Universidad Nacional Autónoma de México

Facultad de Ciencias

Tarea 04:

Introducción al algoritmo genético

Pablo A. Trinidad Paz - 419004279

Trabajo presentado como parte del curso de **Cómputo Evolutivo** impartido por el profesor **Mario Iván Jaen Márquez**.

Fecha de entrega: **Lunes 19 de Marzo de 2019**.

1. Planteamiento

Implemente el algoritmo genético para que dados los siguientes ejercicios, el algoritmo se aproxime a un individuo que cumpla las condiciones necesarias.

1. Genere un cromosoma cuyas características indiquen a una persona con piel oscura, caballo negro, altura de 1.98 metros o cercana y complexidad delgada.
2. Obtenga un cromosoma con las cualidades para ser luchador de sumo japonés [1]
3. Elija un personaje de la serie que prefiera[2] y ajuste el algoritmo para obtener un cromosoma de características similares. Debe considerar por lo menos las siguientes características:
   - Color de cabello
   - Color de piel
   - Color de ojos
   - Altura
   - Complexión

   Adjunte un enlace a la referencia del personaje del cual se está tomando la información.

   La implementación del algoritmo debe cumplir lo siguiente:

   - La población inicial debe ser aleatoria y de 100 individuos.
   - Para la probabilidad de mutación utilice 0.1.
   - Para el operador de cruza utilice el método de torneo en la elección de padres y dos puntos de cruce (two point crossover).
   - En la selección de individuos utilice ruleta.
   - El *fitness* se va definir a discreción cuidando que el algoritmo garantice o se logre aproximar considerablemente a un individuo que cumpla las características solicitadas. Se recomienda la siguiente estructura:

     Sea $\mathcal{X}$ el individuo con las características deseadas y x el individuo a evaluar. La función f de *fitness* es:

     $$f = \sum_{i=0}^{n} \left(\mathcal{X}[i] - x[i]\right)^2$$

     donde n es el número de características.
   - El cromosoma deberá considerar los valores de color de cabello, color de ojos y color de piel usando 3 valores para cada característica donde para cada valor el rango será entero $[0, 255]$; para el cabello y ojos será válido cualquier color, cuide que el valor de la tez de piel se vea natural, asimismo considere valores de altura coherentes (aproximadamente de 1 a 2.20 metros) y complexión, se deja a discreción agregar más valores al cromosoma.
   - Para graficar la imagen del individuo utilice la función indicada en el script adjunto, y para cada ejercicio adjunte la imagen del individui que mejor cumpla con los requisitos del ejercicio, su posible altura, complexión, el cromosoma, y en su caso, las características adicionales.

[1] Investigue los valores promedio de un luchador de sumo y de un hombre japonés. Adjunte sus referencia

[2] Puede ser de cualquier tipo, ya sea un show de televisión, anime o serie animada, de preferencia use características poco comunes

2. Solución

2.1. Implementación del algoritmo genético

En ésta sección abordaremos las múltiples estrategias utilizadas durante la implementación del algoritmo genético.

2.1.1. Representación

Con el objetivo de obtener una representación binaria del cromosoma se acordaron las siguientes abstracciones de cada una de las características:

1. **Color**: Representado por un valor RGB de tal forma que color $\in \{0, 1, ..., 255\}^L,\ L = 3$. Las características que usarán éste tipo de valor serán el color de piel, el color del cabello y el color de ojos.
2. **Altura**: Medida en centímetros y válida en el rango $[0, 255]$. El rango utilizado ignora la recomendación con el objetivo de poder utilizar 8 bits para representarla.
3. **Complexión**: Definida arbitrariamente en el rango $[0, 15]$ donde una complexión con valor $x = 0$ se considera la complexión más pequeña posible y una complexión $x = 15$ la más grande posible, es decir, la complexión crece de manera lineal de la más pequeña a la más grande conforme x se crece. La complexión puede ser representada utilizando 4 bits.

Dicho lo anterior podemos definir la representación binaria del cromosoma de un individuo como la sucesión ordenada de dichas características de forma que los genes estén distribuidos de la siguiente manera:

Distribución de genes dentro del cromosoma (Indexado en 0)		
Rango de bits	Característica	Rango en $\mathbb{Z}^+$
$[0, 23]$	Color de cabello	$[0, 255]$ cada 8 bits
$[24, 47]$	Color de ojos	$[0, 255]$ cada 8 bits
$[48, 71]$	Color de piel	$[0, 255]$ cada 8 bits
$[72, 79]$	Altura	$[0, 255]$
$[80, 83]$	Complexión	$[0, 15]$

Adicionalmente, la codificación de cada gen será mediante Gray code con el objetivo de reducir la distancia de Hamming.

2.1.2. Función objetivo

Gracias a que tenemos una representación binaria del cromosoma, que estamos utilizando Gray code para la codificación de cada valor, y que queremos que nuestras poblaciones asemejen lo más posible a un individuo dado, podemos establecer la función objetivo f como la distancia de Hamming entre la cadena del individuo ideal y el individuo a evaluar. Además, nótese que únicamente estaríamos realizando la decodificación al obtener un resultado aceptable. Si quisiéramos que la función f esté en el rango $[0, 1]$, entonces podemos definirla como:

$$f(x) = 1 - \frac{h(\mathcal{X}, x)}{84}$$

Donde $\mathcal{X}$ es el individuo con características deseadas y $h(x, y)$ es la función que calcula la distancia de Hamming entre dos cadenas x y y.

2.1.3. Inicialización

La inicialización de la población consiste en la creación de cadenas binarias aleatorias de longitud $L = 84$. La población inicial tendrá 100 individuos.

2.1.4. Mecanismo de selección de padres

El método utilizado para la selección de de padres será Torneo ya que nos evitará la necesidad de contar con una noción del *fitness* promedio de la población. La selección por torneo selecciona λ mejores miembros de un subconjunto aleatorio de μ elementos hasta completar la cantidad deseado de padres ρ.

En nuestro caso particular deseamos obtener $\rho = 6$ padres en cada iteración del algoritmo genético y lo haremos a través de conjuntos de $\mu = 33$ elementos escogiendo $\lambda = 2$ elementos en cada ronda del torneo con reemplazo.

2.1.5. Operador de cruza

El método utilizado será n-point crossover con $n = 2$. En nuestro caso, las parejas de cruzamiento serán unidas en el mismo orden en el que hayan sido seleccionadas durante el torneo.

2.1.6. Operador de mutación

Tras la cruza, la mutación sucederá en los hijos resultantes invirtiendo cada bit de su representación genotípica con probabilidad $p_m = 0.1$.

2.1.7. Selección de individuos

La selección de individuos seguirá el método *fitness proportional selection* (también conocido como ruleta) usando **sigma scaling** descrito en [1] para resolver los problemas de convergencia prematura. La nueva función objetivo para **sigma scaling** será la siguiente

$$f'(x) = \max(f(x) - (\bar{f} - c \cdot \sigma_f), 0)$$

Donde $f(x)$ es el fitness original del individuo, $\bar{f}$ es el promedio del fitness de todos los individuos, σ_f la desviación estándar y c un valor constante, en nuestro caso $c = 2$.

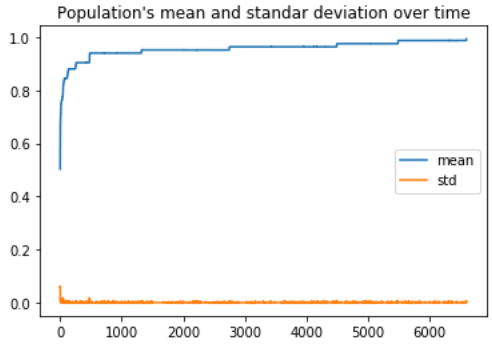
El método *FPS* asignará la probabilidad $P_{FPS}(i)$ a todos los individuos y *girará la ruleta* N veces para obtener N individuos. En nuetro caso $N = 100$ y

$$P_{FPS}(i) = \frac{f'_i}{\sum_{j=1}^{N} f'_j}$$

2.2. Resultados

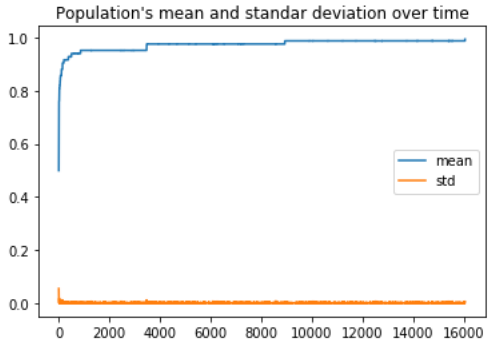
1. Ejercicio 1. **Valores del fenotipo y datos de ejecución:**

```
Starting SGA with:
        Mutation probability = 0.1
        Max trials = 100000000
        Precision = 0.01
        Population size = 100
        Role model attributes:
                height: 198
                body_type: 4
                hair_color: (197, 97, 12)
                eyes_color: (197, 80, 35)
                skin_color: (37, 27, 33)
                genome: 10100111010100010000101010100111011110000011001000110111000101100011000110100101011
                fitness: 1.0
...
SGA converged after 6594 generations
Results:
        Last population mean = 0.9941666666666665
        Last population std = 0.00595119035711902
        Best fitted individual:
                height: 198
                body_type: 4
                hair_color: (197, 97, 12)
                eyes_color: (197, 80, 35)
                skin_color: (37, 27, 33)
                genome: 10100111010100010000101010100111011110000011001000110111000101100011000110100101011
                fitness: 1.0
        Results over time:
```



2. Ejercicio 2. Se consideró la complexión de tipo 15, color de piel (248, 230, 254) y color de cabello y ojos de (99, 50, 110) y la altura de 173 cm[3].

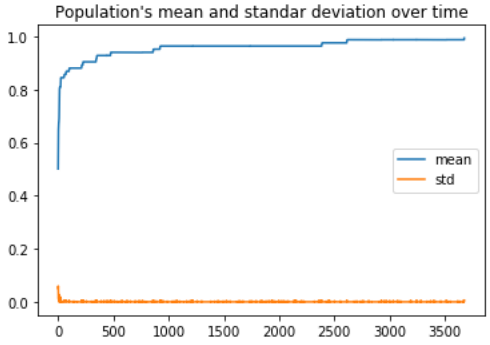
```
Starting SGA with:
        Mutation probability = 0.1
        Max trials = 100000000
        Precision = 0.01
        Population size = 100
        Role model attributes:
                height: 173
                body_type: 14
                hair_color: (99, 50, 110)
                eyes_color: (99, 50, 110)
                skin_color: (248, 230, 254)
                genome: 01010010001010110101100101010010001010110101100110000100100101011000000111111010111001
                fitness: 1.0
...
SGA converged after 16052 generations
Results:
        Last population mean = 0.9942857142857142
        Last population std = 0.005947617141331787
        Best fitted individual:
                height: 173
                body_type: 14
                hair_color: (99, 50, 110)
                eyes_color: (99, 50, 110)
                skin_color: (248, 230, 254)
                genome: 01010010001010110101100101010010001010110101100110000100100101011000000111111010111001
                fitness: 1.0
        Results over time:
```



[3]https://en.wikipedia.org/wiki/Sumo#Minimum_height_requirement

3. Ejercicio 3. Homer Simpson. Color de piel (254, 217, 15), color de cabello y ojos (202, 171, 119) complexión de tipo 10 y altura 183 cm [4]

```
Starting SGA with:
        Mutation probability = 0.1
        Max trials = 100000000
        Precision = 0.01
        Population size = 100
        Role model attributes:
                height: 183
                body_type: 10
                hair_color: (202, 171, 119)
                eyes_color: (202, 171, 119)
                skin_color: (254, 217, 15)
                genome: 10101111111111100100110010101111111111110010011001000000110110101000010001110110011111
                fitness: 1.0
...
SGA converged after 3673 generations
Results:
        Last population mean = 0.9941666666666665
        Last population std = 0.005951190357119021
        Best fitted individual:
                height: 183
                body_type: 10
                hair_color: (202, 171, 119)
                eyes_color: (202, 171, 119)
                skin_color: (254, 217, 15)
                genome: 10101111111111100100110010101111111111110010011001000000110110101000010001110110011111
                fitness: 1.0
        Results over time:
```



[4] https://simpsons.fandom.com/wiki/Homer_Simpson

Referencias

[1] Eiben, A.E., Smith, J.E. (2015). *Introduction to Evolutionary Computing*. Amsterdam, Netherlands: Springer 2nd edition. 287 p.