la selección por ruleta. Este algoritmo es simple y eficiente. Dado un conjunto de k individuos y sus valores objetivos asociados, se ordenan en una ruleta considerando que el tamaño de los cortes asignados a cada individuo sea proporcional al valor objetivo (como en el algoritmo de ruleta). Después, una segunda ruleta, es marcada con k marcadores igualmente espaciados entre sí, donde k es el número de selecciones que se desean efectuar. Por último, se gira la ruleta y se selecciona un individuo por cada marcador. Las posiciones de los marcadores indican los individuos seleccionados.

5. Explique el funcionamiento del operador de selección por torneo.

Todos los métodos de selección discutidos anteriormente consideran información global, es decir, el valor de fitness relativa o el rango. A veces sólo está disponible la información local, es decir, cuál es la mejor en un grupo pequeño. Entonces la selección de torneos es bastante útil.

Para implementar la selección de torneos, solo necesitamos elegir k individuos al azar con reemplazo y comparar los valores de aptitud de estos k individuos, que es el torneo. El

mejor gana el torneo y se selecciona en el grupo de apareamiento. Repita el proceso anterior hasta que se hayan seleccionado individuos de gran tamaño. El k se denomina tamaño del torneo, que controla la presión selectiva. La selección de torneos más utilizada es la selección de torneos binaria, en la que $k = 2$.

- La selección de torneos sólo utiliza información local.
- La selección de torneos es muy fácil de implementar y su complejidad de tiempo es pequeña.
- La selección de torneos se puede implementar fácilmente en un entorno paralelo.

Para disminuir el error selectivo, Sokolov y Whitley sugirieron una selección de torneos imparcial y posteriormente le hicieron mejoras.

Se hacen k permutaciones del conjunto, para el tamaño de torneo k , cada individuo se comparará con otros k veces en una selección de torneo imparcial, lo que garantiza que el mejor individuo tendrá k copias en el grupo de apareamiento y el peor no será seleccionado.

Goldberg añadió probabilidad a la selección de torneos y sugirió la selección de torneos de Boltzmann.

6. Explique el funcionamiento del operador de selección proporcional.

Si un proceso de selección puede cumplir el requisito de que el número esperado de un individuo sea proporcional a su valor de aptitud relativo, es decir, a la probabilidad de ser seleccionado, lo llamamos selección proporcional. RWS es una selección proporcional típica. En el RWS, todos los resultados se determinan aleatoriamente, lo que significa que cada individuo puede perderse sea cual sea su valor de aptitud relativo. Así que el RWS tiene errores de selección, por lo que podríamos decir que el RWS sufre un sesgo selectivo.

El RWS es igual al operador de selección por ruleta (tiene sólo una flecha), pero añade el ordenamiento de su población. Cada vez la flecha gira $\text{rand} \times 2\pi$ ángulos, donde $\text{rand} \sim U(0,1)$. Si la flecha se detiene en el área j , decimos que el individuo j está seleccionado. Por lo tanto, RWS se lleva a cabo de una manera serial.

7. Explique el funcionamiento del operador de selección por ruleta.

La diferencia entre la ruleta real y nuestra selección es que los agujeros de la rueda de la ruleta en nuestra selección tienen diferentes tamaños. Cuanto más grande es la probabilidad p_i , más grande es el agujero que tiene el individuo i . Entonces la bola podría caer en el hoyo con mayor probabilidad. El tamaño del agujero para i es proporcional a p_i .

El ángulo central del individuo i es $2\pi p_i$. Queremos que la flecha comience a girar en el sentido de las agujas del reloj y se detenga aleatoriamente con un ángulo central de $2\pi \cdot \text{rand}$, donde $\text{rand} \sim U(0,1)$. Si la flecha se detiene en un sector, el individuo correspondiente se selecciona en el grupo de apareamiento. Está claro que las personas más en forma tienen más posibilidades de ser seleccionadas.

8. Explique el funcionamiento del operador de selección por emparejado variado inverso (NAM).

Podemos ajustar la diversidad de la población usando una restricción de apareamiento. Si dos individuos similares (genotipo o fenotipo) tienen más posibilidades de reproducirse, este tipo se denomina apareamiento selectivo positivo.

Por el contrario, el apareamiento selectivo negativo promueve el cruce entre padres diferentes. El apareamiento selectivo negativo podría implementarse comparando la distancia entre dos individuos con un umbral predefinido.

Si la distancia es mayor que el umbral, puede ocurrir un cruce y viceversa. Es bastante obvio que el apareamiento selectivo negativo promueve la exploración. Por lo tanto, la diversidad de la población se podría mantener de esta manera. Por otro lado, el apareamiento selectivo positivo promueve la explotación y empuja a la población hacia la convergencia.

9. Proponga las estructuras de datos necesarias para la implementación de los operadores de selección.

- Listas

10. Implemente tres de los operadores antes mencionados.

- Operador de selección por torneo

```
import random

k = 2
popsize = 10
poblacion = []
torneo = []
apareamiento = []

for i in range(popsize):
    poblacion.append(random.uniform(0,20))

print(poblacion)

for i in range(k):
    seleccion = random.randint(0,len(poblacion)-1)
    torneo.append(poblacion[seleccion])
    poblacion.pop(seleccion)

print(poblacion)
```

```
print(torneo)

apareamiento.append(max(torneo))
print(apareamiento)
```

- **Operador de selección por ruleta**

```
import random
import math

popsize = 10
poblacion = []
apareamiento = []
probabilidad = []
probTotal = 0
flecha = random.random()
ruleta=[0]

for i in range(popsize):
    poblacion.append(random.uniform(0,20))
    probTotal += poblacion[i]

print(poblacion)

for i in range(popsize):
    probabilidad.append(poblacion[i]/probTotal)

print(probabilidad)

for i in range(popsize):
    ruleta.append(probabilidad[i]+ruleta[i])

print(ruleta)

ruleta.append(flecha)
ruleta.sort()
posicion = ruleta.index(flecha)

apareamiento.append(poblacion[posicion-1])
print(apareamiento)
```


- **Operador de selección proporcional**

```
import random
import math

popsize = 10
poblacion = []
apareamiento = []
probabilidad = []
probTotal = 0
flecha = random.random()
ruleta=[0]

for i in range(popsize):
    poblacion.append(random.uniform(0,20))
    probTotal += poblacion[i]

poblacion.sort()
print(poblacion)

for i in range(popsize):
    probabilidad.append(poblacion[i]/probTotal)

print(probabilidad)

for i in range(popsize):
    ruleta.append(probabilidad[i]+ruleta[i])

print(ruleta)

ruleta.append(flecha)
ruleta.sort()
posicion = ruleta.index(flecha)

apareamiento.append(poblacion[posicion-1])
print(apareamiento)
```

3. Conclusiones

Los algoritmos genéticos presentan importantes ventajas sobre los algoritmos tradicionales, como son: su capacidad para trabajar con varios puntos simultáneamente, abarcando así un espacio mayor de búsqueda (paralelismo), no requieren manipular directamente los parámetros del problema y poseen una gran facilidad para adaptarse a diferentes tipos de problemas.

A su vez, los algoritmos genéticos también presentan algunas desventajas como son: la dificultad para encontrar una representación apropiada de los parámetros del problema o para hallar una función objetivo adecuada.

Los AG algoritmos no aseguran encontrar la solución óptima, pero en muchos casos pueden proporcionar soluciones bastante adecuadas y que por otros métodos hubiera sido imposible encontrar. Dicho lo anterior, estos algoritmos representan, por todas sus características, una opción que debe ser tomada en cuenta cuando se busca resolver un problema con algún grado de complejidad.

4. Referencias

- [1] Yu & Gen. Introduction to Evolutionary Algorithms. 2010. Capítulos 1 al 3.
- [2] Burke & Kendall. Search Methodologies. 2005 Capítulo 4.
- [3] Russell & Norving. Artificial Intelligence - A Modern Approach – 1995. Capítulo 4.