Optimización Evolutiva con Objetivos Múltiples : Estado del Arte y Tendencias Futuras

Carlos A. Coello Coello[1]
LANIA
Rébsamen 80, Col. Isleta
Xalapa, Veracruz
México
ccoello@xalapa.lania.mx

Resumen

En este artículo se presenta una revisión crítica de las técnicas evolutivas más importantes para optimización con objetivos múltiples que se han desarrollado en los últimos años. En cada caso, se proporcionará una breve descripción del algoritmo principal de la técnica, mencionando además sus ventajas y desventajas. Finalmente, se indicarán las tendencias futuras de investigación en esta disciplina, enfatizando los problemas actuales y los futuros desafíos en esta área.

Palabras clave: optimización con objetivos múltiples, optimización vectorial, algoritmos evolutivos, inteligencia artificial, algoritmos genéticos.

1. Introducción

A partir del trabajo pionero que Rosenberg realizara a fines de los 1960s con respecto a la posibilidad de usar búsqueda genética como una forma de lidiar con objetivos múltiples, esta nueva área de investigación (llamada ahora optimización evolutiva con objetivos múltiples) ha crecido considerablemente y esto se refleja con un notable incremento (sobre todo en los últimos 5 años) de artículos técnicos en revistas internacionales, sesiones especiales en conferencias especializadas y grupos de interés en Internet.

La optimización con objetivos múltiples es, sin duda, un área de investigación muy importante tanto para los científicos como para los ingenieros, no sólo debido a que la mayoría de los problemas del mundo real tienen objetivos múltiples, sino también porque todavía restan por resolver muchas interrogantes en esta disciplina. De hecho, no hay ni siquiera una definición de "óptimo" que sea aceptada universalmente, como en el caso de la optimización con un solo objetivo. Esto hace difícil incluso poder comparar los resultados de una técnica con los de otra, porque normalmente la decisión acerca de cuál es la mejor respuesta corresponde realmente a un humano que aplicará sus propios criterios (los cuales suelen ser subjetivos).

Son pocas las reseñas sobre este tema que se han efectuado en la literatura especializada en los últimos 15 años: Tamaki et al. [1], que hace una revisión muy breve y superficial de algunas de las técnica más importantes y Fonseca y Fleming [2], que hacen una excelente reseña de los principales problemas que se enfrentan al tratar de lidiar con objetivos múltiples usando una

---

[1]Este trabajo fue realizado cuando el autor se encontraba afiliado al *Engineering Design Centre* de la Universidad de Plymouth, en el Reino Unido.

técnica evolutiva. Sin embargo, en ambos casos, se proporciona poca información con respecto a los detalles de cada técnica y sus alcances. Además, desde la publicación de estos dos artículos, ha surgido un número importante de nuevas técnicas y la intención de este trabajo es proporcionar una reseña actualizada de la actividad en investigación que se ha dado en esta disciplina en los últimos años, indicando además cuáles parecen ser las rutas más prometedoras de investigación futura.

2. Descripción del Problema

La optimización con objetivos múltiples (llamada también optimización vectorial y con criterios múltiples) puede definirse como [3]:

> El problema de encontrar un vector de variables de decisión que satisfaga las restricciones y optimice una función vectorial cuyos elementos representen las funciones objetivo. Estas funciones forman una descripción matemática de los criterios de desempeño que usualmente están en conflicto entre sí. Por lo tanto, el término "optimizar" significa encontrar una solución tal que proporcione valores para todos los objetivos que resulten aceptables para el diseñador.

De manera formal, podemos definir el problema de la manera siguiente:

Encontrar el vector $\bar{x}^* = \left[x_1^*, x_2^*, \ldots, x_n^*\right]^T$ que satisfaga las m restricciones de desigualdad:

$$g_i(\bar{x}) \leq 0 \qquad i = 1, 2, \ldots, m \tag{1}$$

las p restricciones de igualdad

$$h_i(\bar{x}) = 0 \qquad i = 1, 2, \ldots, p \tag{2}$$

y que optimice la función vectorial

$$\bar{f}(\bar{x}) = \left[f_1(\bar{x}), f_2(\bar{x}), \ldots, f_k(\bar{x})\right]^T \tag{3}$$

donde $\bar{x} = \left[x_1, x_2, \ldots, x_n\right]^T$ es el vector de variables de decisión. En otras palabras, deseamos determinar de entre el conjunto F de valores que satisfacen (1) y (2) al conjunto $x_1^*, x_2^*, \ldots, x_k^*$ que produzca los valores óptimos de todas las funciones objetivo.

Las restricciones dadas por (1) y (2) definen la *región factible* F y cualquier punto $\bar{x}$ en F constituye una *solución factible*. La función vectorial $\bar{f}(\bar{x})$ mapea al conjunto F en el conjunto X, que representa todos los valores posibles de las funciones objetivo. Los k componentes del

vector $\bar{f}(\bar{x})$ representan los criterios *no conmensurables*[2] que se considerarán. Las restricciones $g_i(\bar{x})$ y $h_i(\bar{x})$ representan las limitantes impuestas sobre las variables de decisión. El vector $\bar{x}^*$ se reservará para denotar las soluciones óptimas (habrá normalmente más de una).
El problema es que el significado de la palabra *óptimo* no está bien definido en este contexto, puesto que en raras ocasiones tenemos una $\bar{x}^*$ tal que para toda $i = 1, 2, \ldots, k$

$$\bigwedge_{x \in F} \left(f_i(\bar{x}^*) \leq f_i(\bar{x}) \right) \qquad (4)$$

Si este fuese el caso, entonces $\bar{x}^*$ sería una solución deseable, pero desgraciadamente, en la práctica casi nunca tenemos situaciones como ésta, en las que todas las $f_i(\bar{x})$ tengan un mínimo en F localizado en un punto común $\bar{x}^*$.

2.1 Optimo de Pareto

El concepto de *Optimo de Pareto* fue formulado por Vilfredo Pareto en el siglo XIX [4], y constituye por sí mismo el origen de la investigación en optimización con objetivos múltiples. Decimos que un punto $\bar{x}^* \in F$ es un *óptimo de Pareto* si para toda $\bar{x} \in F$,

$$\bigwedge_{i \in I} \left(f_i(\bar{x}) = f_i(\bar{x}^*) \right) \qquad (5)$$

o, hay al menos una $i \in I$ (*I* es el conjunto de funciones objetivo del problema) tal que

$$f_i(\bar{x}) > f_i(\bar{x}^*) \qquad (6)$$

En palabras, esta definición dice que $\bar{x}^*$ es un óptimo de Pareto si no existe un vector factible $\bar{x}$ que decrementaría algún criterio sin causar un incremento simultáneo en al menos otro de ellos. Desafortunadamente, el óptimo de Pareto casi siempre produce no una, sino un conjunto de soluciones a las que se les llama *no inferiores* o *no dominadas*.

3. Formas simplistas de lidiar con objetivos múltiples

La noción de búsqueda genética en un problema con varios objetivos se remonta a fines de los 1960s, en que Rosenberg [5] realizó un estudio que contenía una sugerencia que habría conducido a la optimización con objetivos múltiples de haberse realizado de la manera indicada en su trabajo. Su sugerencia fue usar varias *propiedades* (cercanías a alguna composición química) en su simulación de la genética y la química de una población de organismos unicelulares. Puesto que su implementación sólo contenía una sola propiedad, esta técnica para lidiar con objetivos múltiples

---

[2]El término *no conmensurable* se refiere al hecho de que los valores de las funciones objetivo están expresados en diferentes unidades.

no se pudo llevar a la práctica, aunque la sugerencia marcó el inicio de la investigación en esta área usando técnicas evolutivas.
Si sabemos que un algoritmo genético (AG) requiere de información escalar sobre el valor de aptitud de un individuo para operar, probablemente la idea más simple que podríamos proponer para lidiar con varios objetivos sería combinarlos en uno solo usando una suma, una multiplicación o cualquier otra combinación de operaciones aritméticas que se nos pueda ocurrir. Hay, sin embargo, problemas obvios con esta técnica. El primero de ellos es que debemos proporcionar información escalar precisa sobre el rango de los objetivos, a fin de evitar que uno de ellos domine a los demás. Esto implica que debemos saber, en la medida de lo posible, el comportamiento de cada una de las funciones objetivo, lo cual es normalmente (al menos en la mayoría de las aplicaciones del mundo real) un proceso muy costoso (en términos de tiempo de CPU) que suele estar fuera de nuestro alcance. Obviamente, si esta combinación de objetivos es posible (y lo es, en algunos casos), esta técnica no sólo es la más simple de implementar, sino que además es la más eficiente, porque no se requiere posterior interacción con el usuario y si el algoritmo genético tiene éxito en el proceso de optimización, entonces los resultados serán al menos subóptimos en la mayoría de los casos.
Al proceso de combinar objetivos en una sola función se le denomina normalmente *función agregada*, y se le ha utilizado en diversas ocasiones con éxito relativo en problemas en los cuales el comportamiento de las funciones objetivo se conoce más o menos bien. Esta sección incluye las técnicas más populares que se basan en esta idea.

3.1 Suma de pesos

Este método consiste en sumar todas las funciones objetivo usando diferentes pesos para cada una de ellas. Esto significa que nuestro problema con objetivos múltiples se transforma en un problema de optimización de la forma:

$$min \sum_{i=1}^{k} w_i f_i(\bar{x}) \qquad (7)$$

donde $w_i \geq 0$ son los pesos que representan la importancia relativa de los objetivos. Se supone usualmente que

$$\sum_{i=1}^{k} w_i = 1 \qquad (8)$$

Puesto que los resultados de resolver un problema de optimización usando (7) pueden variar significativamente conforme se modifiquen los pesos y puesto que se sabe normalmente muy poco acerca de la forma más adecuada de seleccionar estos coeficientes, se requiere resolver el mismo problema usando diferentes valores de w_i. Pero en este caso, el diseñador todavía tiene que confrontar la decisión de tener que elegir la solución más apropiada en base a su intuición.

Debe advertirse que los pesos no reflejan proporcionalmente la importancia relativa de los objetivos, sino que son sólo factores que, al variarse, localizan puntos diferentes en el conjunto de Pareto. Para los métodos numéricos que puedan usarse para buscar el mímino de (7), esta localización de puntos depende no sólo de w_i, sino también de las unidades en las que se expresen las funciones.

Análisis de la técnica

Este método fue el primero en desarrollarse para la generación de soluciones no inferiores en problemas de optimización con objetivos múltiples. Esto es una consecuencia obvia del hecho de que el método puede inferirse del trabajo seminal de Kuhn y Tucker en optimización numérica [6]. Esta técnica es muy eficiente desde el punto de vista de recursos de cómputo, y puede usarse para generar una solución fuertemente dominada que pueda usarse como un punto inicial para otras técnicas. El principal problema de esta técnica es cómo determinar los pesos apropiados cuando no tenemos suficiente información acerca del problema. En este caso, cualquier punto óptimo obtenido será una función de los coeficientes usados para combinar los objetivos. La mayor parte de los investigadores usan una simple combinación lineal de objetivos y después generan la *superficie de compromisos*[3] mediante la variación de los pesos. Esta técnica es muy simple y fácil de implementar, pero tiene como desventaja principal el perder porciones cóncavas de la curva compromiso[4] [7], lo cual es un inconveniente serio si se pretenden resolver problemas del mundo real.

3.2 Programación de metas

En este método, el tomador de decisiones tiene que asignar objetivos o metas que desee alcanzar para cada objetivo. Estos valores se incorporan en el problema como restricciones adicionales. La técnica tratará entonces de minimizar las desviaciones absolutas de cada objetivo con respecto a lo deseado. La forma más simple de este método puede formularse de la manera siguiente [8]:

$$min \sum_{i=1}^{k} \left| f_i(\bar{x}) - T_i \right|, \quad \text{sujeta a } \bar{x} \in F \qquad (9)$$

donde T_i denota a la meta u objetivo establecida por el tomador de decisiones para la i-ésima función objetivo $f_i(\bar{x})$, y F representa la región factible. El criterio es, entonces, minimizar la suma de los valores absolutos de las diferencias entre los valores deseados y los obtenidos. Una formulación más general de esta técnica es una suma ponderada de la p-ésima potencia de la desviación $\left| f_i(\bar{x}) - T_i \right|$ [9]. A esta formulación normalmente se le denomina *programación de metas generalizada* [10].

---

[3] El término "compromiso" en este contexto se refiere al hecho de que negociamos el valor de una función objetivo por el valor de otra función o funciones.

[4] En otras palabras, no funciona bien con espacios de búsqueda no convexos.

Análisis de la técnica

Esta técnica producirá una solución dominada si la meta seleccionada yace en la región factible [8]. La técnica puede ser muy eficiente (en términos de tiempo de CPU) si conocemos las metas que deseamos obtener y si éstas se encuentran en la zona factible. Sin embargo, el tomador de decisiones tiene la tarea de encontrar pesos o prioridades adecuadas para los objetivos que eliminen las características no conmensurables del problema, lo cual es difícil en la mayor parte de los casos, a menos que exista conocimiento previo acerca de la topología del espacio de búsqueda. Asimismo, si la región factible es difícil de localizar, este método puede volverse muy ineficiente. Sin embargo, la técnica puede ser muy útil en casos en los que se pueden efectuar aproximaciones lineales parciales de las funciones objetivo, debido a la disponibilidad de excelentes programas para esa tarea y a la posibilidad de eliminar las metas dominadas fácilmente. Por otra parte, en los casos no lineales, otras técnicas resultarán generalmente más eficientes.

3.3 Satisfacción de metas

En esta técnica se debe proponer un vector de pesos $w_1, w_2, \ldots, w_k$, el cual describa la sub-satisfacción o sobre-satisfacción de las metas deseadas $b_1, b_2, \ldots, b_k$ para las funciones objetivo $f_1, f_2, \ldots, f_k$. Para encontrar la mejor solución compromiso $\bar{x}^*$ resolvemos el siguiente problema:

$$\text{Minimizar } \alpha \qquad (10)$$

sujeta a:

$$g_j(\bar{x}) \leq 0 \qquad j = 1, 2, \ldots, m$$

$$b_i + \alpha \cdot w_i \geq f_i(\bar{x}) \qquad i = 1, 2, \ldots, k \qquad (11)$$

donde α es una variable escalar sin restricciones en el signo y los pesos $w_1, w_2, \ldots, w_k$ están normalizados de forma que

$$\sum_{i=1}^{k} |w_i| = 1 \qquad (12)$$

Si algún peso $w_i = 0$ $(i = 1, 2, \ldots, k)$, esto significa que el límite máximo de los objetivos $f_i(\bar{x})$ es b_i. Puede demostrarse fácilmente [11] que el conjunto de soluciones no dominadas para un problema puede ser generado variando los pesos, con $w_i \geq 0$ $(i = 1, 2, \ldots, k)$ aún para problemas no convexos. Debe hacerse notar que el valor óptimo de α informará al tomador de decisiones sobre si las metas son posibles de satisfacer o no. Un valor negativo de α implica que la meta es factible de satisfacerse y que se obtendrá una solución mejorada. De lo contrario, si $\alpha > 0$, entonces la meta no puede satisfacerse.

Análisis de la técnica

Como indican Wilson y MacLoud [12], la satisfacción de metas tiene varios problemas, de los cuales probablemente el principal es la presión de selección errónea que puede generar en algunos casos. Por ejemplo, para un problema con 2 objetivos, si tenemos dos soluciones candidatas que son iguales en cuanto al valor de una de sus funciones objetivo, pero diferente en el otro, aún así tendrán el mismo valor de satisfacción de la meta para sus dos objetivos, lo que significa que para el AG ninguna de las dos soluciones es mejor que la otra.

3.4 El método de las restricciones ε

Este método se basa en la minimización de una función objetivo (la principal o preferida) y considera a los demás objetivos como restricciones que están acotadas por ciertos niveles permisibles ε_i. Por lo tanto, se efectúa una minimización con un solo objetivo para la función objetivo más relevante f_1 sujeta a restricciones adicionales en las otras funciones objetivo. Los niveles ε_i se alteran después para generar todo el conjunto de Pareto. El método puede formularse de la manera siguiente:

(1) Encontrar el mínimo de la *r*-ésima función objetivo; es decir, encontrar $\bar{x}^*$ tal que

$$f_r(\bar{x}^*) = \min_{x \in F} f_r(\bar{x}) \qquad (13)$$

sujeta a restricciones adicionales de la forma

$$f_i(\bar{x}) \leq \varepsilon_i \quad i = 1, 2, \ldots, k \ ; \ i \neq r \qquad (14)$$

donde ε_i son los valores supuestos de las funciones objetivo que no deseamos exceder.

(2) Repetir (1) para diferentes valores de ε_i. La información derivada de un conjunto bien seleccionado de ε_i puede ser útil para tomar la decisión final. La búsqueda se detiene cuando el tomador de decisiones encuentra una solución satisfactoria.

Puede ser necesario repetir el proceso anterior para diferentes índices *r*.

Para obtener valores adecuados de ε_i, se efectúan normalmente optimizaciones individuales para cada una de las funciones objetivo en turno, usando técnicas de programación matemática (o AGs independientes). Para cada función objetivo f_i $(i = 1, 2, \ldots, k)$, hay un vector de diseño óptimo $\bar{x}_i^*$ para el cual $f_i(\bar{x}_i^*)$ es un mínimo. Hagamos que $f_i(\bar{x}_i^*)$ sea el límite inferior de ε_i. Es decir,

$$\varepsilon_i \geq f_i(\bar{x}_i^*) \quad i = 1, 2, \ldots, r-1, r+1, \ldots, k \qquad (15)$$

y $f_i(\bar{x}_r^*)$ será el límite superior de ε_i; es decir,

$$\varepsilon_i \leq f_i(\bar{x}_r^*) \quad i = 1, 2, \ldots, r-1, r+1, \ldots, k \qquad (16)$$

Cuando los límites ε_i son demasiado bajos, no hay solución y al menos uno de ellos debe relajarse. Esta técnica fue sugerida por Ritzel y Wayland [7] como una manera simple para lidiar con objetivos múltiples usando un AG. La idea es codificar el AG de tal forma que todos los objetivos, excepto por uno, se mantengan constantes (restringidos a un solo valor), y el objetivo restante se vuelve entonces la función de aptitud del AG. De tal forma, a través de un proceso de ejecuciones múltiples de un AG con diferentes valores para los objetivos restringidos, se puede generar una superficie compromiso.

Análisis de la técnica

El inconveniente obvio de esta técnica es que consume mucho tiempo de CPU y que la codificación de las funciones objetivo puede ser muy difícil o incluso imposible para ciertos problemas, particularmente si hay muchos objetivos. Además, este método tiende a encontrar soluciones débilmente no dominadas, lo cual puede no ser muy adecuado en ciertas aplicaciones (por ejemplo, optimización estructural). Sin embargo, la simplicidad relativa de la técnica la ha hecho muy popular entre algunos investigadores.

4. Técnicas más sofisticadas que no se basan en el óptimo de Pareto

Para superar las dificultades involucradas con los métodos agregados, se han enfocado varios esfuerzos hacia el desarrollo de técnicas alternativas basadas en políticas poblacionales o en el manejo especial de los objetivos [13]. Algunas de las técnicas más populares que caen dentro de esta categoría se examinarán en esta sección.

4.1 VEGA

David Schaffer [14] extendió el programa GENESIS (de Grefenstette [15]) a fin de que pudiera manejar objetivos múltiples. La idea de Schaffer fue usar una extensión del algoritmo genético simple (SGA por sus siglas en inglés) a la que él llamó el *Vector Evaluated Genetic Algorithm* (VEGA), y que difería del primero únicamente en la forma en que se efectuaba la selección. Este operador se modificó de tal manera que a cada generación se generaba un cierto número de subpoblaciones efectuando una selección proporcional de acuerdo a la función objetivo en turno. De tal forma, para un problema con k objetivos, se generaban k subpoblaciones de tamaño N/k cada una (suponiendo un tamaño total de población de N). Estas subpoblaciones se mezclarían luego entre sí a fin de obtener una nueva población de tamaño N, en la que el AG podría aplicar los operadores de cruza y mutación de la forma usual.
Schaffer se percató de que las soluciones generadas por su sistema eran no dominadas localmente, porque esta no dominancia se limitaba a la población en proceso y aunque un individio dominado localmente es dominado también globalmente, esta propiedad no aplica necesariamente a la no dominancia [14]. Un individuo que es no dominado en una cierta generación, puede resultar dominado por un individuo que emerja en una generación posterior. También advirtió un problema al que se denomina en genética como "especiación", el cual consiste en que ciertas "especies" con características específicas evolucionan en direcciones distintas. Este problema se origina porque esta técnica selecciona individuos que son excelentes en una cierta dimensión de desempeño, pero

no necesariamente en las otras. El peligro potencial con esto es que podríamos tener individuos con lo que Schaffer llama desempeño "mediatizado[5]" en todas las dimensiones, el cual podría ser muy útil para soluciones compromiso, pero que no sobrevivirá bajo este esquema de selección, puesto que no es el mejor en ninguno de los objetivos del problema, sino sólo moderadamente bueno en todos. La especiación es un fenómeno indeseable en este contexto porque se contrapone a nuestro objetivo de encontrar una solución compromiso. Schaffer sugirió algunas heurísticas para lidiar con este problema. Por ejemplo, propuso usar una selección mediante preferencias heurísticas para los individuos no dominados en cada generación, a fin de proteger a nuestros cromosomas "mediatizados". Asimismo, propuso motivar la cruza entre las diferentes "especies" agregando alguna heurística específica en vez de hacer cruza aleatoria como en el AG tradicional.

Análisis de la técnica

Aunque Schaffer reportó algo de éxito con su técnica, y ésta es suficientemente fácil de implementar como para estar tentado a probarla, Richardson et al. [16] advirtieron que el barajar y mezclar todas las subpoblaciones es realmente equivalente a promediar las aptitudes de cada uno de los objetivos. Puesto que Schaffer usó selección proporcional [17], estas aptitudes son además proporcionales a los objetivos mismos [2]. Por lo tanto, la aptitud esperada resultante corresponde a una combinación lineal de los objetivos donde los pesos dependen de la distribución de la población a cada generación, tal y como lo indican Richardson et al. [16]. La principal consecuencia de esto es que cuando tenemos una superficie compromiso cóncava, una simple combinación lineal de los objetivos no podrá encontrar algunos de los puntos de esta superficie, sin importar qué conjunto de pesos se use para combinar los objetivos [16].

4.2 Ordenamiento Lexicográfico

En este método, los objetivos se jerarquizan en un orden de importancia definido por el diseñador. La solución óptima $\bar{x}^*$ se obtiene posteriormente minimizando las funciones objetivo, empezando con la más importante y procediendo de acuerdo con el orden de importancia asignado a cada uno de los objetivos. Supongamos que los subíndices de los objetivos indican no sólo el número de función objetivo al que corresponden, sino también la prioridad de cada objetivo. Tendremos entonces que $f_1(\bar{x})$ y $f_k(\bar{x})$ denotan las funciones objetivo más y menos importantes, respectivamente. Entonces, el primer problema se formula como

$$\text{Minimizar } f_1(\bar{x}) \tag{17}$$

sujeta a

$$g_j(\bar{x}) \leq 0; \quad j = 1, 2, \ldots, m \tag{18}$$

y su solución $\bar{x}_1^*$ y $f_1^* = f_1(\bar{x}_1^*)$, se obtiene de ahí. Posteriormente, el segundo problema se formula como

---

[5]Con "mediatizado", Schaffer se refiere a un individuo que tenga buen desempeño en todos los objetivos (quizás por encima del promedio), pero que no es el mejor en ninguno de ellos en particular.

$$\text{Minimizar } f_2(\bar{x}) \tag{19}$$

sujeta a

$$g_j(\bar{x}) \leq 0; \quad j = 1, 2, \ldots, m \tag{20}$$

$$f_1(\bar{x}) = f_1^* \tag{21}$$

y la solución a este problema se obtiene como $\bar{x}_2^*$ y $f_2^* = f_2(\bar{x}_2^*)$. Este procedimiento se repite hasta que todos los *k* objetivos han sido considerados. El *i*-ésimo problema está dado por

$$\text{Minimizar } f_i(\bar{x}) \tag{22}$$

sujeta a

$$g_j(\bar{x}) \leq 0; \quad j = 1, 2, \ldots, m \tag{23}$$

$$f_l(\bar{x}) = f_l^*, \quad l = 1, 2, \ldots, i-1 \tag{24}$$

La solución obtenida al final, es decir $\bar{x}_k^*$, se considera como la solución deseada al problema $\bar{x}^*$.

Análisis de la técnica

Seleccionar aleatoriamente un objetivo es equivalente a una combinación ponderada de objetivos, en la que cada peso se define en términos de la probabilidad de que cada objetivo sea seleccionado. Sin embargo, el uso de selección mediante torneo con esta técnica presenta una importante diferencia con respecto a otras metodologías tales como VEGA, porque las comparaciones binarias de la selección mediante torneo hace que no tenga importancia la información sobre el escalamiento de los objetivos [2]. Esto significa que esta técnica puede ver una superficie cóncava como si fuera convexa, aunque eso realmente depende de la distribución de la población y del problema mismo. Su principal desventaja es que esta técnica tiende a favorecer más a ciertos objetivos cuando el problema en cuestión tiene muchos, debido a la aleatoreidad involucrada en el proceso, y esto tenderá a ser una consecuencia indeseable que hará que la población converja a una parte en particular del frente de Pareto en vez de delinearlo completamente [18].

4.3 Uso de géneros para identificar objetivos

Robin Allenson [19] usó una técnica poblacional basada en VEGA en la que se usan géneros para distinguir entre las dos funciones objetivo de un problema que consiste en la planeación de una ruta compuesta por un número de segmentos rectos de tubo. Con esta técnica, sólo se permite la cruza entre macho-hembra, y los géneros de cada individuo se generan aleatoriamente al momento de nacer. En la población inicial, Allenson se aseguró de que hubiese un número igual de machos y hembras, pero este balance no se mantiene después de aplicar los operadores genéticos. A cada generación, el peor individuo (elegido de acuerdo a uno de los dos géneros) es eliminado y reemplazado por otro (elegido aleatoriamente) del mismo género. Allenson usó estrategias evolutivas para implementar una especie de atractores sexuales que modificaban la forma en que se efectúa la cruza. La idea fue modelar la atracción sexual que algunos individuos tienen hacia otros en la naturaleza, lo que determina una cruza no del todo aleatoria. Lis y Eiber [20] también incorporaron géneros en su AG, pero en un sentido más general. En este caso, el número de géneros (o sexos) no se limitó a dos, sino que podían ser tantos como objetivos se tuvieran. Otra diferencia de esta técnica es que el operador de cruza se modificó de tal forma que permitiera reproducción panmítica, en la que varios padres generan un solo hijo (en vez de que dos padres generen dos hijos como en el algoritmo genético tradicional). La idea era seleccionar un padre para cada sexo a fin de que contribuyera con la cantidad más grande posible de genes (en caso de empate, entonces el sexo se elegía aleatoriamente del de los padres que contribuyeron con la misma cantidad de genes). Si no se efectuaba la cruza, entonces uno de los individuos en la generación anterior se copiaba exactamente igual (incluyendo su sexo) a la siguiente generación. En esta técnica, los individuos son evaluados usando diferentes funciones de aptitud (de acuerdo con su sexo). El operador de mutación sólo está ligeramente restringido para evitar cambios en el sexo de un individuo. Conforme progresan las generaciones, se actualiza una lista de individuos no dominados, removiendo de ella a todos los individuos que dejen de ser no dominados. Al final, la lista contendrá las soluciones óptimas en el sentido de Pareto.

Análisis de la técnica

El uso de géneros es realmente otra forma de definir subpoblaciones independientes para cada objetivo. La diferencia de esta técnica con VEGA [14] es que Lis y Eiber usaron cruza panmítica, lo que introduce restricciones en la mezcla de individuos, evitando la cruza aleatoria de individuos entre diferentes subpoblaciones que Schaffer efectuara. Sin embargo, conforme incrementamos el número de objetivos (o géneros), tendremos muchas subpoblaciones y la cruza panmítica se volverá más ineficiente (computacionalmente hablando), porque necesitaremos usar más padres para generar un hijo. Adicionalmente, el tamaño de la población tendrá que ser suficientemente grande conforme incrementemos el número de objetivos, a fin de mantener una distribución razonable de géneros a través de toda la población.

4.4 Uso de Min-Max con pesos

La idea de definir el *óptimo min-max* y de aplicar este concepto a problemas de optimización con objetivos múltiples fue tomada de la teoría de juegos, que lidia con la solución de situaciones conflictivas. El óptimo min-max compara las desviaciones relativas de los mínimos deseados por separado. Consideremos la *i*-ésima función objetivo para la cual la desviación relativa puede calcularse como

$$z_i'(\bar{x}) = \frac{\left| f_i(\bar{x}) - f_i^0 \right|}{\left| f_i^0 \right|} \qquad (25)$$

o como

$$z_i''(\bar{x}) = \frac{\left| f_i(\bar{x}) - f_i^0 \right|}{\left| f_i(\bar{x}) \right|} \qquad (26)$$

Debe aclararse que para (25) y (26) tenemos que suponer que para toda $i \in I$ y para toda $\bar{x} \in F$, $f_i(\bar{x}) \neq 0$. Si todas las funciones objetivo van a minimizarse, entonces la ecuación (25) define incrementos relativos, mientras que si todas se van a maximizar, define decrementos relativos. La ecuación (26) trabaja a la inversa.

Hagamos que $\bar{z}(\bar{x}) = \left[z_1(\bar{x}), \ldots, z_i(\bar{x}), \ldots, z_k(\bar{x})\right]^T$ sea un vector de los incrementos relativos que se definen en $\Re^k$. Las componentes del vector $z(\bar{x})$ se evaluarán a partir de la fórmula

$$\forall_{i \in I}(z_i(\bar{x})) = \max \left\{z_i^{'}(\bar{x}), z_i^{''}(\bar{x})\right\} \tag{27}$$

Ahora podemos definir el óptimo min-max de la manera siguiente [3]: conociendo los extremos de las funciones objetivo que pueden obtenerse resolviendo los problemas de optimización de cada criterio por separado, la solución deseable es aquella que proporcione los valores más pequeños para los incrementos relativos de todas las funciones objetivo.

Análisis de la técnica

Esta técnica puede crear una presión de selección muy elevada si ciertas combinaciones de pesos se producen en etapas tempranas de la búsqueda [18]. El uso de repartición de aptitud evitará hasta cierto punto que haya convergencia prematura, pero introduce un factor adicional que no es fácil de determinar y que normalmente es sujeto a experimentación.

4.5 Un Algoritmo Genético No Generacional

Valenzuela-Rendón y Uresti-Charre [21] propusieron un AG que usa selección no generacional y en el cual la aptitud de los individuos se calcula de manera incremental. La idea fue derivada de los Sistemas Clasificadores (LCS, por sus siglas en inglés) [17], en los cuales se ha demostrado que el simple reemplazo del peor individuo en la población seguido de una actualización de la aptitud del resto de la población funciona mejor que un AG tradicional (generacional). En el contexto de optimización con objetivos múltiples, lo que los autores hicieron fue transformar un problema con *N* objectivos en otro con sólo 2 objetivos: la minimización de la cuenta de dominación (un promedio del número de individuos que han dominado a un individuo hasta ese momento) y la minimización del conteo de movimientos del nicho (promedio del número de individuos que se encuentran cerca de acuerdo a una cierta función de repartición de aptitud). Posteriormente, estos 2 objetivos se combinan linealmente, a fin de producir uno solo, que será el que se optimice con el AG.

Análisis de la técnica

Esta técnica es realmente una versión más elaborada de las técnicas de jerarquización ponderadas usadas por Bentley y Wakefield [22], particularmente del método denominado *Weighted Average Ranking* (WAR). Aunque esta técnica parece proporcionar buenas distribuciones, no parece posible incorporar dentro del algoritmo preferencias de los objetivos, lo cual puede ser una desventaja al resolver problemas del mundo real. Asimismo, no parece claro cómo definir los parámetros adicionales requeridos por este algoritmo, los cuales parecen estar sujetos a un ajuste empírico similar al de los parámetros tradicionales del AG (por ejemplo, los porcentajes de cruza y mutación).

4.6 Uso de pesos generados aleatoriamente y elitismo

Ishibuchi y Murata [23] propusieron un algoritmo similar a la técnica min-max con pesos de Hajela, pero en este caso los pesos se generan de una forma ligeramente diferente, y el conjunto de soluciones no dominadas que se produce en cada generación se mantiene aparte de la población actual.

Análisis de la técnica

Esta técnica es muy similar a la que Bentley y Wakefield [22] llaman *Sum of Weighted Ratios* (SWR) y a las funciones de valores de los atributos usadas por Greenwood et al. [24]. Bentley y Wakefield [22] afirman que esta técnica mantiene suficiente diversidad como para retener una amplia gama de soluciones a través de muchas generaciones. Sin embargo, Coello [18] ha mostrado (usando una técnica similar), que esta distribución de valores no puede mantenerse en problemas en los que hay un objetivo del vector ideal que puede alcanzarse fácilmente. En tal caso, es necesario usar una técnica de repartición de aptitud o una técnica de búsqueda local (como la propuesta por Ishibuchi y Murata [23]) para mantener la diversidad. Bentley y Wakefield [22] mostraron otra variante de este algoritmo a la que llamaron *Sum of Weighted Global Ratios* (SWGR), la cual reduce visiblemente la distribución de soluciones producidas (es decir, el tamaño del conjunto de Pareto) usando los mejores y peores individuos globales en vez de los locales. La idea es, sin embargo, interesante, y la implementación del algoritmo no sólo parece ser fácil, sino también muy eficiente con respecto a la mayoría de las técnicas basadas en el óptimo de Pareto que se describen a continuación.

5. Técnicas basadas en el óptimo de Pareto

Las técnicas englobadas en esta sección son las que se basan en el concepto del óptimo de Pareto para seleccionar individuos de la población. Estas son las técnicas más populares en la literatura, aunque también presentan ciertos problemas al igual que las técnicas anteriores.

5.1 Jerarquización de Pareto

La idea de usar asignación de aptitud basada en el óptimo de Pareto fue propuesta originalmente por Goldberg [17] para resolver los problemas de la técnica de Schaffer. Goldberg sugirió el uso de asignación de jerarquías y selección basados en no dominancia para mover la población hacia el frente de Pareto. La idea básica es encontrar el conjunto de cadenas en la población que sean no dominadas (en el sentido de Pareto). A estas cadenas se les asigna entonces la jerarquía más alta y se remueven de la población. De la población restante se obtiene otro conjunto de cadenas no dominadas, a las que se les asigna la siguiente jerarquía en turno. Este proceso se repite hasta que toda la población esté jerarquizada. Goldberg también sugirió el uso de algún tipo de técnica de nichos, como la repartición de aptitud [25], a fin de evitar que el AG converja hacia un solo punto del frente de Pareto.

Análisis de la técnica

El principal problema con la jerarquización de Pareto es que, en general, no hay un algoritmo eficiente para checar no dominancia en un conjunto de soluciones factibles [18]. Los algoritmos tradicionales tienen una fuerte degradación en su desempeño conforme incrementamos el tamaño de la población y el número de objetivos. Asimismo, el uso de repartición de aptitud hace necesario estimar un parámetro adicional (σ_r), del cual depende en gran medida el desempeño de la técnica y que, desgraciadamente, no resulta fácil determinar.

5.2 MOGA (Multiple Objetive Genetic Algorithm)

Fonseca y Fleming [26] propusieron una técnica en la cual la jerarquía de un cierto individuo corresponde al número de cromosomas por los cuales es dominado en la población actual. Consideremos, por ejemplo, a un individuo x_i en la generación t, el cual es dominado por p_i^t individuos en la generación actual. Su posición actual en la jerarquía puede obtenerse mediante [26]:

$$\text{jerarquía}(x_i, t) = 1 + p_i^t \qquad (28)$$

A todos los individuos no dominados se les asigna la jerarquía 1, mientras que a los dominados se les penaliza de acuerdo a la densidad de la población en la región correspondiente de la superficie de compromiso. La asignación de aptitud se efectúa de la manera siguiente [26]:

1. Ordenar la población de acuerdo a la jerarquía.

2. Asignar aptitud a los individuos interpolando entre el mejor (jerarquía 1) y el peor (jerarquía $n \le N$) en la forma propuesta por Goldberg [17], de acuerdo a alguna función la cual es usualmente, pero no necesariamente, lineal.

3. Promediar las aptitudes de los individuos que tengan la misma jerarquía, de manera que todos ellos sean muestreados con la misma probabilidad. Este procedimiento mantiene la aptitud global de la población constante, a la vez que mantiene una presión de selección apropiada de acuerdo a la función utilizada.

Goldberg y Deb [27] han hecho ver que este tipo de asignación de aptitud con bloqueo suele producir una gran presión de selección que puede orillar a una convergencia prematura. Para evitar este problema, Fonseca y Fleming usaron un método de formación de nichos para distribuir la población sobre el frente de Pareto, pero en vez de efectuar repartición de aptitud en base a los valores de los parámetros, la efectuaron en base a los valores de las funciones objetivo [28].

Análisis de la técnica

En MOGA, la repartición de aptitud se efectúa en el espacio de los valores de las funciones objetivo, lo que significa que dos vectores diferentes con los mismos valores para sus funciones objetivo no pueden existir simultáneamente en la población usando esta técnica. Esto es, aparentemente indeseable, porque éste es precisamente el tipo de soluciones que normalmente estamos buscando, aunque debe decirse que en la práctica la técnica de Fonseca y Fleming parece funcionar bastante bien [18]. MOGA es una buena técnica, eficiente y relativamente fácil de implementar, aunque, al igual que todas las demás técnicas de jerarquización de Pareto, su desempeño depende en gran medida de una selección adecuada del factor de repartición de aptitud. Sin embargo, cabe mencionar que Fonseca y Fleming [26] desarrollaron una buena metodología para calcular este valor.

5.3 NSGA (Non-dominated Sorting Genetic Algorithm)

Esta técnica fue propuesta por Srinivas y Deb [28], y se basa en el uso de varias capas de clasificación de los individuos. Antes de efectuar la selección, la población es jerarquizada en base a la no dominación: todos los individuos no dominados se clasifican en una misma categoría (con un valor arbitrario de aptitud, el cual es proporcional al tamaño de la población, a fin de proporcionar un potencial reproductivo equitativo para estos individuos). Para mantener la diversidad de la población, se efectúa una repartición de aptitud entre estos individuos clasificados usando los valores arbitrarios de aptitud previamente definidos. Posteriormente, este grupo de individuos clasificados se ignora, considerándose otra capa de individuos no dominados. Este proceso se repite hasta que todos los individuos de la población estén clasificados. Esta técnica utiliza selección proporcional (la variante llamada *stochastic remainder*). Puesto que los individuos en el primer frente tienen el máximo valor de aptitud, siempre obtienen más copias que el resto de la población. Esto permite buscar regiones no dominadas, y produce convergencia rápida de la población hacia estas regiones. La repartición de aptitud, por su parte, ayuda a distribuir los individuos sobre esta región. La eficiencia del NSGA yace en la forma en que los objetivos múltiples de un problema se reducen a una función de aptitud arbitraria usando un procedimiento de ordenamiento basado en no dominancia. Esta técnica puede lidiar con cualquier número de objetivos [28], tanto en problemas de maximización como de minimización.

Análisis de la técnica

En este caso, la repartición de aptitud se efectúa en base a los valores de los parámetros en vez de en base a los valores de las funciones objetivo, lo cual pretende asegurar que se lleve a cabo una mejor distribución de los individuos, así como permitir que existan múltiples soluciones equivalentes. Sin embargo, esta técnica es más ineficiente (tanto computacionalmente como en términos de la calidad de los frentes de Pareto que produce) que MOGA, y es más sensible al valor del factor de repartición de aptitud σ_r [18].

5.4 NPGA (Niched Pareto GA)

Horn y Nafpliotis [29] propusieron una selección mediante torneo basado en dominancia de Pareto. En vez de limitar la comparación a dos individuos, se usa un conjunto más grande

(típicamente de 10 individuos) para efectuar el torneo del cual el individuo no dominado resultará triunfador. Si existe un empate entre dos individuos (ninguno domina al otro), entonces se utiliza repartición de aptitud [25]. Normalmente se deben usar poblaciones considerablemente mayores que lo usual para hacer tolerable el ruido producido por la técnica de selección utilizada para los nichos que emergen de la población [2].
Horn y Nafpliotis [29] llegaron a una forma de repartición de aptitud en el dominio de las funciones objetivo, y sugirieron el uso de una métrica en la que se combinan los dominios de las variables de decisión con el de las funciones objetivo, lo que condujo a lo que ellos denominaron *repartición anidada de aptitud.*

Análisis de la técnica

Puesto que esta técnica no aplica la selección de Pareto a toda la población, sino sólo a un segmento de ella en cada corrida, su ejecución resulta muy rápida y produce buenos frentes de Pareto que suelen mantenerse durante un gran número de generaciones [18]. Sin embargo, además de requerir un factor de repartición de aptitud, esta técnica también requiere que se elija cuidadosamente el tamaño del torneo, ya que de él depende en gran medida su desempeño.

6. Rutas futuras de investigación

Aunque se ha hecho mucho trabajo en esta área, la mayor parte de la investigación actual se ha concentrado en la aplicación de técnicas convencionales o *ad-hoc* para resolver ciertos problemas del mundo real interesantes. Por lo tanto, todavía existen varios temas de interés y problemas sin resolver, de entre los cuales los principales se describirán a continuación:

- Puesto que el tamaño del conjunto de Pareto es normalmente muy grande, y al usar el AG, éste está limitado por el tamaño de la población, es deseable en algunos casos encontrar formas de reducir el número de elementos de este conjunto de tal manera que se le facilite la labor de análisis al tomador de decisiones. Cunha, Oliveira y Covas [30] propusieron la incorporación del algoritmo de Roseman y Gero [31] en el AG, a fin de agrupar los puntos que se encuentren a cierta distancia (definida por el usuario) entre sí en el frente de Pareto. Sin embargo, se ha realizado poco trabajo a este respecto.

- Probablemente uno de los problemas más difíciles dentro de la optimización con objetivos múltiples sea cómo determinar la forma de medir la calidad de una solución. Hasta el momento, prácticamente la única forma de hacerlo es mediante inspección visual, a menos que exista cierto conocimiento previo acerca de la ubicación del frente de Pareto (en cuyo caso es innecesario usar una técnica de optimización con objetivos múltiples). Fonseca y Fleming [26] propusieron la definición de ciertas metas (arbitrarias) que deseamos que sean alcanzadas por el AG; posteriormente, podemos efectuar múltiples corridas del AG y aplicar procedimientos estadísticos no paramétricos estándares para evaluar la calidad de las soluciones (es decir, los frentes no dominados) producidas por la técnica en estudio, a fin de compararla con otras técnicas similares. Sin embargo, estas metas arbitrarias tampoco son fáciles de definir, y se necesita realizar más investigación en esta dirección a fin de desarrollar una forma adecuada y justa de medir la calidad de las soluciones producidas por las diferentes técnicas de optimización con objetivos múltiples.

• En algunos casos puede ser necesario asignar más importancia a cierto objetivo u objetivos. Resulta curioso hacer notar que en estos casos las funciones agregadas nos permiten cambiar con gran facilidad la importancia de los objetivos, mientras que las técnicas de jerarquización (es decir, las basadas en el óptimo de Pareto) no proporcionan normalmente los medios para hacer esto de manera directa. Fonseca y Fleming [26] propusieron el uso de una función de utilidad combinada con MOGA [2] para producir un método para la articulación progresiva de preferencias. La idea que propusieron fue tener un ciclo de realimentación entre el tomador de decisiones y el AG, de forma que ciertas soluciones (del conjunto de Pareto) reciban mayor preferencia que otras. Idealmente, este proceso puede hacerse automáticamente reemplazando al tomador de decisiones con un sistema experto [26]. El problema de esta propuesta es que se requiere conocimiento previo de los rangos de cada función objetivo, lo cual podría resultar excesivamente costoso o incluso imposible en algunos casos. Bentley y Wakefield [22] propusieron el uso de pesos para estimar la importancia de las soluciones que ya estuvieran identificadas como óptimas en el sentido de Pareto, en un intento por superar los problemas con la técnica de Fonseca. Asimismo, Greenwood et al. [24] propusieron, en una técnica más elaborada, un compromiso entre las funciones agregadas (el uso de pesos) y las técnicas de jerarquización. En esta propuesta, debe definirse el nivel de preferencia de cada objetivo. Greenwood et al. [24] usaron una técnica llamada *Iterative Specified Multi-Attribute valUe Theory* (ISMAUT) [32] la cual, combinada con un AG, permite que el AG defina las preferencias por sí mismo, en vez de requerir la intervención de un humano. Sin embargo, el tomador de decisiones todavía debe decidir qué área en particular de la superficie compromiso quiere explorar, de forma que el AG restrinja la búsqueda a esa área específica. Adicionalmente, Greenwood et al. [24] definieron cierta métrica que nos permita obtener un solo valor (o función de utilidad) que guíe la búsqueda a la región de Pareto en particular que resulte de interés para el tomador de decisiones. Finalmente, Voget y Kolonko [33] propusieron el uso de un controlador difuso que regula la presión de selección automáticamente usando un conjunto de metas predefinidas que describen el comportamiento 'deseable' de la población. Un aspecto interesante de este trabajo es que realmente combinan la jerarquización de Pareto con VEGA durante la misma corrida del AG, a fin de permitir la reducción deseada de desviaciones de las metas especificadas por los autores [33]. Estas 3 propuestas son muy interesantes, pero se requiere todavía más trabajo en esta área, preferentemente con problemas del mundo real, de forma que puedan derivarse lineamientos más generales de las diferentes técnicas propuestas.

• Directamente relacionado con el problema de medir la calidad de una solución se encuentra el problema de contar con un conjunto de funciones de prueba para las técnicas de optimización con objetivos múltiples. Este conjunto debe incluir problemas con y sin restricciones[6], ejemplos con pocos objetivos (2 ó 3) que sean adecuados para inspección visual, problemas con pocas y varias variables y problemas en los que sea posible obtener el vector ideal (estos problemas podrían usarse para ajustar inicialmente los parámetros de una nueva técnica a analizarse). Asimismo, también hay necesidad de efectuar estudios detallados de desempeño de diferentes AGs (suponiendo ciertas medidas de calidad) usando estos problemas de prueba y derivar información

---

[6]La mayor parte de los artículos actuales que introducen una nueva técnica evolutiva de optimización con objetivos múltiples usan 2 ó 3 ejemplos con 2 objetivos y sin restricciones, particularmente las propuestas originalmente por Schaffer [14].

más precisa sobre el comportamiento de cada uno de los algoritmos usados. Coello [18] condujo un estudio de este tipo usando varios problemas de ingeniería, pero se requiere extender este tipo de trabajo a problemas más generales.

• También es importante definir criterios de detención para las técnicas evolutivas de optimización con objetivos múltiples, porque no es obvio saber en qué momento la población ha alcanzado un punto a partir del cual no puede lograrse mayor progreso. Actualmente, las maneras típicas de detención son el uso de un número máximo de generaciones o el monitoreo de la población a ciertos intervalos sumados a la interpretación visual de los resultados.

• El uso de repartición de aptitud en estas técnicas introduce otro problema, porque el valor de σ_r se convierte en un parámetro más con el que el usuario tiene que experimentar hasta encontrarse un valor razonable. Aunque se ha efectuado trabajo importante en esta área (ver por ejemplo Deb y Goldberg [34] y Fonseca y Fleming [26]), la mayor parte de la investigación realizada hasta el momento se concentra en optimización con un solo objetivo o multimodal.

• Finalmente, un tema muy importante que ha sido abordado sólo superficialmente por los investigadores es el desarrollo de una teoría que pueda explicar cuestiones tales como el efecto de los parámetros utilizados (tamaño de la población, porcentajes de cruza y mutación y tamaño de los nichos) y la forma en que la técnica de selección utilizada afecta el desempeño de un algoritmo.

Conclusiones

En este trabajo se han presentado las técnicas evolutivas de optimización con objetivos múltiples más populares reportadas en la literatura en los últimos años. Además de describir brevemente cada algoritmo, cada uno de ellos se ha analizado, mostrando sus fortalezas y debilidades. Finalmente, se han descrito brevemente algunos de los problemas más interesantes que permanecen sin resolver en esta área, a fin de motivar a los estudiantes e investigadores a incursionar en esta disciplina que se ha tornado tan popular en los últimos años.

Bibliografía

[1] Tamaki, H., Kita, H., y Kobayashi, S., "Multi-Objective Optimization by Genetic Algorithms : A Review", En T. Fukuda & T. Furuhashi (Eds.), *Proceedings of the 1996 International Conference on Evolutionary Computation*, Nagoya, Japón, 1996, pp. 517-522, IEEE Press, 1996.

[2] Fonseca, C. M. y Fleming, P. J. "An overview of evolutionary algorithms in multiobjective optimization", *Evolutionary Computation*, **3**(1), Spring, pp. 1-16, 1995.

[3] Osyczka, A., *Multicriterion Optimization in Engineering with FORTRAN programs*, Ellis Horwood Limited, 1984.

[4] Pareto, V., *Cours D'Economie Politique*, Volume I and II, F. Rouge, Lausanne, 1896.

[5] Rosenberg, R. S., *Simulation of genetic populations with biochemical properties*, Ph. D. thesis, University of Michigan, Ann Harbor, Michigan, 1967.

[6] Kuhn, H. W. y Tucker, A. W. "Nonlinear programming", En J. Neyman (Editor), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, California, pp. 481-492, University of California Press, 1951.

[7] Ritzel, B. J., Eheart, J. W., y Ranjithan, S. "Using genetic algorithms to solve a multiple objective groundwater pollution containment problem", *Water Resources Research*, **30**(5), may, pp. 1589-1603, 1994.

[8] Duckstein, L. "Multiobjective Optimization in Structural Design: The Model Choice Problem", En E. Atrek, R. H. Gallagher, K. M. Ragsdell, y O. C. Zienkiewicz (Editores), *New Directions in Optimum Structural Design*, pp. 459-481, John Wiley and Sons, 1984.

[9] Haimes, Y. Y., Hall, W., y Freedman, H., *Multi-Objective Optimization in Water Resources Systems: The Surrogate Trade-Off Method*, Elsevier, Amsterdam, 1975.

[10] Ignizio, J. P., "The determination of a subset of efficient solutions via goal programming", *Computing and Operations Research*, **3**, pp. 9-16, 1981.

[11] Chen, Y. L. y Liu, C. C. "Multiobjective VAR planning using the goal-attainment method", *IEE Proceedings on Generation, Transmission and Distribution*, **141**(3), may, pp. 227-232, 1994.

[12] Wilson, P. B. y Macleod, M. D. "Low implementation cost IIR digital filter design using genetic algorithms", *IEE/IEEE Workshop on Natural Algorithms in Signal Processing*, Chelmsford, U.K., pp. 4/1-4/8, 1993.

[13] Powell, D. and Skolnick, M. M. "Using genetic algorithms in engineering design optimization with non-linear constraints", En S. Forrest (Editora), *Proceedings of the Fifth International Conference on Genetic Algorithms*, San Mateo, California, julio, pp. 424-431, University of Illinois at Urbana-Champaign, Morgan Kaufmann Publishers, 1993.

[14] Schaffer, J. D. "Multiple objective optimization with vector evaluated genetic algorithms", En *Genetic Algorithms and their Applications: Proceedings of the First International Conference on Genetic Algorithms*, pp. 93-100, Lawrence Erlbaum, 1985.

[15] Grefenstette, J. J. "GENESIS: A system for using genetic search procedures", *Proceedings of the 1984 Conference on Intelligent Systems and Machines*, pp. 161-165, 1984.

[16] Richardson, J. T., Palmer, M. R., Liepins, G., y Hilliard, M., "Some guidelines for genetic algorithms with penalty functions", En J. D. Schaffer (Editor), *Proceedings of the Third International Conference on Genetic Algorithms*, George Mason University, pp. 191-197, Morgan Kaufmann Publishers, 1989.

[17] Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1989.

[18] Coello, C. A. C. *An Empirical Study of Evolutionary Techniques for Multiobjective Optimization in Engineering Design*, Ph. D. thesis, Department of Computer Science, Tulane University, New Orleans, LA, 1996.

[19] Allenson, R. "Genetic algorithms with gender for multi-function optimisation", Technical Report EPCC-SS92-01, Edinburgh Parallel Computing Centre, Edinburgh, Scotland, 1992.

[20] Lis, J. y Eiben, A. E. "A Multi-Sexual Genetic Algorithm for Multiobjective Optimization", En T. Fukuda y T. Furuhashi (Eds.), *Proceedings of the 1996 International Conference on Evolutionary Computation*, Nagoya, Japón, pp. 59-64. IEEE Press, 1996.

[21] Valenzuela-Rendón, M. y Uresti-Charre, E. "A Non-Generational Genetic Algorithm for Multiobjective Optimization", En T. Bäck (Editor), *Proceedings of the Seventh International Conference on Genetic Algorithms*, San Mateo, California, July, pp. 658-665, Michigan State University, Morgan Kaufmann Publishers, 1997.

[22] Bentley, P. J. y Wakefield, J. P. "Finding Acceptable Solutions in the Pareto-Optimal Range using Multiobjective Genetic Algorithms", *Proceedings of the 2nd On-Line World Conference on Soft Computing in Engineering Design and Manufacturing* (WSC2), June, *http://users.aol.com/docbentley/dispaper.htm*, 1997.

[23] Ishibuchi, H. y Murata, T., "Multi-Objective Genetic Local Search Algorithm", En T. Fukuda y T. Furuhashi (Eds.), *Proceedings of the 1996 International Conference on Evolutionary Computation*, Nagoya, Japón, pp. 119-124. IEEE Press, 1996.

[24] Greenwood, G. W., Hu, X. S., y D'Ambrosio, J. G. "Fitness Functions for Multiple Objective Optimiztion Problems: Combining Preferences with Pareto Rankings", En R. K. Belew y M. D. Vose (Eds.), *Foundations of Genetic Algorithms 4*, pp. 437-455, San Mateo, California, Morgan Kaufmann, 1997.

[25] Goldberg, D. E. y Richardson, J. "Genetic algorithm with sharing for multimodal function optimization", En J. J. Grefenstette (Editor), *Genetic Algorithms and Their Applications: Proceedings of the Second International Conference on Genetic Algorithms*, pp. 41-49, Lawrence Erlbaum, 1987.

[26] Fonseca, C. M. y Fleming, P. J. "Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization", En S. Forrest (Editor), *Proceedings of the Fifth International Conference on Genetic Algorithms*, San Mateo, California, pp. 416-423, University of Illinois at Urbana-Champaign, Morgan Kauffman Publishers, 1993.

[27] Goldberg, D. E. y Deb, K. "A comparison of selection schemes used in genetic algorithms", En G. E. Rawlins (Editor), *Foundations of Genetic Algorithms*, pp. 69-93, San Mateo, California, Morgan Kaufmann, 1991.

[28] Srinivas, N. y Deb, K. "Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms", *Evolutionary Computation*, **2**(3), fall, pp. 221-248, 1994.

[29] Horn, J. y Nafpliotis, N. "Multiobjective Optimization using the Niched Pareto Genetic Algorithm", Technical Report IlliGAl Report 93005, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA, 1993.

[30] Cunha, A. G., Oliveira, P., y Covas, J. A. "Use of Genetic Algorithms in Multicriteria Optimization to Solve Industrial Problems", En T. Bäck (Editor), *Proceedings of the Seventh International Conference on Genetic Algorithms*, San Mateo, California, July, pp. 682-689, Michigan State University, Morgan Kaufmann Publishers, 1997.

[31] Roseman, M. A. y Gero, J. S., "Reducing the Pareto optimal set in multicriteria optimization", *Engineering Optimization*, **8**, pp. 189-206, 1985.

[32] White, C., Sage, A., y Dozono, S., "A model of multiattribute decision-making and tradeoff weight determination under uncertainty", *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-14**, pp. 223-229, 1984.

[33] Voget, S. y Kolonko, M., "Multidimensional Optimization with a Fuzzy Genetic Algorithm", *Journal of Heuristics*, (por aparecer).

[34] Deb, K. y Goldberg, D. E., "An investigation of niche and species formation in genetic function optimization", En J. D. Schaffer (Editor), *Proceedings of the Third International Conference on Genetic Algorithms*, San Mateo, California, junio, pp. 42-50, George Mason University, Morgan Kaufmann Publishers, 1989.