COMPUTACIÓN BIOLÓGICA

Pedro Isasi[1]

[1]Departamento de Informática
Universidad Carlos III de Madrid
Avda. de la Universidad, 30. 28911 Leganés (Madrid). Spain
email: isasi@ia.uc3m.es

Presentación

Temario

Diferencias

- Codificación con valores continuos
- No hay mutación discreta
- Nueva mutacion con $N(0, \sigma)$
- AG $\Longrightarrow$ Potencia de búsqueda en sobrecruzamiento
- EE $\Longrightarrow$ Potencia de búsqueda en mutación

EE. Tipos

- Varios tipos de Estrategias Evolutivas
  - Dependiendo del tamaño de la población. Si está compuesta por uno o más individuos
  - De la forma en que las nuevas generaciones son generadas

(1+1)-EE

- Método de escalada puro
- La población está compuesta por un solo individuo, que variará de generación en generación (a_1^t):

$$P^t = (a_1^t) = ((x_1^t, \sigma_1^t))$$

- El individuo está formado por dos vectores de números reales:
  - El vector de la codificación x^t. Codifica una solución alternativa al problema
  - El vector de varianzas σ^t. Indica lo alejada que se encuentra dicha solución, de la solución óptima

(1+1)-EE. Mutación

- Cada población en una generación t produce una población intermedia de tamaño 2, añadiendo un nuevo individuo generado a partir del existente mediante el operador de mutación:

$$P'^t = (a_1^t, a_2^t) \in I \times I$$

$$a_2^t = m(P^t) = (x_2^t, \sigma_2^t)$$

- El operador de mutación ($m(P^t)$) actúa de forma diferente con cada parte del individuo:
  - Se produce una mutación de la parte funcional, mediante una distribución Gaussiana
  - El vector de varianzas es mutado siguiendo la regla del 1/5

MUTACIÓN DE LA PARTE FUNCIONAL

- A cada elemento del vector se le aplica un variación obtenido de forma aleatoria siguiendo una distribución Gaussiana:

$$x_2^t = x_1^t + N_0(\sigma_1^t)$$

- La distribución normal $N_0(\sigma)$ generará valores cercanos al cero, tanto positivos como negativos, con una probabilidad tanto más grande cuanto menor sea el valor de σ
  - Si la σ es pequeña, es porque se está cerca de la solución, y habrá que "explotar" la solución. Se generarán valores pequeños
  - Si la σ es grande, es porque se está lejos de la solución, y habrá que "explorar" nuevas soluciones. Se generarán valores grandes

Mutación de las varianzas

- Las varianzas deben modificarse de forma que se obtengan valores grandes cuando se esté lejos de la solución, y pequeños cuando se esté cerca

- Si ultimamente se generan buenas soluciones $\Longrightarrow$ lejos de la solución óptima (la solución propuesta es facilmente mejorable)
- Si ultimamente se generan malas soluciones $\Longrightarrow$ cerca de la solución óptima (la solución propuesta es dificilmente mejorable)

Regla del 1/5

- Se almacena en ψ_s^t la proporción de veces que la población ha cambiado en las s últimas generaciones

Example

Si $s = 10$, estamos en la generación t y en las últimas 10 generaciones se ha mejorado 7 veces la solución, entonces $\psi_{10}^t = 0,7$

- Se aplica la regla:

$$\sigma^{t+n} = \begin{cases} c_d \sigma^t & si\ \psi_s^t < \frac{1}{5} \\ c_i \sigma^t & si\ \psi_s^t > \frac{1}{5} \\ \sigma^t & si\ \psi_s^t = \frac{1}{5} \end{cases} \quad (1)$$

Regla del 1/5

- Si el número de éxitos relativos recientes está por encima de $\frac{1}{5}$ el valor de la varianza se decrementa
- Si está por debajo se incrementa
- Es necesario que $c_d < 1$ y que $c_i > 1$
- Los valores habituales de estos dos parámetros son: $c_d = 0,82$ y $c_i = 1,18$
- Esta regla asegura la convergencia del método:
  - Si las varianzas aumentan, se están obteniendo soluciones cada vez mejores y en algún momento se encontrara el óptimo
  - Si se encuentra el óptimo, las siguientes soluciones serán peores, y la varianza decrecerá hasta cero
  - Si no se encuentra el óptimo, las soluciones no mejorarán, y pasará lo mismo
  - Cuando la varianza es cero, no se generan nuevas soluciones

Evolución

- En la generación $t+1$, se evalúan los dos individuos de la población intermedia
- Se selecciona el mejor de los dos
- Si el nuevo individuo generado es mejor que el anterior, se genera una nueva población con el generado; sino se mantiene la misma población:

$$P^{t+1} = s(P'^t) = \begin{cases} a_2^t & \text{si } f(x_1^t) \leq f(x_2^t) \\ a_1^t = P^t & \text{en caso contrario} \end{cases}$$

EE Múltiples

- Se distinguen por utilizar poblaciones con más de un individuo
- Ahora sí se podrá utilizar el operador de sobrecruzamiento, hay que definirlo para vectores de números reales
- No se podrá utilizar la mutación mediante la regla de $1/5$, no hay una estadística de mejoras para cada individuo
- Se define un nuevo operador de cruce y un nuevo operador de mutación

Cruce en EE

- Actúa de forma separada en el vector de codificación, y en el vector de varianzas
  - Vector de codificación. Cada nuevo valor se calcula como la media de los valores de los progenitores en posiciones análogas:

$$\vec{x'} = (\frac{1}{2}(x_1^1 + x_1^2), \frac{1}{2}(x_2^1 + x_2^2), \cdots, \frac{1}{2}(x_\mu^1 + x_\mu^2))$$

  - Para las varianzas se realiza una operación similar:

$$\vec{\sigma'} = (\sqrt{\sigma_1^1 + \sigma_1^2}, \sqrt{\sigma_2^1 + \sigma_2^2}, \cdots, \sqrt{\sigma_\mu^1 + \sigma_\mu^2})$$

- Se le denomina cruce uniforme. No se eligen los genes, se combinan para generar alelos nuevos

Mutación en EE-Múltiples

- El vector de codificación muta de la misma forma
- Para el vector de varianzas ya no se utiliza la regla de $1/5$, sino que se van decrementando paulatinamente siguiendo un esquema Gaussiano descrito por:

$$\sigma' = e^{N(0,\tau)\cdot\sigma}$$

EE. Evolución

- Selección por muestreo aleatorio simple, proporcional al valor de fitness
- Se puede utilizar el método de la ruleta, o cualquier otro método similar
- No se genera una población completa, los nuevos individuos sustituyen a algunos de los de la generación precedente $\Longrightarrow$ Solapamiento
- Solapamiento. Hay que definir cómo se van a insertar los nuevos individuos y cómo se van a eliminar los sobrantes. Políticas de inserción y reemplazo

Políticas de inclusión y reemplazo

Si se generan λ descendientes a partir de μ progenitores, y dependiendo de las políticas de inserción y reemplazo, habrá:

- Reemplazo por inclusión. Se unen las poblaciones de progenitores (μ), y descendientes (λ), y se seleccionan los μ mejores, los que tienen mejor valor de su función de adaptación
- Reemplazo por inserción
  - $\mu \leq \lambda$.- Se elabora una nueva población con λ individuos, y se elige a los μ mejores
  - $\mu \geq \lambda$.- Se eliminan los λ peores de entre los μ y se insertan todos los λ nuevos generados

Políticas de inclusión y reemplazo

- En el reemplazo por inclusión, tanto los padres como los descendientes compiten por ser incluidos en la nueva población. $(\mu + \lambda) - EE$
- En el reemplazo por inserción, los descendientes siempre pasan a la siguiente generación y sustituyen a los progenitores menos aptos. $(\mu, \lambda) - EE$
- Existe un procedimiento diferente dependiendo del tipo de Estrategia Evolutiva

Procedimiento (1+1)-EE

1. Generar aleatoriamente un vector de números reales y sus varianzas. Las varianzas tomarán valores grandes
2. Evaluar el individuo generado
3. Repetir hasta cumplir criterio de convergencia
   1. Generar una nueva solución a partir de la anterior mediante mutación
   2. Evaluar el individuo generado
   3. Eliminar el individuo cuyo valor de adaptación sea menor
   4. Si el individuo que queda es el nuevo, aumentar la frecuencia de éxitos, sino disminuirla
   5. Mutar el vector de varianzas de acuerdo a la regla $1/5$
4. Producir como resultado el individuo resultante

Procedimiento (μ, λ)-EE

1. Generar aleatoriamente un conjunto de vectores de números reales y sus varianzas. Las varianzas tomarán valores grandes
2. Repetir hasta cumplir criterio de convergencia
   1. Evaluar la población generada
   2. Ordenar la población de forma decreciente al fitness
   3. Repetir λ veces
      1. Seleccionar dos individuos de forma proporcional a su fitness
      2. Generar un nuevo individuo aplicando sobrecruzamiento
      3. Mutar este último individuo
   4. Generar una población intermedia con los λ descendientes generados en el paso anterior
   5. Eliminar de la población original los últimos λ elementos
   6. Generar una nueva población mediante la unión de la población original y la intermedia
3. Producir como resultado el mejor individuo de la población resultante

Procedimiento $(\mu + \lambda)$-EE

1. Generar aleatoriamente un conjunto de vectores de números reales y sus varianzas. Las varianzas tomarán valores grandes
2. Repetir hasta cumplir criterio de convergencia
   1. Evaluar la población generada
   2. Repetir λ veces
      1. Seleccionar dos individuos de forma proporcional al fitness
      2. Generar un nuevo individuo aplicando sobrecruzamiento
      3. Mutar este último individuo
   3. Añadir a la población todos los elementos generados en el paso anterior. El tamaño de la población será ahora $\mu + \lambda$
   4. Eliminar de la población los peores λ elementos
3. Producir como resultado el mejor individuo de la población resultante

Temario

Descripción del ejemplo

- Se quiere elaborar un insecticida eficaz contra moscas tsetse
- Se conocen un grupo de sustancias que son tóxicas para las moscas y a la vez inocuas para los humanos
- Se desea elaborar el insecticida mezclando cada una de las sustancias, pero se da el caso de que hay sustancias que en determinada mezcla refuerzan el poder insecticida de la mezcla, y otras que lo atenúan
- Se desconoce en qué medida y de qué manera se producen dichas interacciones
- Se desea saber cuáles de las sustancias, y en qué cantidades, hay que mezclar para que el producto resultante sea un insecticida lo más eficaz posible.

Ejemplo. Codificación del individuo

- Hay que determinar:
  - La codificación en forma de vector de valores reales
  - La función de adecuación de cada codificación
- El tamaño del individuo será el número de sustancias que se va a considerar en la mezcla, digamos 10
- Cada individuo será un vector de 10 posiciones indicando qué cantidad de la sustancia correspondiente se incluirá en la mezcla final
- Si el individuo es el vector: $I_i = \{3{,}8; 12{,}3; 0{,}0; \cdots ; 1{,}8\}$, estará codificando un insecticida preparado con: 3,8 partes de la sustancia 1, 12,3 partes de la sustancia 2, no se introduce nada de sustancia 3, y así sucesivamente

Ejemplo. Función de evaluación

- Supóngase también que se tiene una manera de determinar la eficacia del insecticida, probándolo con un conjunto de 100 moscas
- El fitness podría ser el número de moscas que el insecticida es capaz de matar en un determinado intervalo de tiempo
- La mezcla perfecta sería aquella que diera una valoración de 100 en su función de fitness

Ejemplo. Restricciones

- Los valores negativos no tienen sentido
- La función de restricciones podría ser:
  - Generar el individuo y si resultan valores negativos, eliminarlo y seguir generando nuevos hasta que uno no tenga valores negativos
  - Considerar a todos los valores negativos como 0
  - Considerar a todos los valores negativos como positivos
- El primer caso hace que el proceso pueda ser muy lento, ya que si en un determinado momento se explora en una zona del espacio de búsqueda muy próxima al cero para determinados atributos, la probabilidad de generar valores negativos aumenta, y el algoritmo pasaría mayor parte del tiempo en descartar soluciones no válidas
- Las otras dos posibilidades introducen mucha redundancia en la codificación, se aumenta inecesariamente el espacio de búsqueda

Ejemplo. Restricción en codificación

- Otra posibilidad es incluir la restricción en la codificación
- Se podría realizar un paso intermedio de normalización:
  - Si existe algún valor negativo en algún atributo del individuo, restar a todos los atributos el valor del atributo más negativo:

$$I'_i = \begin{cases} I_i & si\ \forall i\ x_i \geq 0 \\ \{x_1 - min_i(x_i), x_2 - min_i(x_i), \cdots, x_{10} - min_i(x_i)\} & \text{en caso contr} \end{cases}$$

  - Se normaliza el vector anterior entre 0 y 1 dividiendo cada valor por la suma de todos ellos:

$$I_i = \{\frac{x'_1}{\sum_{i=0}^{10} x'_i}, \frac{x'_2}{\sum_{i=0}^{10} x'_i}, \cdots, \frac{x'_{10}}{\sum_{i=0}^{10} x'_i}\}$$

- Siempre se generan individuos válidos y la suma de todos los genes es siempre 1. Esto facilita la realización de la mezcla

EJEMPLO. CASO CONCRETO (1+1)-EE

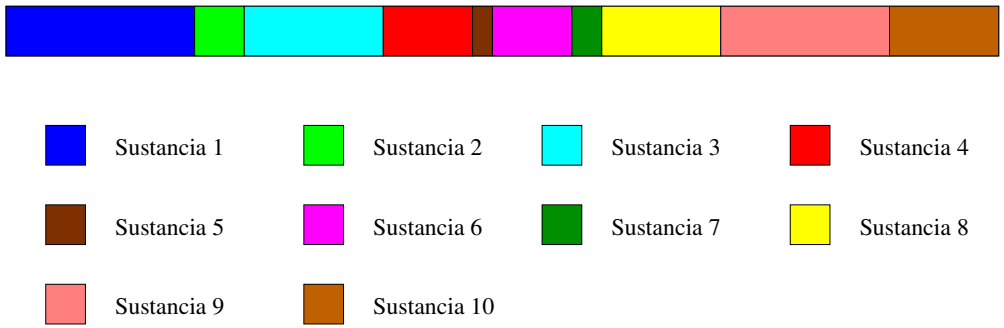
- Se genera aleatoriamente el individuo:

$$I = \{8{,}3; 2{,}2; 6{,}1; 4{,}1; 0{,}9; 3{,}7; 1{,}5; 5{,}4; 7{,}7; 4{,}8\}$$

$$\sigma = \{5{,}2; 0{,}7; 8{,}3; 9{,}0; 1{,}5; 7{,}6; 3{,}1; 6{,}4; 2{,}8; 3{,}4\}$$

- Se normaliza:

$$I = \{0{,}19; 0{,}05; 0{,}14; 0{,}09; 0{,}02; 0{,}08; 0{,}03; 0{,}12; 0{,}17; 0{,}11\}$$



EJEMPLO. CASO CONCRETO (1+1)-EE

- A continuación se calcula la frecuencia de éxitos Ψ
- Hay que decidir el tamaño de las estadísticas. 10, por ejemplo
- Se genera un vector $\vec{\psi}$ de 10 posiciones que se inicializa a valores nulos (por ejemplo -1)

Si en la posición i-ésima del vector hay un 0 es porque en la generación $t - i$, siendo t la generación actual, el individuo que se generó no fue mejor que el padre; si en la posición i-ésima del vector hay un 1 es porque en la generación $t - i$ el individuo que se generó sí fue mejor que el padre

EJEMPLO. CASO CONCRETO (1+1)-EE

- De esta manera el peso del vector binario $\vec{\psi}$, indica el número de éxitos que hubo en las 10 últimas generaciones
- Se asigna como valor de la frecuencia de éxitos el peso de $\vec{\psi}$:

$$\Psi = \begin{cases} \sum_{i=0}^{10} \psi_i & \textit{si } \forall i \; \psi_i \geq 0 \\ 0 & \text{en caso contrario} \end{cases}$$

Ejemplo. Generación de un nuevo individuo

- Se genera su parte funcional mutando la parte funcional del padre mediante una $N(0, \sigma_i)$:

$$I' = \{8{,}3 + N_0(5{,}2), 2{,}2 + N_0(0{,}7), \cdots, 4{,}8 + N_0(3{,}4)\}$$

- Se copia el vector de varianzas del padre en el nuevo individuo, $\sigma' = \sigma$
- Se calcula la adaptación del nuevo individuo

EJEMPLO. PROCEDIMIENTO

- Se actualizan las estadísticas de éxito:
  - Se copia el valor de cada elemento del vector $\vec{\psi}$ en el elemento anterior:

$$\forall i < 10 \; \psi_i = \psi_{i+1}$$

  - Si el nuevo individuo es mejor que el viejo se introduce en la última posición un 1, $\psi_{10} = 1$
  - Si el nuevo individuo es peor que el viejo se introduce en la última posición un 0, $\psi_{10} = 0$
  - Se calcula el valor de Ψ

Ejemplo. Procedimiento

- Se modifica el vector de varianzas mediante la regla del $1/5$:

$$\sigma' = \begin{cases} c_d\sigma & si\ \Psi < \frac{1}{5} \\ c_i\sigma & si\ \Psi > \frac{1}{5} \\ \sigma & si\ \Psi = \frac{1}{5} \end{cases}$$

- El proceso se repite hasta que todos los valores del vector de varianzas estén todos por debajo de cierto umbral
- Criterio de convergencia del método

Temario

Temario