

Introducción a la Computación Evolutiva

Dr. Carlos Artemio Coello Coello

Tarea No. 3

4 de junio de 2009

Variantes de representación y selección

1. El objetivo de esta tarea es que experimente con técnicas alternativas de representación y selección. Usando el mismo programa desarrollado en la tarea anterior (con los cambios indicados a continuación), deberá **minimizar** la siguiente función:

$$\frac{1}{1/K + \sum_{j=1}^{25} f_j^{-1}(x_1, x_2)}, \text{ donde } f_j(x_1, x_2) = c_j + \sum_{i=1}^2 (x_i - a_{ij})^6, \quad (1)$$

$$\text{donde: } -65.536 \leq x_i \leq 65.536, K = 500, c_j = j, \text{ y:} \quad (2)$$

$$[a_{ij}] = \begin{bmatrix} -32 & -16 & 0 & 16 & 32 & -32 & -16 & \cdots & 0 & 16 & 32 \\ -32 & -32 & -32 & -32 & -32 & -16 & -16 & \cdots & 32 & 32 & 32 \end{bmatrix} \quad (3)$$

Considere una precisión de 3 dígitos después del punto decimal. Esta es la función F5 de DeJong.

Para esta tarea se usará un algoritmo genético con las siguientes características:

1. Representación binaria con códigos de Gray y representación entera (un programa separado para cada caso).
2. Selección en la variante indicada más adelante.
3. Cruza de dos puntos.
4. Mutación uniforme (como en la tarea anterior). En el caso de la representación entera, para mutar genere un entero aleatorio entre 0 y 9 para cada posición a ser mutada y haga el reemplazo correspondiente.

Advierta que la técnica de selección que le corresponderá implementar será de acuerdo a la lista siguiente:

- **Arturo** : Muestreo determinístico.

- **Xavier y Miriam** : Implementar el método del sobrante estocástico con reemplazo.
- **Rosa y José Carlos** : Adicionar selección mediante jerarquías lineales (usar $\text{Max}=1.2$) a la selección universal estocástica.
- **Gustavo** : Adicionar escalamiento sigma al método del sobrante estocástico sin reemplazo.

Asegúrese de indicar claramente en sus 2 programas y en su reporte qué técnica le tocó implementar. Si, por cualquier razón implementa una técnica diferente a la que le corresponde, su tarea será automáticamente anulada (es decir, tendrá una calificación de CERO PUNTOS).

Lo que debe entregarse es lo siguiente:

1. **(60 puntos)** El código fuente en C/C++ de dos programas: uno que implemente el algoritmo genético simple con representación binaria con códigos de Gray y otro que implemente una representación entera en la que se fije la posición del punto decimal y se genere un dígito en cada alelo de la cadena (como lo vimos en clase). Es importante asegurarse de que ambos programas pueden lidiar con valores decodificados fuera de rango a todo momento. En cada caso, los listados deberán contar con comentarios y se deberán entregar también los ejecutables, además de una impresión de ambos listados. En el diskette (o CD) deberán incluirse los archivos de TODAS las corridas de ejemplo efectuadas (compactados con PKZIP). Se deberá mencionar en qué plataforma se probó (incluyendo el compilador de C/C++ utilizado) y, de ser necesario, el autor deberá demostrar los programas personalmente al instructor. Los programas deberán pedir al usuario (de forma interactiva y NO a través de la línea de comandos) los parámetros principales del algoritmo genético: tamaño de la población, porcentaje de cruce, porcentaje de mutación, número máximo de generaciones y nombre del archivo donde se almacenarán los datos de cada corrida. En cada generación deberá retenerse a la mejor solución (elitismo), y deberán imprimirse al menos las siguientes estadísticas por generación: media de aptitud de la población, aptitud máxima y mínima, número de cruces efectuadas, número de mutaciones efectuadas y cadena (binaria o decimal) correspondiente al mejor individuo (el que tenga la aptitud más alta). Adicionalmente, deberá imprimirse el mejor individuo global (es decir, contando todas las generaciones), su aptitud y su cadena (binaria o decimal).
2. **(30 puntos)** Dos corridas de ejemplo (una usando una representación binaria con códigos de Gray y la otra usando codificación entera), y los resultados promediados de AL MENOS 20 corridas independientes para cada programa en las que se usen los mismos parámetros (porcentaje de cruce y mutación, tamaño de la población y máximo número de generaciones) pero diferente semilla para generar los números aleatorios (un

total de 40 corridas). Es importante mencionar en su reporte los nombres de los archivos que corresponden a cada conjunto de corridas, así como la colocación de los mismos en subdirectorios separados (es decir, un subdirectorio para la representación con códigos de Gray y otro para la representación entera). En el reporte deberá aparecer la mejor solución obtenida, la media de aptitud, la peor solución obtenida, la varianza y la desviación estándar de cada uno de los dos grupos de 20 corridas efectuadas (para los dos casos antes mencionados). Deberán mostrarse los valores de x_1 , x_2 y $f(x)$ en cada caso. Asimismo, se incluirá la gráfica de convergencia correspondiente a la solución que esté en la mediana de los resultados obtenidos (de las 20 corridas) para cada caso. En esta gráfica se mostrará en el eje de las x el número de generaciones (de cero al máximo permisible) y en el eje de las y se mostrará la aptitud **máxima** de cada generación. TODAS las corridas efectuadas deberán aparecer en archivos diferentes en un diskette o en un CD.

3. **(10 puntos)** Una comparación de resultados entre las dos implementaciones. Usando los resultados producidos en el punto anterior, se deberá analizar lo que ocurrió en cada caso y concluir si la representación hizo alguna diferencia en el desempeño del algoritmo genético (por ejemplo, si aceleró la convergencia hacia el óptimo, si ayudó a encontrar una mejor solución, etc.). Elabore cuidadosamente sus argumentos y diga claramente en qué se basan.
4. **(Bonificación : 10 puntos)** Se bonificará al que presente una corrida que muestre la solución óptima a este problema con cualquiera de las 2 representaciones adoptadas. Esta corrida se deberá presentar en el diskette o CD (en un subdirectorio llamado "OPTIMO"), pero la tabla que resuma los resultados deberá incluirse en el reporte. Debe indicarse claramente cuál de los dos programas (o sea, cuál de las representaciones) se utilizó para obtenerla, así como los parámetros correspondientes (tamaño de población, porcentaje de cruza y mutación y número máximo de generaciones). Es necesario que en su reporte indique claramente cuál fue el resultado óptimo obtenido (es decir, los valores de x_1 , x_2 y $f(x_1, x_2)$) y el nombre del archivo (incluido en el diskette o CD correspondiente) que contenga la corrida que obtuvo ese resultado. De no especificarse claramente en su reporte cuál es el resultado que Ud. considera óptimo **NO RECIBIRÁ BONIFICACIÓN ALGUNA**. Nótese que no se proporcionarán bonificaciones parciales por acercarse al óptimo. Sólo se considerarán las soluciones que lleguen al óptimo con la precisión indicada al inicio de este documento.

Fecha de entrega: *Martes 16 de junio a las 12:00hrs.* Toda tarea entregada tarde será penalizada con 10% (sobre la calificación obtenida) por cada periodo de 24 horas que se retrase su entrega.