

Algoritmos Genéticos aplicados a la Ingeniería biomédica

Genetic algorithms applied to Biomedical Engineering

Ing. Carlos Galeano Uruena, MSc. Ing. Diego A. Garzón-Alvarado Ph.D, Ing. Juan Miguel Mantilla González, PhD

Departamento de Ingeniería Mecánica y Mecatrónica. Facultad de Ingeniería. Universidad Nacional de Colombia.

RESUMEN

En este trabajo se describe las aplicaciones y alcances del método de los algoritmos genéticos (AG) en la investigación en bioingeniería, mecanobiología y medicina. Para este fin, se ha desarrollado el trabajo sobre tres artículos recientes que describen las aplicaciones de los AG en problemas de ingeniería biomédica. Este trabajo pone de manifiesto la importancia del uso de nuevas metodologías de optimización en las investigaciones biomédicas.

Palabras clave: Simulación, algoritmos genéticos, ingeniería biomédica, optimización.

ABSTRACT

In present paper are described the applications and scope of the genetic algorithms method (GA) in the case of the research in the bioengineering, mechanobiology and medicine. For this aim, the paper on three recent articles was developed describing the applications of the GA in problems related to biomedical engineering. Present paper emphasizes the significance of the use of new methodologies of optimization in the biomedical researches.

Key words: Simulations, genetic algorithms, biomedical engineering, optimization.

INTRODUCCIÓN

Las aplicaciones de los AG en Medicina e Ingeniería biomédica son múltiples. Tan solo en la base de datos de Elsevier, la búsqueda de AG con aplicaciones médicas, pueden alcanzar las 5 000 entradas (mayo de 2011). En este universo de información se pueden encontrar investigaciones en la identificación del cáncer,¹ manipulación de imágenes médicas,² análisis de proteínas,^{3,4} evaluación de desempeño de fármacos,⁵ entre otros. En especial, en esta revisión se han elegido tres artículos relacionados con la mecanobiología, es decir, la relación existente entre la biología y la medicina en acción directa con los efectos de las cargas mecánicas sobre los tejidos y órganos. El primer artículo se basa en el desarrollo de modelos de hiperelasticidad en el comportamiento de tejidos blandos, como los músculos, tendones y arterias.⁶ El desarrollo de los modelos matemáticos que determinan el comportamiento mecánico de un tejido tiene gran relevancia para los ingenieros, en especial, los mecánicos y civiles. El problema consiste en encontrar las constantes que determinan el comportamiento de cada material.

El segundo artículo⁷ desarrolla un modelo de algoritmos genéticos en el diagnóstico del cáncer de seno. Por último, el tercer artículo desarrolla un modelo para el diseño de andamios estructurales en tejidos lesionados.

Los algoritmos genéticos en el desarrollo de modelos de hiperelasticidad

Una dificultad que tienen los ingenieros en modelar tejidos humanos y animales, es la determinación de la ley constitutiva que rige el comportamiento mecánico del tejido ante fuerzas externas. El problema radica en la identificación de los parámetros del modelo que son requeridos para las leyes hiperelásticas. La búsqueda de los parámetros es, en ocasiones compleja y tediosa, debido a que se requiere un punto inicial para hacer la búsqueda, tal como lo refiere *Holzapfel*.⁸ Estos modelos son requeridos en la descripción del comportamiento mecánico de tejidos blandos (o suaves), como la pared arterial del sistema circulatorio.

El proceso de optimización se utiliza para encontrar las constantes del modelo hiperelástico. En este contexto, la función objetivo es la norma del residuo entre la cantidad medida y aquella obtenida mediante la experimentación. Cuando dicha función alcanza un valor cero (o cercano) se obtienen los parámetros del modelo. La norma de la aproximación es un problema convexo, por tanto la solución es única. Esta norma del residuo está dada por:

$$f = \frac{\|S_n - S_m\|^2}{\|S_m\|^2} * 100 \quad (1)$$

Donde S_n y S_m son los valores de la función a optimizar y la experimental, respectivamente. En particular, la ecuación constitutiva (relación esfuerzo-deformación) tiene dos componentes, uno isotrópico y otro anisotrópico, que corresponde al comportamiento de las fibras. De esta misma forma, la función a optimizar también se puede dividir en dos términos, según la ley constitutiva:

$$S_n = aM + bN \quad (2)$$

Donde a y b son los parámetros lineales, mientras M y N agrupan los parámetros no lineales a identificar. Al reemplazar (1) en (2) se tiene una función de la forma:

$$f = 100 * \|Ax - S_m\|^2 / \|S_m\|^2 \quad \text{con } A = [M, N] \text{ y } b = [a, b]$$

Esto permitirá escribir la función objetivo como una función cuadrática convexa, en términos de x , tal como:

$$f = \frac{x^T A^T A x - 2 A^T S_m x + \|S_m\|^2}{\|S_m\|^2} * 100 \quad (3)$$

Por tanto, la función objetivo se puede dividir en dos procesos. El proceso analítico y por AG. El primero se puede obtener en función de x , siempre y cuando el gradiente de f sea cero, por lo que se obtiene:

$$A^T A x = A S_m \quad (4)$$

De esta forma, mediante manipulación matemática, se puede encontrar el valor de x .

De otro lado, mediante algoritmos genéticos, se deben encontrar los valores no lineales que se encuentran en M y N . El método evita el uso de una única condición inicial, como es tradicional en los métodos clásicos.⁸ Por tanto, se parte de múltiples puntos semillas (cromosomas/individuos) que forman la población inicial, y van reproduciendo la población mediante operadores de AG (crossover y mutación). Como es común, el proceso se detiene cuando todos los individuos comparten los mismos genes. Gracias a la separación de la parte lineal y la no lineal, los algoritmos genéticos solo son requeridos para los valores no lineales. En la figura 1, en la parte A, se puede observar la metodología propuesta para hallar los valores no lineales que pertenecen a las ecuaciones $A = [M, N]$ y $b = [a, b]$, que se muestran en la parte B de la figura 1. En ésta última, se puede observar la relación entre los experimentos y el modelo matemático propuesto en la función objetivo.

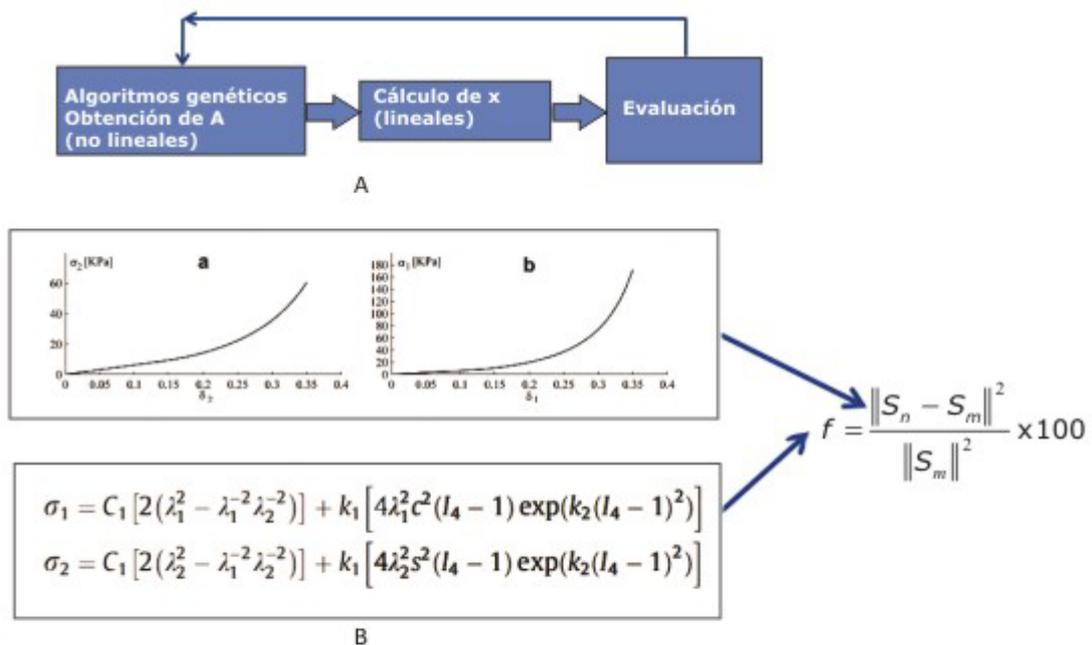


Fig. 1. A: proceso en la estimación de los parámetros de la relación constitutiva de la pared arterial. B: relación del modelo y los experimentos en la función objetivo.

Algoritmos genéticos aplicados en el diagnóstico de cáncer de seno

Tseng y Liao⁹ proponen una metodología de AG para encontrar la incidencia y relación entre el cáncer/fibroadenoma y los virus existentes en el ADN (DNA virus). Diversos investigadores han desarrollado trabajos similares,¹⁰⁻¹³ donde se utilizan estrategias de relación entre un marcador molecular y/o genético para detectar el cáncer. Basados en el trabajo de Kim y Han,¹⁴ se propone un algoritmo genético de forma tradicional con los tres operadores a saber: selección, *crossover* (cruzamiento) y mutación. Sin embargo, la diferencia radica en:

1. Cambiar el operador *crossover*, esto es, el operador *crossover* actual se sustituye por otro en cualquier tiempo.
2. Ajuste adaptativo de las tasas de *crossover* y mutación para incrementar la diversidad genética y guía del sistema.

De esta forma, el proceso completo se puede ver en la figura 2.

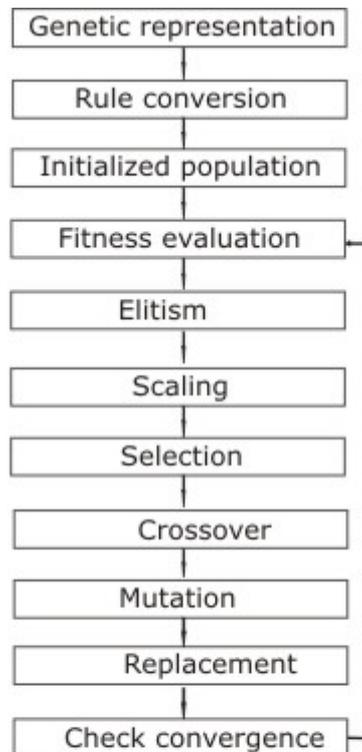


Fig. 2. Representación del proceso de algoritmo genético utilizado (Tseng M, Liao H. The genetic algorithm for breast tumor diagnosis-The case of DNA viruses. *Applied Soft Computing*. 2009;9[2]:703-10).

Para determinar la representación genética (del algoritmo) se toman los posibles virus del ADN y la enfermedad de cáncer/fibroadenoma. Por tanto, en el mismo vector se dispone de la incidencia en el factor de riesgo y la enfermedad, separados por un punto y coma, tal como se representa en la figura 3.

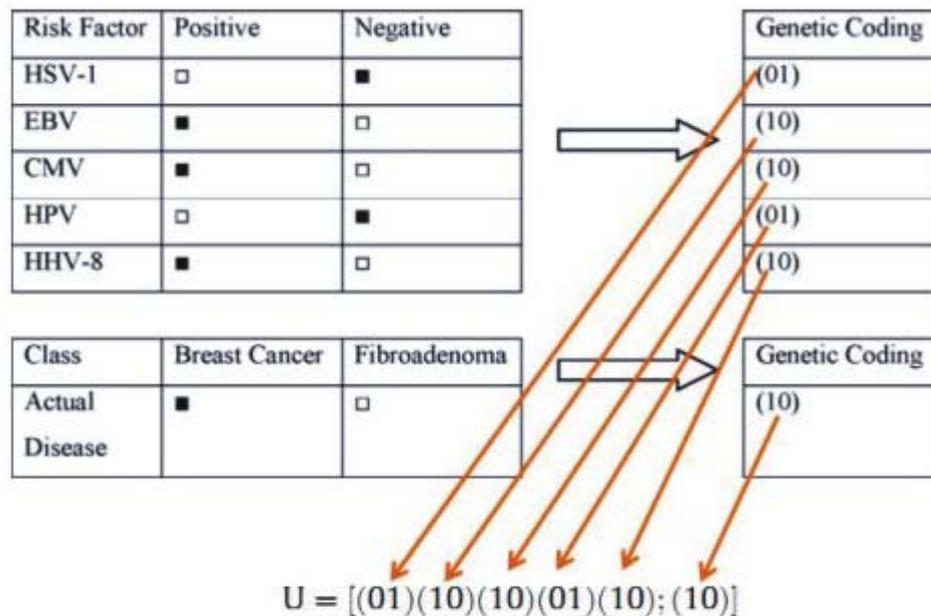


Fig. 3. Ejemplo de la representación genética del problema.

Por tanto, el vector U de la figura produce una cadena de factores de riesgo (enlazadas mediante un *and* o *or*) que entregan una determinada enfermedad (*=then*). Para evaluar la categoría de los datos se ha utilizado una función objetivo de la forma:

$$\text{maximizar} \left(\frac{\text{Número de casos correctamente clasificados}}{\text{Número total de casos}} \right) \quad (5)$$

De otro lado, al igual que en la naturaleza, los operadores de selección, cruzamiento y mutación son esenciales para aplicar presión a la población para la selección natural. Reglas que tienen un buen *fitness*, servirán para la reproducción. En este sentido se utilizó la regla de rueda de ruleta.¹⁵ El cruzamiento permite intercambiar la información entre diferentes organismos, de igual manera, en este trabajo se empleó cruzamiento de un punto, el cual ocurre en conjunto con la probabilidad, llamada *tasa de cruzamiento* (0,4 a 0,9). Por último la mutación es utilizada para cambiar, aleatoriamente, un miembro de la población y cambiar un bit en su cadena de representación de bits. Por tanto, la selección y cruzamiento producen nuevas cadenas. En contraposición, la mutación evita alcanzar cualquier punto en el espacio de búsqueda, por tal razón, este valor es manejado con cuidado, con una probabilidad de *tasa de mutación* de 0,001 a 0,3. Este proceso se repite en 100 generaciones.

Como la rueda de ruleta no garantiza la selección de un individuo particular, se utiliza el individuo de la "elite", para mantenerlo y propagar su descendencia, además de mantenerlo durante varias generaciones. Para este fin, se ha utilizado un operador de elitismo.

Para el entrenamiento y prueba del algoritmo, se utilizaron una fuente de 80 datos, 52 mujeres con cáncer de seno y 28 con fibroadenoma mamario, obtenidos en el

Hospital Universitario Chung-Shan.⁹ En ellas se midieron los virus DNA y mediante ANOVA se estableció la significancia de cada uno de los posibles virus de DNA en la incidencia de Cáncer o fibroadenoma. Se encontraron 5 virus importantes (ver figura 2), de los cuales, el HHV-8 y HSV-1, son los más importantes en la identificación de las enfermedades. Por tanto, mediante el análisis estadístico ANOVA se establecen las características (reglas del modelo) para ser usadas en AG. Por consiguiente, se establecen varias reglas, según el tipo de virus, las cuales se pueden observar en la figura 3. Con 60 datos se alimenta el algoritmo genético para establecer las reglas de desempeño (que fueron 5, según el tipo de virus de ADN) y los 20 datos restantes sirven para probar si el algoritmo ha funcionado correctamente.

En la tabla 1, se observan los resultados, donde en la parte superior están las reglas para la diagnosis del fibroadenoma, y en la parte inferior, para el diagnóstico del cáncer.

Tabla 1. Resultados sobre las muestras con las reglas obtenidas para el fibroadenoma (arriba) y cáncer (abajo)

Regla no.	Modelo de la regla					Datos de entrenamiento (60)	Datos de ensayo (20)
	HSV-1	EBV	CMV	HHV-8	Clase	Fitness	Fitness
1	(10)	(11)	(11)	(10)	(01)	0,8625	0,9
2	(10)	(11)	-	(10)	(01)	0,8625	0,9
3	(10)	-	-	(10)	(01)	0,8625	0,9
4	-	-	-	(10)	(01)	0,65	0,8
5	(10)	-	-	-	(01)	0,80	0,8

Regla no.	Modelo de la regla					Datos de entrenamiento (60)	Datos de ensayo (20)
	HSV-1	EBV	CMV	HHV-8	Clase	Fitness	Fitness
1	(01)	(11)	(11)	(11)	(10)	0,80	0,8
2	(01)	(11)	-	(11)	(10)	0,80	0,8
3	(01)	-	-	(11)	(10)	0,80	0,8
4	-	-	-	(01)	(10)	0,65	0,8
5	(01)	-	-	-	(10)	0,80	0,8

Algoritmos genéticos aplicados en el diseño de andamios estructurales para tejidos suaves (scaffolds)

Existe un incremento considerable de lesiones y enfermedades que conducen a la pérdida funcional de tejido blando o duro. Por ejemplo, enfermedades como la osteoporosis desarrolla una pérdida progresiva de tejido óseo confiriéndole a los huesos una alta posibilidad de quebrarse. En este sentido se desarrollan andamios estructurales, denominados *scaffolds* que tienen la doble intención de, por un lado, permitir el desarrollo de tejido en aquellas zonas con gran pérdida de tejido y,

conferir resistencia mecánica por otro. *Rezende* y otros¹⁶ han desarrollado un modelo matemático obtenido experimentalmente que describe el comportamiento mecánico de degradación del material del *scaffold*. El material utilizado es alginato. De otro lado, para el diseño han utilizado algoritmos genéticos para maximizar el módulo elástico.

Los *scaffolds* tienen características únicas de alta porosidad (macro-porosidad), apropiada morfología de superficie (micro-porosidad), gran cantidad de área, tamaño de poro adecuado y alta conectividad de la estructura porosa. Además de estas características, el *scaffold* debe ser biocompatible y biodegradable, como lo es el alginato. Por tanto, el objetivo del artículo¹⁶ es encontrar los mejores valores de la cantidad de alginato para su fabricación que maximice el módulo elástico.

El problema de optimización, por consiguiente, se enfoca en determinar las características para la fabricación del *scaffold*. En especial se deben encontrar los valores óptimos de la composición del alginato y la porosidad inicial, para que, en un tiempo específico, se presente el mayor módulo elástico. Este problema se puede escribir como:

$$\begin{aligned} &\text{maximizar} && E(\phi_0, \alpha, t) \\ &\text{sujeto } \alpha && 1\% \leq \alpha \leq 8\% \\ &&& 30\% \leq \phi_0 \leq 80\% \end{aligned} \tag{6}$$

Donde, E es el módulo elástico, α es el porcentaje de alginato y ϕ_0 la porosidad inicial.

Para conocer la evolución de las propiedades mecánicas en el tiempo, se determinó el cambio de esta en conjunto con la degradación y la porosidad (el cambio en el tiempo se denomina contracción $C(\phi_0, \alpha, t)$). Este paso fue desarrollado mediante experimentación.¹⁶ En este sentido, se halla que la evolución de la porosidad en el tiempo es función de la porosidad inicial ϕ_0 , el porcentaje de alginato α , el término de degradación (o contracción $C(\phi_0, \alpha, t)$) y las funciones ζ y ψ , dado por la expresión:

$$\phi(\phi_0, \alpha, t) = \phi_0 + \zeta(\phi_0, \alpha) * C(\phi_0, \alpha, t) + \psi(\phi_0, \alpha) * C^2(\phi_0, \alpha, t)$$

De igual manera, se puede dar una expresión para el módulo de elasticidad:

$$E(\phi_0, \alpha, t) = E_0 + k_1(\phi_0, \alpha) * \phi(\phi_0, \alpha, t) + k_2(\phi_0, \alpha) * \phi^2(\phi_0, \alpha, t) + k_3(\phi_0, \alpha) * \phi^3(\phi_0, \alpha, t)$$

Por tanto, el problema de optimización se amplía para que el nivel de degradación (o contracción) y la porosidad final sean mayor que 25 y 80 %, respectivamente. Para este fin se utilizan los algoritmos genéticos, los cuales inician con poblaciones aleatorias de cromosomas que son un conjunto de soluciones para el problema de optimización original. En cada generación la función objetivo es evaluada en cada individuo y múltiples individuos son estocásticamente elegidos de la población actual (basado sobre el valor del fitness). Luego, son modificadas (recombinadas y mutadas aleatoriamente) para formar una nueva población. El AG termina cuando

el número máximo de generaciones se ha producido o, cuando la población entera cumple la función objetivo. El código fue desarrollado en FORTRAN.

Se han empleado los operadores genéticos de selección del tipo *tournament*, cruzamiento uniforme, *creep* y salto de mutación. Además se ha utilizado Niching y elitismo (figura 4).



Fig. 4. Metodología propuesta en Rezende R, Rezende M, Bártolo P, Mendes A, Filho R. Optimization of scaffolds in alginate for biofabrication by genetic algorithms. Computer Aided Chemical Engineering. 2009;27:1935-40.

Los parámetros de entrada del modelo se observan en la tabla 2:

Tabla 2. Valores de entrada del ASG¹⁶

GA input parameters	Value
Population size per generation	50
Maximum number of generations	30
Crossover probability	0,60
Jump mutation probability	0,077
Creep mutation probability	0,077
Initial random number seed for the GA run	-1000

Por tanto, se desarrollaron dos tipos de ejemplos, con contracción mayor al 25 % y porosidad final del 80 %. La figura de la evolución según cada generación se ve en la figura 5.

En este trabajo se llevó a cabo la recopilación de 3 trabajos sobre algoritmos genéticos aplicados a la ingeniería biomédica. Se observa la importancia de los algoritmos en el desarrollo de nuevas teorías sobre el comportamiento de enfermedades y para la optimización de problemas en medicina. En siguientes estudios se desarrollará un algoritmo genético aplicado a problemas de diagnóstico en enfermedades ortopédicas.

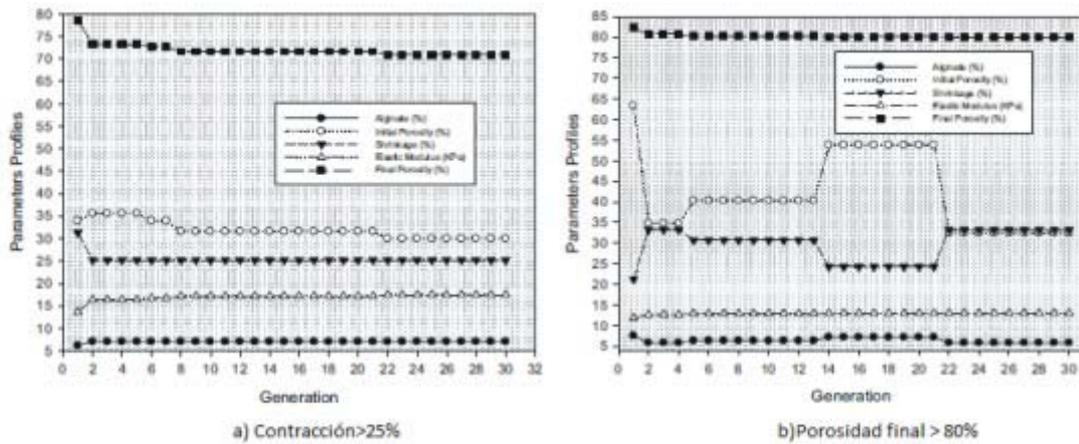


Fig. 5. Gráfica de evolución del algoritmo genético en función de la generación; en Rezende R, Rezende M, Bártolo P, Mendes A, Filho R. Optimization of scaffolds in alginate for biofabrication by genetic algorithms. *Computer Aided Chemical Engineering*. 2009;27:1935-40.

REFERENCIAS BIBLIOGRÁFICAS

1. Ahmad S, Bergen. S. A genetic algorithm approach to the inverse problem of treatment planning for intensity-modulated radiotherapy. *Biomedical Signal Processing and Control*. 2010;5(3):189-95.
2. Yeh J, Fu J. A hierarchical genetic algorithm for segmentation of multi-spectral human-brain MRI. *Expert Systems with Applications*. 2008;34(2):1285-95.
3. Arora V, Bakhshi, A. Theoretical investigations on model ternary polypeptides using genetic algorithm. Some new results. *Chemical Physics*, In Press. online 15 March 2011.
4. Hill Tobias, Lundgren Andor, Fredriksson Robert, Schiöth Helgi B. Genetic algorithm for large-scale maximum parsimony phylogenetic analysis of proteins. *Biochimica et Biophysica Acta (BBA)-General Subjects*. 2005;1725(1):19-29.
5. Qiang Fei, Ming Li, Bin Wang, Yanfu Huan, Guodong Feng, Yulin Ren. Analysis of cefalexin with NIR spectrometry coupled to artificial neural networks with modified genetic algorithm for wavelength selection. *Chemometrics and Intelligent Laboratory Systems*. 2009;97(2):127-31.
6. Harb N, Labeed N, Domaszewski M, Peyraut F. A new parameter identification method of soft biological tissue combining genetic algorithm with analytical optimization. *Computer Methods in Applied Mechanics and Engineering*. 2011;200(1-4):208-15.
7. Tseng, Ming-Hseng, Liao Hung-Chang. The genetic algorithm for breast tumor diagnosis—The case of DNA viruses. *Applied Soft Computing*. 2009;9(2):703-10.
8. Holzapfel GA. Determination of material models for arterial walls from uniaxial extension tests and histological structure. *J. Theor Biol*. 2006;238:290-302.

9. Tseng M, Liao H. The genetic algorithm for breast tumor diagnosis—The case of DNA viruses. *Applied Soft Computing*. 2009;9(2):703-10.
10. Cheng SH, Tsou MH, Liu MC, Jian JJ. Unique features of breast cancer in Taiwan. *Breast Cancer Res Treat*. 2000;63(3):213-20.
11. Ziegler RG, Hoover RN, Pike MC, Hildesheim A, Nomura AM, West DW, et al. Migration patterns and breast cancer risk in Asian-American women. *J Natl Cancer Inst*. 1993;85:1819-27.
12. Yu H, Rohan TE, Cook MG, Howe GR, Miller AB. Risk factors for fibroadenoma: a case-control study in Australia. *Am J Epidemiol*. 1992;135(3):247-59.
13. Dimmock NJ, Primrose SB. *Carcinogenesis and tumor viruses: introduction to modern virology*. 4th ed. London: Blackwell Science Ltd. p. 256-275.
14. Kim MJ, Han I. The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms. *Expert Syst Appl*. 2003;25(4):637-46.
15. Coley DA. *An Introduction to Genetic Algorithms for Scientists and Engineers*. London: World Scientific; 1998.
16. Rezende R, Rezende M, Bártolo P, Mendes A, Filho R. Optimization of Scaffolds in Alginate for Biofabrication by Genetic Algorithms. *Computer Aided Chemical Engineering*. 2009;27:1935-40.

Recibido: 5 de febrero de 2011.

Aprobado: 20 de febrero de 2011.

Carlos Galeano Truena. Departamento de Ingeniería Mecánica y Mecatrónica.
Facultad de Ingeniería. Universidad Nacional de Colombia.

Autor para la correspondencia: Diego A. Garzón-Alvarado. Correo electrónico:
dagarzona@bt.unal.edu.co