



Dirección General de Educación Superior Tecnológica

Instituto Tecnológico de Veracruz

***“UNA PROPUESTA DE MODELADO MATEMÁTICO
PARA EL PROBLEMA DE BALANCEO DE
ECUACIONES QUÍMICAS Y SU RESOLUCIÓN
APLICANDO UN ALGORITMO GENÉTICO”***

**OPCIÓN I:
TÉSIS**

**PARA OBTENER EL TÍTULO DE:
INGENIERO EN SISTEMAS COMPUTACIONALES**

**PRESENTA:
LETICIA PALOS SÁNCHEZ**

**ASESORES:
M.C. RAFAEL RIVERA LÓPEZ
I.Q. MARCO ANTONIO CRUZ CHAVEZ**

Resumen

El balanceo de ecuaciones químicas es estudiado por la comunidad científica y se han desarrollado diferentes métodos para resolver dicho problema. Algunos de los métodos propuestos resuelven una formulación matemática. La presente tesis propone un modelo matemático para resolver cualquier ecuación química. Para resolver dicho modelo se aplica una metaheurística basada en el Cómputo Evolutivo para balancear una ecuación química. Este enfoque se utiliza como una alternativa a los métodos tradicionales de balanceo de ecuaciones químicas, ya que estos no son algoritmos generales que se puedan aplicar a cualquier formulación química.

El presente trabajo de investigación se estructura en 4 capítulos que se describen a continuación:

El capítulo I contiene el marco metodológico, mencionando los antecedentes de la investigación, para luego definir el planteamiento del problema y la justificación de la investigación señalando la importancia que ésta tiene.

En el capítulo II referente al marco teórico se presentan las bases teóricas de la investigación tanto de la Ciencia Química como de la Inteligencia Artificial. Respecto a la ciencia química se muestran los nuevos enfoques que se han estado

desarrollando dentro de la investigación científica. Referente a las ciencias computacionales, en el campo de investigación de la Inteligencia Artificial se define la Computación Evolutiva. Dentro de los algoritmos que forman parte de la Computación evolutiva se encuentran los Algoritmos Genéticos.

En el capítulo III se describe el desarrollo de la formulación matemática, describiéndola de forma general y empleándola en una ecuación de ejemplo. También se describe el algoritmo genético para el balanceo de ecuaciones químicas, junto con la explicación de los operadores genéticos.

El capítulo IV describe la etapa de experimentación y análisis de resultados de la implementación del algoritmo desarrollado, incluyendo los problemas escogidos para las pruebas, de donde se obtiene el análisis de resultados.

ÍNDICE

Contenido	pag.
Resumen	i
Índice	iii
Índice de figuras	v
Índice de tablas.....	vii
CAPÍTULO 1: MARCO METODOLÓGICO	01
1.1.- Antecedentes.....	02
1.2.- Planteamiento del problema	04
1.3.- Objetivos.....	05
1.3.1.- Objetivo general	06
1.3.2.- Objetivos específicos.....	06
1.4.- Justificación del problema	06
1.5.- Alcances y limitaciones.....	08
CAPÍTULO 2: MARCO TEÓRICO	09
2.1.- Antecedentes.....	10
2.2.- Métodos de balance de ecuaciones químicas	13
2.2.1.- Técnicas clásicas.....	13
2.2.2.- Técnicas basadas en matriz	22
2.2.3.- Técnicas heurísticas	24
2.3.- Computación evolutiva	25
2.3.1.- Motivación	26
2.3.2.- Estructura de un algoritmo de computación evolutiva	28
2.3.3.- Ventajas de los algoritmos de computación evolutiva.....	30
2.3.4.- Programación Evolutiva.....	31
2.3.5.- Estrategias Evolutivas	32
2.3.6.- Algoritmos Genéticos.....	33
2.3.7.- Programación Genética	33
2.4.- Algoritmos Genéticos.....	36
2.4.1.- Antecedentes.....	37
2.4.2.- Codificación	40
2.4.3.- Evaluación de los individuos.....	41
2.4.5.- Selección	43
2.4.6.- Cruzamiento	44
2.4.7.- Mutación	45
2.5.- Comentarios Finales.....	46

CAPÍTULO 3: DESCRIPCIÓN DEL ALGORITMO GENÉTICO PARA EL BALANCEO DE ECUACIONES QUÍMICAS	48
3.1.- Introducción	49
3.2.- Formulación matemática	49
3.3.- Descripción general del algoritmo	56
3.3.1.- Codificación	58
3.3.2.- Creación de la población inicial	61
3.3.3.- Selección	63
3.3.4.- Cruzamiento	64
3.3.5.- Mutación	67
3.3.6.- Evaluación	68
3.4.- Comentarios Finales.....	69
CAPÍTULO 4: PRUEBAS Y ANÁLISIS	71
4.1.- Introducción	72
4.2.- Diseño e implementación de la lectura de la ecuación química desde un archivo de texto plano	72
4.2.1.- Diseño	73
4.2.2.- Implementación	75
4.3.- Variables de control	77
4.4.- Selección de problemas	78
4.5.- Convergencia del algoritmo implementando Codificación Directa y Codificación Reducida	81
4.6.- Convergencia del algoritmo implementando evaluación del individuo con f y g	83
4.7.- Análisis de resultados	85
4.8.- Comentarios finales	90
Conclusiones	91
Referencias.....	94

ÍNDICE DE FIGURAS

Figura	pag.
Figura 1.1.- Ejemplo 1 de ecuación química balanceada.	03
Figura 1.2.- Ejemplo 2 de ecuación química balanceada	03
Figura 2.1.- Sistema de ecuaciones de la ecuación 2.3	16
Figura 2.2.- Asignación de 1 a c	16
Figura 2.3.- Sistema de ecuaciones después de la sustitución.....	16
Figura 2.4.- Despeje y resultado de la incógnita d	17
Figura 2.5.- Calculo de número de oxidación.....	18
Figura 2.6.- Elementos con cambio de estado en el método de oxido-reducción. ...	19
Figura 2.7.- Balanceo de masa de semireacciones.	19
Figura 2.8.- Procedimiento de Mn.....	19
Figura 2.9.- Procedimiento de Fe.....	19
Figura 2.10.- Balanceo de electrones intercambiados	20
Figura 2.11.- Resultado de proceso redox.	20
Figura 2.12.- Esquema general de un algoritmo evolutivo.	29
Figura 2.13.- Representación gráfica de un proceso evolutivo.	30
Figura 2.14.- Ejemplo de una mutación de una MEF en programación evolutiva. ...	31
Figura 2.15.- Efecto de una mutación de punto en la programación genética	36
Figura 2.16.- Ejemplo de cruzamiento de sub-árboles en programación genética ..	36
Figura 2.17.- Esquema general de un algoritmo genético	40
Figura 2.18.- Individuo genético entero.....	41
Figura 2.19.- Cruzamiento de P1 con P2.....	45
Figura 2.20.- Mutación del individuo.	46
Figura 3.1.- Sustitución del símbolo químico por E	51
Figura 3.2.- Sustitución del símbolo químico por e_{ij}	52
Figura 3.3.- Sumatoria de los elementos	53
Figura 3.4.- Generalización de la figura 3.3	53
Figura 3.5.- Sustitución del símbolo químico por e_{ij} de la figura 3.4.....	53
Figura 3.6.- Matriz A de la ecuación 3.1.....	54
Figura 3.7.- Algoritmo Genético para el balanceo de ecuaciones químicas.....	57
Figura 3.8.- Codificación directa de la ecuación 3.22.	59
Figura 3.9.- Cromosoma de individuo con genes bloqueados	60
Figura 3.10.- Despeje del sistema de ecuaciones.	61
Figura 3.11.- Cromosoma completo.....	61
Figura 3.12.- Población inicial	62
Figura 3.13.- Vector booleano de los individuos elegidos para la selección	63
Figura 3.14.- Individuos elegidos para la selección	63

Figura 3.15.- Primer enfrentamiento.	64
Figura 3.16.- Segundo enfrentamiento.	64
Figura 3.17.- Cromosoma de los individuos seleccionados como padres.	65
Figura 3.18.- Cromosoma de los hijos creados por el cruce.	66
Figura 3.19.- Cromosoma de los hijos sin cruzamiento.	66
Figura 3.20.- Mutación del hijo 1	67
Figura 3.21.- Individuo con grado de adaptación igual con cero.....	69
Figura 3.22.- Individuo con solución factible y óptima.....	69
Figura 3.23.- Decodificación del cromosoma.	69
Figura 4.1.- Estructura del archivo de un problema	73
Figura 4.2.- Reactivos (EjemploEcuacion)	74
Figura 4.3.- Productos (EjemploEcuacion).....	74
Figura 4.4.- Variables de Control (EjemploEcuacion)	74
Figura 4.5.- Procedimiento para lectura de archivo de entrada.	75
Figura 4.6.- Archivo de salida	76
Figura 4.7.- Archivo de salida 2	77
Figura 4.8.- Convergencia de la CD y CR para el problema 1.	82
Figura 4.9.- Convergencia de la CD y CR para el problema 14.	82
Figura 4.10.- Convergencia de función objetivo GF y G para el problema 1.....	84
Figura 4.11.- Convergencia de función objetivo FG y G para el problema 8.....	85

ÍNDICE DE TABLAS

Tabla	pag.
Tabla 2.1.- Enfoques de estudio del balanceo de ecuaciones químicas	11
Tabla 2.2.- Métodos de balanceo más conocidos.	14
Tabla 2.3.- Conceptos base para la oxido-reducción.	18
Tabla 2.4.- Ejemplo de matriz en [Camp5].	23
Tabla 3.1.- Matriz de subíndice	51
Tabla 3.2.- Relaciones de los coeficientes de la ecuación 3.22	60
Tabla 4.1.- Variables de control.....	78
Tabla 4.2.- Problemas resueltos por métodos exactos.	79
Tabla 4.3.- Características de los problemas	80
Tabla 4.4.- Valores de variables de control para pruebas de convergencia.	81
Tabla 4.5.- Valores de variables de control para pruebas de convergencia.	83
Tabla 4.6.- Espacio de Solución para cada problema	85
Tabla 4.7.- Valores de variables de control para cada problema.	86
Tabla 4.8.- Soluciones del AGBEQ	87
Tabla 4.9.- Resultados de las pruebas experimentales.....	88

CAPÍTULO I

MARCO METODOLÓGICO

1.1.- Antecedentes

El balanceo de ecuaciones químicas (BEQ) juega un papel muy importante tanto en el ambiente académico como en la industria química, ya que es una herramienta para el modelado de reacciones químicas. El BEQ se utiliza tanto para aplicaciones cuantitativas como cualitativas como el estudio y análisis de los procesos químicos, la estimación de reactivos y productos, y la determinación de las condiciones de una reacción química.

En general, una reacción química es un proceso mediante el cual una o más sustancias (elementos o compuestos) denominados reactivos sufren una transformación para dar lugar a sustancias diferentes conocidas como productos. De acuerdo a [Lir11], a una escala microscópica, una reacción química produce que los enlaces químicos entre los elementos de un compuesto se rompan o que se formen nuevos enlaces, debido al desplazamiento de electrones entre los átomos. En base a la ley de la conservación de la masa establecida por Lavoisier, una reacción química no puede modificar la cantidad de elementos químicos.

En la figura 1.1 se muestra un ejemplo gráfico de una ecuación química balanceada. Se ilustra con una esfera a cada elemento químico, en cuyo interior se coloca el símbolo químico del elemento al que representa.

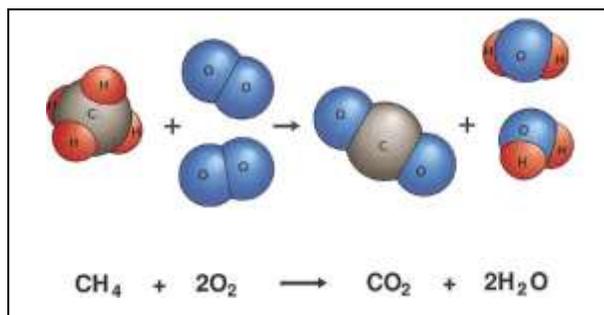


Figura 1.1.- Ejemplo 1 de ecuación química balanceada.

Otro ejemplo de una ecuación balanceada se muestra en la figura 1.2, donde se tienen 3 átomos de Carbono tanto en los reactivos o como en los productos, y así para todos los demás elementos presentes en la ecuación. Como se indica en [BOP06], el balanceo de ecuaciones químicas consiste en encontrar un conjunto de coeficientes que se colocan delante de cada compuesto o elemento químico que aparece en una reacción de manera que el número de átomos de cada elemento sea el mismo en ambos lados de la ecuación.

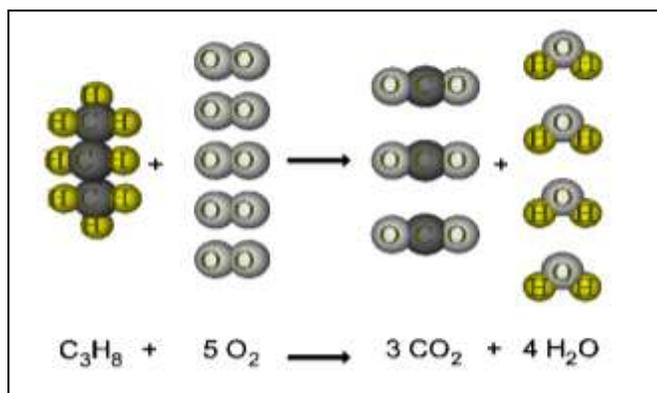


Figura 1.2.- Ejemplo 2 de ecuación química balanceada.

El balanceo de ecuaciones químicas ha sido ampliamente estudiado por la comunidad científica y se han desarrollado diferentes métodos para resolverlo. Los métodos más conocidos en el ámbito académico son el de inspección o tanteo, el oxido-reducción y el método algebraico. Estos métodos se consideran como técnicas

clásicas para resolver el problema del balanceo químico. Por otro lado, y de acuerdo a la literatura consultada, se han desarrollado otros métodos que buscan balancear cualquier ecuación química. La mayoría de estos métodos se enfocan en el uso de matrices, y recientemente se presentó una propuesta basada en una heurística novedosa para el balanceo de ecuaciones químicas.

En cada uno de estos nuevos enfoque se plantea la formulación de un modelo matemático para el problema de balanceo de ecuaciones químicas ([Ris08], [Ris09], [SAS06] y [Zou10]).

También se han desarrollado aplicaciones para computadora utilizando diferentes enfoques para balancear una ecuación química. En [Kum01], Kumar realiza un estudio de las aplicaciones en computadora para el balanceo de ecuaciones químicas, donde se presentan un análisis de métodos con diferentes enfoques.

Aún cuando ya existen varias propuestas para modelar matemáticamente el problema del balanceo de ecuaciones químicas y varios métodos para balancear una ecuación química basados en diferentes técnicas de resolución, el balanceo de ecuaciones químicas sigue siendo del interés de la comunidad científica el estudiar y proponer nuevos enfoques para resolver este tipo de problemas.

1.2.- Planteamiento del Problema

Como cualquier método propuesto para resolver algún problema, todos los métodos de balanceo de ecuaciones químicas existentes tienen ventajas y desventajas. Por ejemplo, en ecuaciones con gran cantidad de compuestos químicos aumenta la complejidad para resolverlo con el método de tanteo [Chu97]; por otro lado, para

utilizar el método algebraico se debe cumplir con la condición de que el número de elementos es igual al número de compuestos más uno [SGC95]. Finalmente, una desventaja para balancear una ecuación química utilizando métodos basados en matrices es que la matriz producida por el problema generalmente no es cuadrada y no es invertible de forma directa [Dep10]. En [SAS06] se propone un enfoque de Programación Lineal Entera (PLE) para el balanceo de ecuaciones químicas haciendo uso de una matriz .

Como es conocido, un problema modelado como PLE tiene un carácter combinatorio y en general se considera, dentro de la teoría de la complejidad computacional, como un problema muy difícil de resolver, por lo que, para este tipo de modelos, se han propuesto un sinnúmero de enfoques de solución, y el uso de metaheurísticas para encontrar buenas soluciones a estos problemas es un enfoque que ha sido estudiado recientemente por la comunidad científica.

Basándose en el hecho de que el problema del balanceo de ecuaciones se ha modelado matemáticamente como un problema de PLE y que existen metaheurísticas que se ha aplicado a este tipo de problemas, se puede formular la siguiente pregunta:

¿Es factible aplicar un algoritmo genético para resolver el problema de balanceo de ecuaciones químicas usando una nueva propuesta de modelado matemático como un problema de programación lineal entera?

1.3.- Objetivos

El presente trabajo de tesis tiene los siguientes objetivos general y específicos:

1.3.1.- Objetivo General

Desarrollar un modelo matemático para el problema de balanceo de ecuaciones químicas resolviéndolo mediante un algoritmo genético que permita hacer búsqueda inteligente en el espacio de soluciones del problema de balanceo de ecuaciones químicas para determinar los coeficientes de los compuestos de cualquier ecuación química que permitan establecer la ecuación balanceada.

1.3.2.- Objetivo Específicos

- Definir un modelo matemático para el problema de balanceo de ecuaciones usando un enfoque de optimización.
- Identificar los elementos que se modifican en el balanceo de ecuaciones químicas.
- Establecer la representación de las soluciones de una ecuación química para su balanceo.
- Establecer las dependencias de los coeficientes estequiométricos entre los elementos de una ecuación química.
- Definir los operadores de selección, cruzamiento y mutación para implementar un algoritmo genético.

1.4.- Justificación

Las ecuaciones químicas desempeñan un papel importante en el plano teórico, así como la química industrial. El balanceo de ecuaciones químicas es un problema con más de un siglo de antigüedad y es uno de los temas más estudiados en la

educación química. Se han definido diversos métodos para realizar el balanceo de las ecuaciones químicas, algunos más complicados que otros.

Como ya se indicó, en el 2006 se propone un enfoque de PLE para el balanceo de ecuaciones químicas. La formulación matemática propuesta es presentada en (1.1).

$$\min \quad f = \sum_{i=1}^n x_i \quad (1.1)$$

$$\text{sujeto a } Ax = 0$$

$$x \geq 1$$

$$x \in Z$$

Donde x_i representa cada coeficiente de los compuestos químicos presentes en la fórmula química a balancear, y $x = [x_1, x_2, \dots, x_n]^T$. La matriz A es la matriz de reacción, donde el elemento a_{ij} de la matriz es el número de átomos de tipo i en cada compuesto o elemento (reactivo/producto) j . a_{ij} es positivo si corresponde a un reactivo y es negativo si es producto. La matriz A tiene una dimensión de $m \times n$, donde m representa a los elementos y n a los compuestos químicos en una la fórmula química. Finalmente, los elementos de x deben ser enteros.

Existen en la literatura varios problemas combinatorios que se modelan como un problema de PLE, por ejemplo el problema del agente viajero y el problema de asignación de horarios escolares. Si se considera al problema del balanceo de ecuaciones químicas como un problema combinatorio, es evidente modelarlo matemáticamente y poderlo resolver usando una metaheurística extraída del Cómputo Evolutivo como un Algoritmo Genético.

1.5.- Hipótesis

Para este trabajo de tesis, se formula la siguiente hipótesis:

“El modelado matemático propuesto permite desarrollar un algoritmo genético para balancear cualquier tipo de ecuaciones químicas”.

1.6.- Alcances y limitaciones

El alcance de esta tesis es desarrollar un modelo matemático que para el balanceo de ecuaciones químicas, que verifique que la existe la misma cantidad de átomos de cada elemento tanto en los reactivos como en los productos. Dicho modelo se resolverá mediante un algoritmo genético para balancear ecuaciones químicas con un enfoque de optimización buscando el valor mínimo de los coeficientes necesarios para el balanceo y verificar si es factible la aplicación como método de resolución.

Así como también la definición de una ecuación química suficientemente completa para la evaluación de la eficiencia y eficacia. Cabe mencionar que se desconoce si existe información en la literatura sobre la aplicación de los algoritmos genéticos para el balanceo de ecuaciones químicas.

CAPÍTULO II

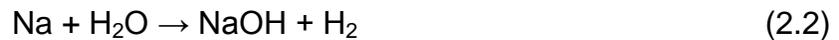
MARCO TEÓRICO

2.1.- Antecedentes

La importancia del estudio de las reacciones químicas se puede apreciar en el desarrollo industrial. Considerando lo expuesto en [Ang60], para que las industrias puedan surgir y vivir son necesarias las materias primas, que en pocos casos pueden ser empleadas sin ser modificadas substancialmente; estas modificaciones se pueden obtener solamente por medio de transformaciones químicas adecuadas, debido a que cualquier tratamiento que no produzca reacciones químicas sólo es capaz de modificar la forma o el aspecto exterior de los materiales sin influir sobre la constitución de las sustancias. De acuerdo a [Gut86], el término estequiometría se usa en Química para hacer referencia a la relación entre los elementos que constituyen las sustancias. Se dice que una sustancia es estequimétrica, cuando tiene una composición atómica definida, y se dice que una reacción química es estequiométrica, cuando existe una correspondencia definida entre la composición y cantidad de las sustancias reaccionantes, y las que se forman en la reacción.

De acuerdo a [BW77], una reacción química es un proceso por el que una o varias sustancias reaccionan químicamente entre sí para producir nuevas sustancias. Las sustancias de partida se denominan reactivos, las sustancias formadas se denominan productos (con propiedades diferentes a los reactivos). Un ejemplo de una reacción química es el siguiente: el sodio es un metal blando y brillante que reacciona vigorosamente con el agua, cuando se coloca una pequeña cantidad de

sodio en un recipiente de agua, se forma rápidamente gas hidrógeno e hidróxido de sodio en la solución, dicha reacción se describe en la ecuación 2.1 y se representa simbólicamente según la ecuación 2.2.



Una ecuación química no está completa si no respeta las ley de conservación de la materia; al proceso de encontrar los coeficientes estequiométricos de las sustancias involucradas en una reacción química se conoce como balanceo de ecuaciones químicas. Se puede entender que el problema es balancear una ecuación química y el algoritmo a desarrollar buscará los coeficientes estequiométricos que cumplan la ley de la conservación de la materia; donde los coeficientes serán el resultado al problema.

El balanceo de las ecuaciones químicas es un problema importante, dado que una reacción química, cuando es factible, es un proceso natural y la consecuente ecuación es siempre coherente. Si la reacción es inviable, entonces la única solución es trivial, donde todos los coeficientes son iguales a cero.

El balanceo de ecuaciones químicas es estudiado por la comunidad científica y se han desarrollado diferentes métodos para resolver dicho problema. De acuerdo a [Ris08] su estudio se ha desarrollado desde cuatro puntos de vistas, como se presenta en la tabla 2.1.

Tabla 2.1.- Enfoques de estudio del balanceo de ecuaciones químicas.

Enfoque	Descripción
Matemático	Se enfoca en modelar matemáticamente el problema utilizando sistemas de ecuaciones y otros artefactos matemáticos.
Informático	Existen muchos artículos publicados que consideran el uso de

Tabla 2.1.- Enfoques de estudio del balanceo de ecuaciones químicas.

Enfoque	Descripción
	computadoras para balancear ecuaciones químicas.
Químico	Los libros de texto universitarios de química general se apoyan generalmente en la técnica de ion-electrón, método algebraico, inspección, por mencionar algunos, como procedimiento básico para el balanceo de una ecuación química.
Pedagógico	Debido a que el balanceo de ecuaciones químicas es un tema considerado difícil en la enseñanza de la química, se han propuesto muchos enfoques para integrar al balanceo de ecuaciones químicas dentro del proceso de enseñanza-aprendizaje.

Para balancear ecuaciones químicas existen diversos métodos que invariablemente tienen el objetivo de que los coeficientes estequiométricos asignados a la ecuación química mediante el método cumplan con la ley de la conservación de la materia.

En los últimos años se han desarrollado aplicaciones de computadora para balancear ecuaciones químicas, utilizando diferentes enfoques. Kumar en [Kum01] realiza un estudio de las aplicaciones en computadora para el balanceo de ecuaciones químicas, donde se presentan métodos con diferentes enfoques, basándose en manipulaciones matriciales, en programas interactivos y en diseño de ingeniería.

En el año 2006, en [SAS06] se propone un enfoque de Programación Lineal Entera (PLE) para el balanceo de ecuaciones químicas haciendo uso de una matriz. En 2008, en [Ris08] se propone resolver el problema de balanceo de ecuaciones usando manipulaciones matriciales. Lo más importante de dicho trabajo es la formulación matemática, sin embargo, se limita a ecuaciones donde contienen el mismo número de elementos que de compuestos, ya que hace uso la matriz cuadrada. En [Ris09] Risteski plantea resolver el balanceo de ecuaciones químicas mediante un matriz singular.

En [Zou10] se propone resolver el balanceo de ecuaciones químicas mediante la utilización de una metaheurística. Se plantea un algoritmo basado en una búsqueda de armonía (HS) incluyendo operadores de actualización de posición y mutación genética.

De acuerdo a la investigación realizada en diferentes artículos, se puede determinar que Sen [SAS06] es el pionero en modelar el problema de balanceo de ecuaciones como un problema de Programación Lineal Entera (PLE). La formulación matemática contempla el balanceo de ecuaciones donde el número de elementos y de compuestos son diferentes, sin embargo el número de elementos es menos o igual al número de compuestos en las ecuaciones que se presentan en dicho artículo.

2.2.- Métodos de balanceo de ecuaciones químicas

Existen diferentes métodos para resolver el problema de balanceo de ecuaciones químicas. De acuerdo a la investigación realizada los métodos de balanceo se pueden clasificar como técnicas clásicas, técnicas basadas en matriz y técnicas basadas en heurísticas. En todos los métodos el objetivo que se persigue es que la ecuación química cumpla con la ley de la conservación de la materia.

2.2.1.- Técnicas clásicas

Los métodos de balanceo de ecuaciones químicas clásicos son aquellos que generalmente se enseñan en nivel medio superior y superior. Dentro de los más importantes se pueden citar el método de tanteo, método algebraico y método oxido-reducción. En la Tabla 2.2 se describen los métodos de balanceo químico más

conocidos.

Tabla 2.2.- Métodos de balanceo más conocidos.

Método	Autor	Descripción
Algebraico	J. Bottomley 1878	El método algebraico está basado en la aplicación del álgebra. Requiere construir un sistema de ecuaciones de varias variables y resolverlas simultáneamente. El número de pasos para balancear una ecuación por el método algebraico puede variar, dependiendo del grado de complejidad de tal ecuación.
Inspección/ tanteo	H. G. Deming 1943	Consiste en fijar arbitrariamente un coeficiente e ir deduciendo los demás haciendo balances a los átomos implicados en la elemento inicial. Si aparecen fracciones, se multiplican todos los coeficientes por el mínimo común múltiplo de los denominadores.
Oxido- Reducción	E. Jette y V.K. La Mer	Para entender este método hay que conocer el número de oxidación de un elemento. Se basa en los cambios de los números de oxidación de las elementos que reaccionan.

A continuación se describen detalladamente el procedimiento para realizar el balanceo de ecuaciones mediante el método algebraico, el método de óxido-reducción y el método de inspección o tanteo.

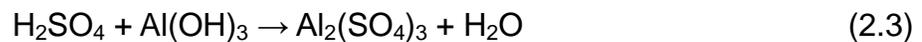
Método algebraico

El método algebraico requiere construir un sistema de ecuaciones de varias variables y resolverlas simultáneamente. El número de pasos para balancear una ecuación por el método algebraico puede variar, dependiendo del grado de complejidad de tal ecuación. Los pasos generales para aplicar el método algebraico, tomados de [McL02], se presentan a continuación:

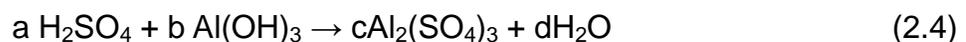
1. Escribir una reacción utilizando n diferentes símbolos (letras) para representar los coeficientes estequiométricos desconocidos.
2. Identificar los $n-1$ elementos involucrados en la reacción.

3. Para cada elemento, construir una ecuación lineal algebraica que iguale el número de átomos de el elemento en ambos lados de la reacción:
 - a) Para cada molécula en la ecuación química en la cual el elemento aparece, multiplique el subíndice del elemento por el coeficiente desconocido de la molécula.
 - b) Para cada elemento, sume el producto de los subíndices y los coeficientes de las moléculas de la izquierda y la derecha de la ecuación e iguale las sumas.
4. Seleccione una de las incógnitas y asigne a algún valor conveniente.
5. Resuelva las n ecuaciones con n incógnitas.
6. Si existen factores comunes en los coeficientes, divida cada coeficiente por ellos para reducir los coeficientes al mínimo conjunto de enteros. Si existe un coeficiente fraccionario, multiplique todos los coeficientes por el valor indicado de forma que todos sean enteros.
7. Verifique los coeficientes encontrados.

La aplicación del método algebraico se explica mediante la ecuación 2.3.



Primero se sustituyen los coeficientes estequiométricos de la ecuación por variables (hasta este momento con valor desconocido) siguiendo preferentemente el orden alfabético. La ecuación 2.3 queda de la siguiente forma 2.3.



La letras a, b, c y d son los coeficientes estequiométricos de la ecuación química balanceada. Luego se define una ecuación por cada uno de los elementos en la

reacción. Las ecuaciones se construyen multiplicando la variable por el número de átomos de ese elemento en específico en cada uno de los compuestos.

Para el hidrógeno la ecuación es $2a + 3b = 2d$. En H_2SO_4 hay 2 átomos de H por lo tanto, 'a' se multiplica por 2. En $Al(OH)_3$ hay 3 átomos de H por eso queda $3b$, en el $Al_2(SO_4)_3$ no hay H por lo tanto sería $0c$ y no se pone. En el caso de agua H_2O sería $2d$ porque tiene 2 átomos de H.

Usando este mismo razonamiento para los elementos restantes, el sistema de ecuaciones queda como en la figura 2.1:

H	$2a + 3b = 2d$
S	$a = 3c$
O	$4a + 3b = 12c + d$
Al	$b = 2c$

Figura 2.1.- Sistema de ecuaciones de la ecuación 2.3

El sistema queda con 4 ecuaciones con 4 incógnitas para resolver. Usualmente las ecuaciones químicas se balancean utilizando números enteros. Este método permite asignarle un valor (el que uno desee) a la letra que aparece en la mayoría de las ecuaciones algebraicas. Por lo tanto se asigna a c el valor de 1 (figura 2.2) y realizar la multiplicación (figura 2.3).

H	$2a + 3b = 2d$
S	$a = 3(1)$
O	$4a + 3b = 12(1) + d$
Al	$b = 2(1)$

Figura 2.2.- Asignación de 1 a c.

H	$2a + 3b = 2d$
S	$a = 3$
O	$4a + 3b = 12 + d$
Al	$b = 2$

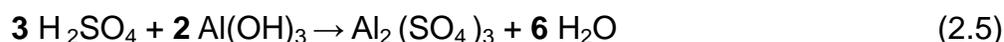
Figura 2.3.- Sistema de ecuaciones después de la sustitución.

Hasta este punto ya se conocen los valores de a, b y c, para encontrar el valor del coeficiente 'd' hay que despejar dicha incógnita (figura 2.4) en cualquiera de las ecuaciones de la figura 2.3 y hacer la sustitución de a y b.

$$d = \frac{2a + 3b}{2} = \frac{2(3) + 3(2)}{2} = \frac{6 + 6}{2} = 6$$

Figura 2.4.- Despeje y resultado de la incógnita d

Finalmente los valores obtenidos en el sistema de ecuaciones se asignan como coeficientes de las moléculas presentes en la ecuación química. Sustituyendo los valores de las variables en la ecuación química original queda de la forma 2.5.



Inspeccionando la ecuación química se observa tanto en los reactivos como en el producto de la ecuación que existen 12H, 3S, 18O y 2Al, por lo tanto la ecuación está balanceada.

Ventajas y Desventajas: El método algebraico para el balanceo de ecuaciones químicas es un método exacto. Sin embargo si la ecuación a balancear está conformada por más elementos que compuestos se obtiene un sistema de ecuaciones indeterminado [SGC95].

Método Óxido-Reducción

Como su nombre lo indica, este método de balanceo se basa en los cambios de los números de oxidación de los elementos que reaccionan. De acuerdo a [Ols97], existen conceptos base para aplicar este método, como se presenta en la tabla 2.3. El número de oxidación, de acuerdo a [Bol03], representa la carga eléctrica que asumiría un elemento si los electrones del enlace estuviesen distribuidos, de modo que se le atribuye el par de electrones del enlace al átomo más electronegativo.

Tabla 2.3.- Conceptos base para la oxido-reducción

Concepto	Descripción
Oxidación	Se refiere a la media reacción donde un átomo o un grupo de átomos <i>pierden e⁻</i> .
Reducción	Se refiere a la media reacción donde un átomo o un grupo de átomos <i>ganan e⁻</i> .
Agente Oxidante	Es la sustancia que se reduce (<i>gana e⁻</i>) provocando la oxidación.
Agente Reductor	Es la sustancia que se oxida (<i>pierde e⁻</i>) provocando la reducción.

Los pasos generales para aplicar el método de oxido-reducción son los siguientes:

1. Se calculan los números de oxidación.
2. Se identifican los elementos que cambian su estado de oxidación o carga y se escriben semi-reacciones con esos elementos.
3. Se efectúa el balance de masa en las semi-reacciones.
4. Se efectúa el balance de carga en las semi-reacciones.
5. Se balancean los electrones intercambiados (perdidos y ganados) en las semi-reacciones balanceadas.
6. Se introducen los coeficientes obtenidos a la reacción global.
7. Se ajustan los coeficientes de los elementos que no cambiaron.

Para explicar de la aplicación de este método, se utiliza la ecuación 2.6. El primer paso es calcular los números de oxidación como se muestra en la figura 2.5.

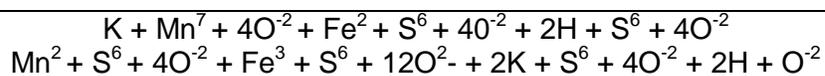
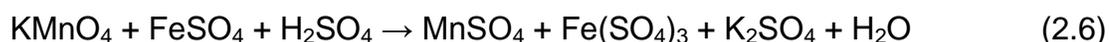


Figura 2.5.- Cálculo de número de oxidación.

Se identifican los elementos que cambian su estado de oxidación o carga. Se escriben como semi-reacciones de oxidación y de reducción (figura 2.6).

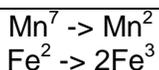


Figura 2.6.- Elementos con cambio de estado en el método de óxido-reducción.

Se efectúa el balance de masa, para igualar el número de elementos en ambos lados de la reacción (figura 2.7). En el caso del Mn, no es necesario efectuar el balance de masa pues hay un número igual de átomos en ambos miembros de la semireacción. Sin embargo, en el caso del Fe, hay un coeficiente de 2 en el Fe³⁺ que también debe aparecer del mismo modo en el Fe²⁺.

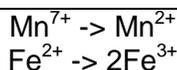


Figura 2.7.- Balanceo de masa de semireacciones.

Se efectúa el balance de carga, ya que debe haber igual número de cargas en ambos lados de las semi-reacciones. Lo único que puede utilizarse para el balance de carga son los electrones que se pierden o se ganan en el proceso. El balance de carga siempre debe hacerse después del balance de masa.

El planteamiento de una desigualdad matemática puede servir para realizar el balance de carga. Al mismo tiempo se pueden identificar los procesos de oxidación y de reducción, dependiendo del lado de donde se agreguen los electrones. La desigualdad se plantea utilizando los números de oxidación de los elementos que cambian en el proceso. En el caso del Mn el procedimiento se describe en la figura 2.8 y para el Fe el procedimiento se describe en la figura 2.9.

$$\begin{array}{l} 7^+ \geq 2^+ \\ 5e^- + 7^+ = 2^+ \\ 2^+ = 2^+ \\ 5e^- + \text{Mn}^{7+} = \text{Mn}^{2+} \end{array}$$

Figura 2.8.- Procedimiento de Mn.

$$\begin{array}{l} 2\text{Fe}^{2+} \leq 2\text{Fe}^{3+} \\ 4^+ \leq 6^+ \\ 4^+ = 6^+ + 2e^- \\ 4^+ = 4^+ \\ 2\text{Fe}^{2+} = 2\text{Fe}^{3+} + 2e^- \end{array}$$

Figura 2.9.- Procedimiento de Fe.

El número de oxidación del Mn disminuye de 7+ a 2+, por lo que la semi-reacción de reducción. El número de oxidación de Fe aumenta de 2+ a 3+ indicando que la semi-reacción de oxidación. Con lo anterior quedan balanceadas las semi-reacciones por masa y carga.

Se balancean los electrones intercambiados (perdidos y ganados) en las semi-reacciones balanceadas. El número de electrones que se intercambian en las semi-reacciones debe ser el mismo. Este se obtiene al multiplicar de manera cruzada los electrones perdidos y ganados. Al final se simplifica la ecuación, como se observa en la figura 2.10. El proceso total queda como la figura 2.11.

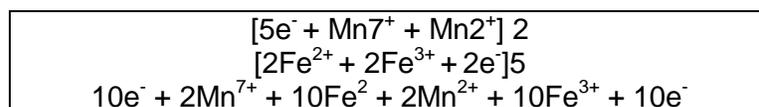


Figura 2.10.- Balanceo de electrones intercambiados

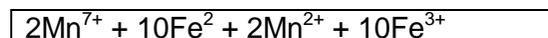


Figura 2.11.- Resultado de proceso redox.

Se introducen los coeficientes obtenidos por el proceso en la reacción original. Los coeficientes que se obtienen hasta este paso corresponden únicamente a las sustancias químicas que intervinieron en el proceso y se colocan como coeficientes de los compuestos correspondientes en la ecuación 2.6, como se muestra en la ecuación 2.7.



Se ajustan los coeficientes de los elementos que no cambiaron en el proceso. En esta reacción, no cambiaron su estado de oxidación el H+, S6+ K+ y O2- de modo que debe haber igual número de estas especies en ambos miembros de la ecuación para que ésta quede balanceada (ecuación 2.8).



Ventajas y Desventajas: Este método de balanceo es exacto, sin embargo se vuelve tedioso en su aplicación tanto al identificar que elemento se reduce o se oxida como en la utilización de las valencias. Efectivo para el balanceo de carga y masas.

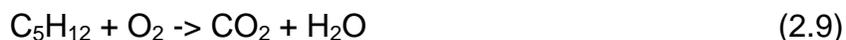
Método de Inspección o tanteo

Como se indica en [Cor02], el método más elemental empleado para balancear una reacción, es la simple inspección de la ecuación, y proceder a igualar el número de átomos de cada elemento, tanto en los reactivos como en los productos. Este método es especialmente útil para reacciones simples, pues presenta ecuaciones poco complicadas.

Para realizar el proceso de balanceo de ecuaciones se siguen una serie de pasos, generalmente se debe:

1. Escribir la ecuación sin balancear usando la fórmula química correcta para todos los reactivos y productos.
2. Usar coeficientes (números que se colocan al frente de cada fórmula química de reactivos y/o productos de acuerdo a la cantidad de átomos necesarios para balancear la ecuación.
3. Expresar los coeficientes con los números enteros más bajos posibles.
4. Verificar su contestación, determinando si la cantidad de átomos es igual en ambos lados de la flecha.

Para explicar de la aplicación de este método, se utiliza la ecuación 2.9.



Se identifican los átomos de cada elemento en los reactivos y productos de la ecuación presentada. Se comienza balanceando los átomos que se repiten una vez en ambos lados de la flecha. Existen cinco átomos de C en el lado de reactivos y un átomo C en los productos. Se prosigue a colocar el coeficiente 5 al frente de CO₂. La ecuación ha sido modificada a la siguiente manera (2.10).



Existen entonces 12 átomos de H en los reactivos y 2 átomos H en los productos. Se asigna el coeficiente 6 al frente de H₂O (2.11).



Ahora solo los átomos oxígenos los que no se encuentran balanceados. Existen 2 átomos de O en los reactivos y 16 átomos de O en los productos. El coeficiente correcto para terminar de balancear la ecuación es 8, que se asigna en los reactivos frente a O₂ (2.12).



Ventajas y Desventajas: Este método no es exacto, lo que no garantiza encontrar una solución. Es fácil de utilizar el método de tanteo cuando la ecuación es reducida y se complica su aplicación en ecuaciones donde se tiene gran número de compuestos y elementos, lo que hace que aumente la dificultad.

2.2.2.- Técnicas basadas en matriz

Dentro de la literatura se han propuestos diferentes métodos de solución basados en

la utilización de una matriz, cada método procesa de diferente manera la matriz.

Como se indica en [Cam95], la idea del método matricial es convertir una ecuación química en una matriz. Esta matriz representa el conjunto de ecuaciones lineales que, a su vez, representan tanto la conservación de masa y de carga de las sustancias. Cada columna de la matriz representa el número de átomos de un elemento que aparece en todas los compuestos químicos que intervienen en la ecuación química. La tabla 2.4 muestra un ejemplo. Los coeficientes correspondientes a los compuestos químicos en el lado derecho de la ecuación química de la ecuación son negativos. A continuación, el conjunto de ecuaciones lineales pueden ser resueltos mediante el uso de un método estándar. Hay algunas ventajas de utilizar el método de la matriz. Que sólo se basa en la idea fundamental de compuestos y en la conservación de la carga eléctrica y puede ser usado para establecer de forma inequívoca el número apropiado de las ecuaciones químicas necesarias. La principal ventaja del método de la matriz es que se puede traducir fácilmente en un programa informático para ordenadores personales.

Tabla 2.4.- Ejemplo de matriz en [Cam95]

	Zn	NO ₃ H	(NO ₃) ₂ Zn	H ₂ O	N ₂
Zn	1	0	-1	0	0
N	0	1	-2	0	-2
O	0	3	-6	-1	0
H	0	1	0	-2	0

Risteski en los últimos años ha enfocado sus investigaciones a resolver el balanceo de ecuaciones químicas basado en la resolución de matriz del tipo $n \times n$, aplicándola a ecuaciones químicas con elementos que poseen número de oxidación fraccionario. Risteski en [Ris08] propone un nuevo método generalizado de matriz inversa para el balanceo de ecuaciones químicas. El método se basa en la solución de una matriz homogénea de ecuaciones usando la matriz pseudoinversa de Von Neumann. Dichas ecuaciones químicas poseen números de oxidación fraccionarios. La formulación matemática se presenta en 2.13.

$$\sum_{j=1}^n x_j \prod_{i=1}^n \psi^i a_{ij} = 0 \quad (2.13)$$

Donde $x_j (1 \leq j \leq n)$ son los coeficientes a encontrar, $\psi^i (1 \leq i \leq n)$ son los elementos químicos y $a_{ij} (1 \leq i, j \leq n)$ representa la cantidad de átomos del elemento ψ^i . Por lo tanto cualquier ecuación química se puede representar con la ecuación 2.13. Este método de una manera formal de balanceo de ecuaciones químicas generales con análisis de matrices. Resultados obtenidos muestran que el método funciona para las ecuaciones química representada como una ecuación de matriz cuadrada.

De acuerdo con [Ris08] Este es un problema necesita un análisis más profundo científica para su resolución. Ese tipo de problemas busca un fundamento lógico de la química, algo similar como fundamentos de las matemáticas.

Sen en [SAS06] propone resolver el balanceo de ecuaciones químicas desde programación lineal entera. Dicho enfoque genera una matriz de átomos donde se contiene toda la información de la ecuación química a balancear. La matriz es del orden $m \times n$ resuelta mediante un procedimiento implementado en Matlab. Analizando el modelo propuesto por Sen se determina que la formulación matemática propuesta solo contempla que la ecuación esté balanceada en base a la cantidad de átomos de los reactivo y de los productos.

2.2.3.- Técnicas heurísticas

De acuerdo a la literatura hasta la fecha solo existe una heurística para resolver el problema del balanceo de ecuaciones químicas. Zou en [Zou10] propone el algoritmo de búsqueda de armonía global (NGHS) para balancear ecuaciones químicas. El algoritmo de búsqueda de armonía global es una mejora del algoritmo de búsqueda

de armonía incluyendo operadores de actualización de posición y mutación genética con baja probabilidad.

Dicho algoritmo se basa utiliza el modelo matemático propuesto por Sen en [SAS06]. NGHS se modifica las improvisaciones de HS y crea una nueva armonía que es mejor que la anterior. El procedimiento que realiza NGHS consiste en 6 pasos.

1. Inicialización del problema y parámetros del algoritmo.
2. Inicialización de la memoria de armonía.
3. Improvisación de una nueva armonía.
4. Reemplazo de la peor armonía con la armonía improvisada.
5. Verifica el criterio de paro.

El criterio de paro contempla el número máximo de iteraciones, si no se repite el procedimiento a partir del paso 3. NGHS es evaluado con los problemas que aparecen en [SAS06].

2.3.- Computación evolutiva

La computación evolutiva define un número de métodos diseñados para simular la evolución. Estos métodos son todos basados en poblaciones, y tienen una combinación de variación aleatoria y selección para resolver un problema. Diferentes enfoques existen en este campo, incluyendo los algoritmos evolutivos (AE), las estrategias evolutivas (EE), la programación evolutiva (PE), los algoritmos genéticos (AG) y la programación genética (PG). Aún cuando los griegos tenían cierto conocimiento rudimentario de la evolución, fue Darwin quien popularizó la teoría moderna de la evolución. El principio fundamental de la evolución es la optimización, donde la meta es la supervivencia de las especies. Los paradigmas de la

computación evolutiva ayudan a resolver problemas que han sido considerados como computacionalmente intratables.

Aún cuando se les critica por no tener una fundamentación matemática fuerte (por lo que se les ha colocado en lo que se conoce como computación suave o soft computing), su aplicación exitosa en la resolución de problemas prácticos de ciencia e ingeniería han hecho que estos enfoques llamen la atención de la comunidad científica.

2.3.1.- Motivación

Si se observa a los seres vivos, se ve que están bien adaptados en forma, función y estilo de vida. En muchos casos, las estructuras biológicas aún sobrepasan las capacidades de los más sofisticados mecanismos artificiales. De acuerdo a la teoría del desarrollo de las especies, todos los organismos que existen hoy en día son el resultado de un largo proceso: la evolución. Se puede establecer que el desarrollo de estructuras óptimas (o al menos muy buenas) es una propiedad de la evolución, esto es, la evolución posee una estrategia de optimización. En evolución, el mecanismo de variación es la ocurrencia de cambios aleatorios, a menudo “errores”, en la transferencia de información genética de una generación a otra. Los criterios de selección favorecen a los individuos mejor situados en lo que se llama la “lucha por sobrevivir”. Bremermann establece que la evolución debe considerarse como un proceso secuencial que explota la información de sucesos anteriores exitosos y fallidos, para seguir una trayectoria, en forma no determinística, dentro del espacio que representa sus parámetros [Fog00].

Los métodos de optimización natural representan procesos en la naturaleza que son notablemente exitosos en el fenómeno de la optimización natural. Las raíces

de la computación evolutiva emerge de algunos algoritmos inspirados en la evolución desarrollados en los años cincuenta. El propósito de estas investigaciones fueron el aprendizaje en máquinas y la optimización numérica. Se consideran como pioneros en esta área a los siguientes [DLJ00]:

1. A. M. Turing (1950) consideró la posibilidad de una búsqueda genética o evolutiva para propósitos de inteligencia en máquinas.
2. John von Neuman (1966) sugirió a un autómata auto-replicante como un modelo implícito de evolución en máquinas.
3. R. M. Friedberg (1958) consideró el uso de un proceso evolutivo para aprendizaje de máquinas y resolvedores de problemas. La idea era enseñar a la computadora para generar programas para resolver simples problemas.
4. G. Box (1957) desarrolló una técnica evolutiva para el diseño y análisis de experimentos industriales.
5. H. J. Bremermann (1962) desarrolló una teoría de algoritmos evolutivos, e intentó aplicar técnicas de evolución simulada para problemas de optimización numérica (optimización convexa) y para resolver ecuaciones no lineales.

La teoría evolutiva clásica de Darwin, junto con el principio de selección de Weismann y la genética de Mendel se conocen en su conjunto como el paradigma neo-darwiniano. El neo-darwinismo asegura que la historia de la mayoría de la vida está caracterizada por cuatro procesos estadísticos que actúan sobre las especies y sus poblaciones: reproducción, mutación, competencia y selección. La reproducción es una propiedad intrínseca de toda vida, afectada a través de la transferencia de un programa genético de un individuo a su descendencia; la mutación representa los errores en la transferencia, que pueden mejorar o afectar al organismo resultante; la competencia y la selección llegan a ser consecuencias de una población que se expande sobre un área restringida. La evolución es entonces el resultado de estos procesos estocásticos que interactúan sobre las poblaciones, generación tras

generación ([Fog00], [Berg01]). Ernest W. Mayr, uno de los más notables biólogos evolutivos del siglo XX, ha resumido algunos de las características más importantes del paradigma neo-darwiniano:

1. Los individuos son el objetivo de la selección.
2. Las variaciones genéticas son principalmente un fenómeno de cambio. Los procesos estocásticos juegan un papel significativo en la evolución.
3. La variación genotípica¹ es principalmente un producto de la recombinación.
4. La evolución gradual puede incorporar discontinuidades fenotípicas².
5. No todos los cambios fenotípicos son necesariamente consecuencias de selección natural “ad-hoc”.
6. La evolución es un cambio en la adaptación y la diversidad, no solo un cambio en la frecuencia de los genes.
7. La selección es probabilística, no determinística.

2.3.2.- Estructura de un algoritmo de computación evolutiva

El término algoritmo evolutivo se refiere a una familia de algoritmos que pueden ser descritos en términos de un esquema evolutivo general. La forma exacta de los operadores, así como la relación entre el tamaño de las poblaciones de padres e hijos, define al método en específico.

De acuerdo a [ACA00], la estructura de todos los algoritmos evolutivos es muy parecida (figura 2.12). El algoritmo mantiene una población de n individuos en la iteración t , $Pob(t)=\{x_1(t), x_2(t), \dots, x_n(t)\}$, donde cada $x_i(t)$ representa una posible

1 Es el conjunto de genes que contiene un organismo heredado de sus progenitores. En organismos como el ser humano, la mitad de los genes se heredan del padre y la otra mitad de la madre [AWT09].

2 Es la manifestación externa del genotipo, es decir, la suma de los caracteres observables en un individuo. El fenotipo es el resultado de la interacción entre el genotipo y el ambiente. El ambiente de un gen lo constituyen los otros genes, el citoplasma celular y el medio externo donde se desarrolla el individuo [AWT09].

solución al problema. Cada solución tiene asociada una aptitud $f(x_i)$; la aptitud de la población está representada por $F(t)=\{f(x_1(t)), f(x_2(t)), \dots, f(x_n(t))\}$. Esta población es afectada por las operaciones de recombinación y mutación, aplicadas utilizando las probabilidades P_r y P_m , para crear una nueva población $Pob''(t)$ de tamaño n .

```
Procedimiento algoritmo_evolutivo;  
inicio  
   $t = 0$ ;  
   $Pob(t) = Inicializar(n)$ ;  
   $F(t) = Evaluar(Pob(t), n)$ ;  
mientras(!condicion_de_paro())  
  inicio  
     $Pob'(t) = Recombinar(Pob(t), P_r)$ ;  
     $Pob''(t) = Mutar(Pob'(t), P_m)$ ;  
     $Pob(t+1) = Pob''(t)$ ;  
     $F(t+1) = Evaluar(Pob(t+1), n)$ ;  
     $t = t + 1$ ;  
  fin  
fin
```

Figura 2.12.- Estructura general de un algoritmo evolutivo.

Para recombinar individuos se necesita hacer una selección de los individuos de la población actual. La forma mas común de seleccionar individuos para construir nuevas poblaciones está basada en variantes de la selección por torneo. El propósito de la recombinación es tomar varios elementos de la población actual y producir varios descendientes combinando sus rasgos genotípicos. La mutación toma un elemento de la población y construye un descendiente perturbando el rastro genotípico de su padre.

Como se observa en la figura 2.12., los algoritmos evolutivos son procedimientos iterativos, cuyo número de iteraciones está determinada por una condición de paro (figura 2.13). En general, se aplican dos criterios para la condición de paro de un algoritmo evolutivo: alcanzar un número de iteraciones (generaciones) o cuando la diferencia de aptitudes entre los mejores individuos de cada población se menor a un rango preestablecido.

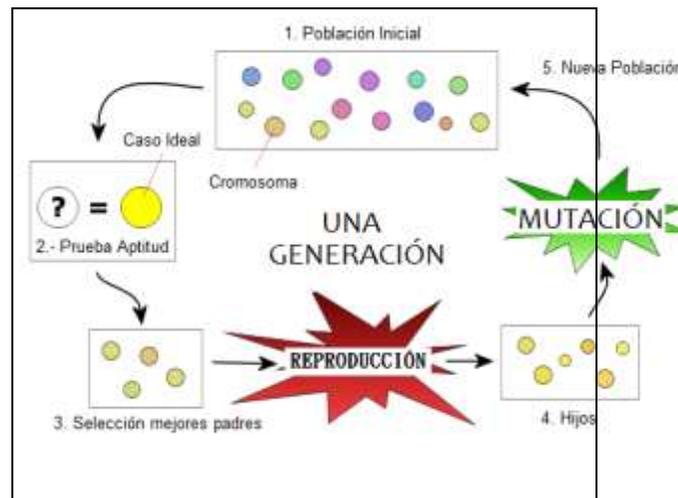


Figura 2.13.- Representación gráfica de un proceso evolutivo.

2.3.3.- Ventajas de los algoritmos de computación evolutiva

Los algoritmos evolutivos son una meta-heurística aplicable a un amplio rango de problemas. No están desarrollados sobre un dominio de problemas específicos y no están orientados hacia un área de aplicación en particular. Es ampliamente compartido que en general los algoritmos evolutivos son muy buenos para resolver problemas complejos, donde los modelos analíticos no son eficientes. En particular, un algoritmo evolutivo es mejor para usarse en comparación con otros enfoques cuando:

1. El problema tiene muchas variables y un espacio de búsqueda muy grande.
2. El problema tiene variables de diferentes tipos.
3. Existen interacciones no lineales entre las variables dentro de una función objetivo no lineal.
4. La función objetivo tiene muchos mínimos locales.
5. La función objetivo cambia con el transcurso del tiempo.

6. Existe ruido en la recuperación de la información.

En las siguientes secciones se describen los cuatro enfoques más importantes de la computación evolutiva (programación evolutiva, estrategias evolutivas, algoritmos genéticos, programación evolutiva).

2.3.4.- Programación Evolutiva

La programación evolutiva, introducida por Fogel en 1962, fue originalmente ideada como un intento de crear inteligencia artificial. El enfoque fue evolucionar máquinas de estados finitos (MEF)³ para predecir eventos en base a observaciones anteriores. El desempeño de un MEF con relación a su ambiente entonces puede ser medido con base en la capacidad de predicción de la máquina (comparando cada símbolo de salida con el siguiente símbolo de entrada y midiendo el valor de una predicción por alguna función de costo) [BHS97].

La técnica de programación evolutiva mantiene una población de MEFs, cada individuo representa una posible solución al problema (representa un comportamiento particular). Cada MEF tienen asociada una aptitud que es calculada como sigue: Cada MEF es expuesta al entorno en el sentido de examinar todos los símbolos previamente vistos. Para cada subsecuencia (a_1, a_2, \dots, a_i) se produce una salida a'_{i+1} , la cual es comparada con el siguiente símbolo observado a_{i+1} . Por ejemplo, para n símbolos, la MEF hace n predicciones; la función de aptitud toma en cuenta el desempeño global [MM97].

³ Una máquina de estados finitos es una máquina abstracta la cual transforma una secuencia de símbolos de entrada en una secuencia de símbolos de salida. La transformación depende de un conjunto finito de estados y un conjunto finito de reglas de transición [BHS97].

Esta técnica primero crea descendientes y después selecciona individuos para la siguiente generación. Cada padre produce un solo descendiente, por lo que el tamaño de la siguiente población se duplica. Los descendientes son creados por mutaciones aleatorias (mutaciones gaussianas) sobre los padres: cambiar un símbolo de salida, cambiar una transición de estado, adicionar un estado, eliminar un estado o cambiar el estado inicial (figura 2.14). Las mutaciones son aplicadas usando una distribución de probabilidad uniforme.

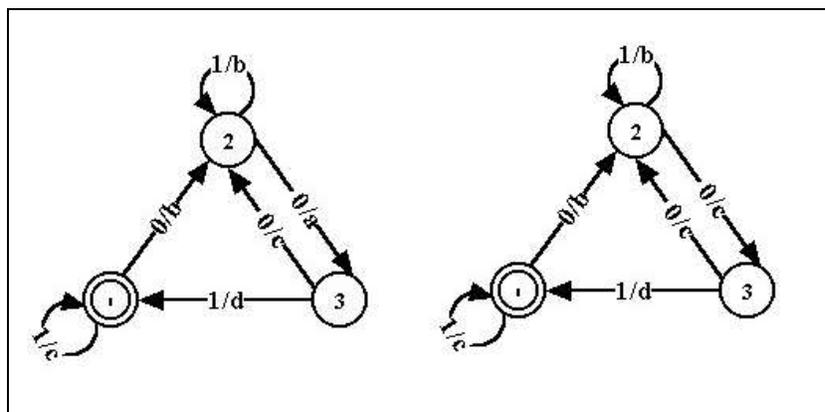


Figura 2.14.- Ejemplo de una mutación de una MEF en Programación Evolutiva.

2.3.5.- Estrategias Evolutivas

Las estrategias evolutivas, diseñadas por Rechenberg [Rec65] y Schwefel, generalmente se aplican a problemas de optimización con valores reales. Dada la imposibilidad de resolver problemas de hidrodinámica y de optimización de estructuras utilizando métodos tradicionales, se desarrolló un método simple basado en cambios aleatorios de valores experimentales. Estos cambios se consideran mutaciones muestreadas sobre una distribución normal. Los primeras aplicaciones numéricas de enfoque fueron realizadas por Hartmann y Höfler en 1976, y los primeros intentos por extender la estrategia para resolver problemas de optimización de parámetros binarios y discretos fue realizada por Schwefel.

Schwefel produjo en 1965 la primera estrategia evolutiva de dos miembros, (1+1)-ES, donde se construye un vector n -dimensional de valores reales basado en las variables del problema, que, aplicando mutaciones, sirve para crear un solo descendiente. La primera estrategia multi-miembro, llamada (λ +1)-ES, fue introducida por Rechenberg, donde varios padres se recombinan para crear un solo descendiente que sustituye al peor padre de la población, siempre y cuando tenga una mejor aptitud [BHS91].

Los puntos de búsqueda en este enfoque son vectores n -dimensionales que permiten definir la función de aptitud como una función similar dentro de un método tradicional. Los individuos contienen los valores de la posible solución, así como de información adicional llamada “parámetros de la estrategia”; esto es, utilizan un conjunto de vectores de valor flotante $v = (x, \sigma, \theta)$, donde x representa un punto en el espacio de búsqueda, σ incluye varias desviaciones estándar de una distribución normal y θ contiene a los ángulos de rotación representando la co-varianza de la distribución normal n -dimensional.

El operador de mutación es implementado manejando variaciones distribuidas normalmente basados en tamaños de paso adaptables (basadas en las varianzas y co-varianzas de la distribución normal). La mutación es el principal operador de variación, mientras la recombinación juega solo un papel secundario.

2.3.6.- Algoritmos Genéticos

Aún cuando se considera a Holland en 1969 como el creador de los algoritmos genéticos, Fraser en 1957, Bremerman en 1958 y Reed en 1967 proponen algoritmos similares que simulan sistemas genéticos [Fog98]. Esencialmente estos algoritmos incorporan una población de individuos codificados como cromosomas

(generalmente como una representación binaria) que es propagada basándose en su aptitud (determinada por una función asociada al problema); la propagación genera nuevos individuos utilizando operaciones de mutación y recombinación.

Estos procedimientos son conocidos con el término de algoritmos genéticos y se implementan como sigue:

8. El problema es definido por una función objetivo que indica la aptitud de cualquier posible solución.
9. Una población de soluciones candidatas es inicializada sujeta a ciertas restricciones. Cada solución es codificada como un vector x (cromosoma), con elementos descritos como genes, que pueden tener asignados valores (alelos). Holland sugiere que todas las soluciones sean representadas por cadenas binarias [Hol75].
10. Cada cromosoma de la población es decodificada para evaluarla y asignarse así un valor de su aptitud, de acuerdo al objetivo del problema.
11. Para cada cromosoma es asignada una probabilidad de reproducción, que es proporcional a su aptitud respecto a los demás cromosomas de la población. Si la aptitud es solo positiva y el problema es de maximización, entonces la selección se conoce como “de ruleta”.
12. Se genera una nueva población por selección probabilística, utilizando las probabilidades de reproducción. Los cromosomas seleccionados generan descendientes utilizando operadores genéticos (cruce y mutación).
13. El proceso se detiene cuando una buena solución es encontrada, o cuando el tiempo de ejecución ha sido agotado.

2.3.7.- Programación Genética

Este enfoque, desarrollado por Koza en 1989 [Koz89], tiene el objetivo de desarrollar en forma automática, programas de computadora que resuelvan problemas específicos. La programación genética es un enfoque independiente de dominio para la programación automática.

La programación genética es un paradigma de computación evolutiva en la cual una población de programas de computadora evolucionan para resolver problemas. En este enfoque, un programa de computadora es un individuo; los resultados que estos programas producen son evaluados, lo que representa el valor de aptitud del individuo. La meta de la programación genética es descubrir un programa de computadora que produzca una salida específica a partir de una entrada específica.

Cada individuo es representado con una estructura de árbol, lo cuales son de tamaño variable, las funciones definidas para el problema aparecen en el interior del árbol y las variables y constantes en las hojas del árbol. Los elementos necesarios para un programa genético son:

2. **Conjunto de terminales:** Comprende las variables y las constantes del problema.
3. **Conjunto de funciones:** Operadores aritméticos, lógicos, funciones matemáticas, funciones recursivas, iterativas, condicionales, etc. Cada función requiere de un número de parámetros (aridad)
4. **Medida de aptitud:** Es la forma de evaluar el desempeño del programa, que generalmente es inversamente proporcional al error producido por la salida del programa.
5. **Parámetros de control:** El tamaño de la población, el número máximo de generaciones, probabilidades de cruce y mutación.

Un algoritmo genético es una metaheurística del cómputo evolutivo que está basado en una analogía de la evolución de las especies.

2.4.1.- Antecedentes

La naturaleza genera seres óptimos, seres adaptados a su entorno, constituido por una infinidad de circunstancias: temperatura, presión atmosférica, precipitación, velocidad del viento, altitud, nivel de insolación, depredadores, alimentos, etc. Los individuos compiten al lado de otros individuos, donde los individuos más aptos logran sobrevivir. La selección natural obra solamente mediante la conservación y acumulación de pequeñas modificaciones heredadas.

Estas “modificaciones heredadas”, señaladas por Darwin como las generadoras de organismos mejores, son llamadas mutaciones hoy en día y constituyen el motor de la evolución. Un organismo mutante ha sufrido una modificación que lo hace diferente al resto de sus congéneres. Esta modificación puede ser un inconveniente para él, pero puede ocurrir también que le confiera alguna cualidad que le permita sobrevivir más fácilmente que al resto de los individuos de su especie. Este organismo tendrá mayor probabilidad de reproducirse y heredará a sus descendientes la característica que le dio ventaja. Con el tiempo, gracias a la competencia, los organismos que en un principio eran raros se volverán comunes a costa de la desaparición del “modelo anterior”. Se habrá dado entonces un paso en el proceso evolutivo.

Esta cualidad del proceso natural de la evolución (generar organismos óptimos sobre los que influyen infinidad de variables) llamó la atención de algunos científicos en las décadas de los cincuenta y sesenta. Un alemán, de apellido Rechenberg, introdujo en 1965 lo que denominó *evolutions strategie*, o estrategias evolutivas,

como un método para optimizar funciones de varias variables que definían dispositivos tales como perfiles de alas de avión.

En 1966 los investigadores Fogel, Owens y Walsh se dieron a la tarea de dejar evolucionar máquinas de estados finitos sometiendo sus diagramas de transición a cambios aleatorios (mutaciones), creando con ello lo que se conoce como *programación evolutiva*.

También en los 60's, John Holland, junto con algunos de sus colegas y alumnos de la Universidad de Michigan, desarrolló lo que se conoce como *algoritmos genéticos* (AG's). Sin embargo, el objetivo de Holland no era inventar un método de optimización basado en los mecanismos de la evolución natural, sino elaborar un estudio formal acerca del fenómeno de adaptación en sistemas naturales y artificiales, es decir, definir un modelo para la adaptación.

De acuerdo con [Mar10], los AG's son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización. Están basados en el proceso genético de los organismos vivos. Los AG's usan una analogía directa con el comportamiento natural. Trabajan con una población de individuos, cada uno de los cuales representa una solución factible a un problema dado. A cada individuo se le asigna un valor ó puntuación, relacionado con la bondad de dicha solución. En la naturaleza esto equivaldrá al grado de efectividad de un organismo para competir por unos determinados recursos. Cuanto mayor sea la adaptación de un individuo al problema, mayor será la probabilidad de que el mismo sea seleccionado para reproducirse, cruzando su material genético con otro individuo seleccionado de igual forma. Este cruce producirá nuevos individuos descendientes de los anteriores los cuales comparten algunas de las características de sus padres. Cuanto menor sea la adaptación de un individuo, menor será la probabilidad de que dicho individuo sea

seleccionado para la reproducción, y por tanto de que su material genético se propague en sucesivas generaciones.

De esta manera se produce una nueva población de posibles soluciones, la cual reemplaza a la anterior y verifica la interesante propiedad de que contiene una mayor proporción de buenas características en comparación con la población anterior. Así a lo largo de las generaciones las buenas características se propagan a través de la población. Favoreciendo el cruce de los individuos mejor adaptados, van siendo exploradas las áreas más prometedoras del espacio de búsqueda.

No se garantiza que el AG encuentre la solución óptima del problema, existe evidencia empírica de que se encuentran soluciones de un nivel aceptable, en un tiempo competitivo con el resto de algoritmos de optimización combinatoria. En el caso de que existan técnicas especializadas para resolver un determinado problema, lo más probable es que superen al AG, como es el caso del balanceo de ecuaciones que ya existen métodos que garantizan encontrar una solución óptima. El gran campo de aplicación de los AG's se relaciona con aquellos problemas para los cuales no existen técnicas especializadas. Incluso en el caso en que dichas técnicas existan, y funcionen bien, pueden efectuarse mejoras de las mismas hibridándolas con los AG's.

La figura 2.17 se presenta un esquema general de un algoritmo genético. Los algoritmos genéticos tradicionales utilizan una representación binaria de individuos en una cadena de longitud fija. Aunado a esta representación binaria, los algoritmos genéticos a menudo utilizan función de codificación y decodificación que permiten la relación de los dos espacios (soluciones y codificación binaria).

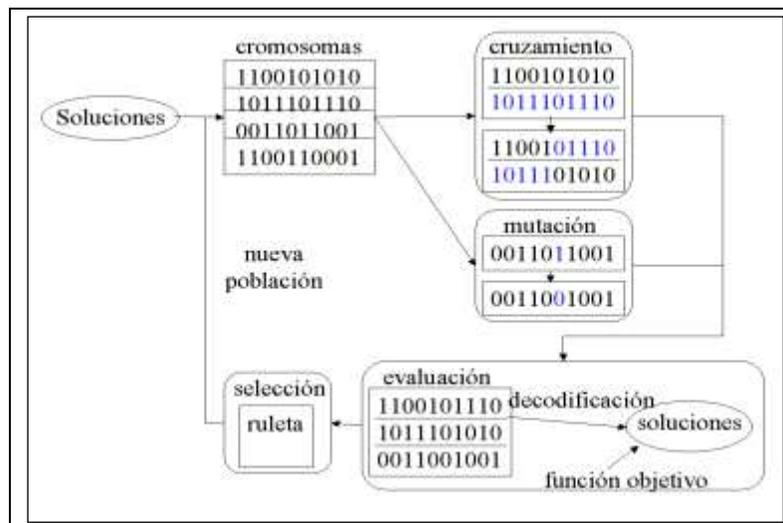


Figura 2.17.- Esquema general de un algoritmo genético.

2.4.1.- Codificación

De acuerdo [KG99] en la naturaleza las características de los seres vivos, incluso aquellas que los hacen óptimos para habitar en su medio, están determinadas por las proteínas que producen. A su vez, como hemos visto, estas proteínas (o más bien, los aminoácidos que las forman) se codifican en el material genético contenido en cada una de las células del individuo. Así pues, la naturaleza ha mapeado cada posible solución al problema de crear un individuo óptimo en una secuencia particular de bases que producirá ciertas proteínas, ha *codificado* el dominio del problema (todos los posibles individuos) mapeándolo al conjunto de todas las posibles secuencias de nucleótidos.

Así, para un algoritmo genético lo primero que se requiere es determinar en qué espacio se encuentran las posibles soluciones al problema que se pretende resolver. El algoritmo opera sobre "códigos genéticos", sobre genotipos que se deberán mapear al espacio de soluciones. Es decir, es necesario *codificar* de alguna manera el dominio del problema para obtener estructuras manejables que puedan ser manipuladas por el AG. Cada una de estas estructuras constituye el equivalente

al genotipo de un individuo en términos biológicos. Es frecuente que el código de los elementos del dominio del problema utilice un alfabeto binario (0's y 1's), también se puede realizar codificación entera y de punto flotante. La figura 2.18 muestra un individuo con codificación entera.

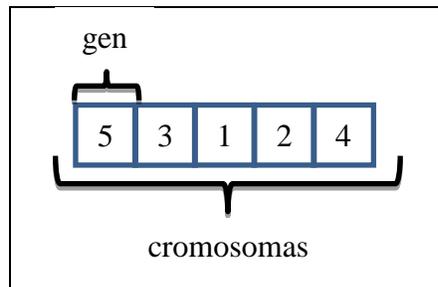


Figura 2.18.- Individuo genético entero

El proyecto se utiliza el sistema decimal con número enteros para la codificación del individuo. Cabe mencionar que los genes de los individuos serán los coeficientes estequiométricos de la ecuación a balancear.

2.4.3.- Evaluación de los individuos

Una vez que se ha definido la manera de codificar los elementos del dominio del problema y se conoce la forma de pasar de un elemento a su código y viceversa, es necesario fijar un punto de partida. Los AG's manipulan conjuntos de códigos en generaciones sucesivas. Nuevamente haciendo una analogía, manipulan poblaciones de códigos. En éstas un código puede aparecer más de una vez. El algoritmo se encargará de favorecer la aparición en la población de códigos que correspondan a elementos del dominio que estén próximos a resolver el problema. En resumen, el algoritmo recibirá como entrada una población de códigos (individuos) y a partir de ésta generará nuevas poblaciones, donde algunos individuos desaparecerán mientras que otros, que se mapean en mejores soluciones

posibles, aparecen con más frecuencia hasta que se encuentra una satisfactoria o hasta que se cumple alguna otra condición de terminación.

En una población hay un individuo más débil que otro, ya que existen diferencias entre los individuos. Esta diferencia es *relativa*, es decir, siempre está referida al resto de la población. Al igual que en la naturaleza, en los algoritmos genéticos es necesario establecer algún criterio que permita decidir cuáles de las soluciones propuestas en una población son mejores respecto del resto de las propuestas y cuáles no lo son. Es necesario establecer, para cada individuo, una medida de desempeño relativa a la población a la que pertenece.

Para determinar cuáles de estos individuos corresponden a buenas propuestas de solución y cuáles no, es necesario calificarlos de alguna manera. Cada individuo de cada generación de un algoritmo genético recibe una *calificación* o, para usar el término biológico, una medida de su *grado de adaptación (fitness)*. El objetivo de este número es que permita distinguir propuestas de solución buenas de aquéllas que no lo son. Si el problema a resolver consiste en maximizar una función, entonces la calificación asignada a un individuo determinado debe indicar qué tan alto es el valor de la función en el elemento de su dominio codificado por el individuo. Si, en cambio, el problema es determinar la ruta más corta entre dos puntos, la calificación deberá ser tanto más alta cuanto más corto sea el camino codificado en el individuo que esté siendo calificado. En el caso del proyecto la función tiende a minimizar ya que se trata de encontrar que los elementos presentes tanto en los reactivos como en los proyectos sean lo mismo y se determina encontrando que la diferencia de átomos entre ambos sea igual a cero.

Evidentemente, al hablar de que a cada individuo de la población se le asigna una y sólo una calificación, se está hablando de una función que se denomina *función de adaptación*, cuya evaluación puede no ser sencilla y es, de hecho, lo que

en la mayoría de los casos consume más tiempo en la ejecución de un AG. Hay que tener en cuenta que se evalúa una vez en cada individuo de cada generación. Si un AG es ejecutado con una población de tamaño 100 durante 100 generaciones, la función es evaluada 10,000 veces. Además, puede darse el caso de que la función de evaluación no tenga una regla de correspondencia explícita, esto es, una expresión algebraica, y puede ocurrir incluso que la función cambie de generación en generación.

2.4.5.- Selección

Una vez calificados todos los individuos de una generación, el algoritmo debe, al igual que lo hacen la naturaleza y el hombre, seleccionar a los individuos mejor adaptados al medio, para que tengan mayor oportunidad de reproducción. De esta forma se incrementa la probabilidad de tener individuos “buenos” (con alta calificación) en el futuro. Si de una determinada generación de individuos se seleccionaran sólo aquellos con una calificación mayor o igual que cierto número c para pasarlos a la siguiente generación, es claro que en ésta la calificación promedio superará c y por tanto al promedio de la generación anterior.

La selección ocasiona que haya más individuos buenos, explota el conocimiento que se ha obtenido hasta el momento, procurando elegir lo mejor que se haya encontrado, elevando así el nivel de adaptación de toda la población. En principio podría parecer que es conveniente tener una estrategia de selección estricta para que mejore rápidamente la población y converga el algoritmo, es decir, que la población se acumule alrededor de un genotipo óptimo. Esto no es cierto. Lo que ocurrirá es que la población se acumulará rápidamente alrededor de algún individuo que sea bueno, comparativamente con el resto de los individuos considerados a lo largo de la ejecución del algoritmo, pero este individuo puede no ser el mejor posible. A esto se le suele llamar *convergencia prematura*.

No se puede asegurar pero sí procurar que lo anterior no ocurra. Además de la explotación es necesario que exista exploración. El AG debe, no sólo seleccionar de entre lo mejor que ha encontrado, sino procurar encontrar mejores individuos. A esto se dedican los operadores que serán descritos a continuación, los que aseguran que en todo momento exista cierto grado de variedad en la población.

En la estrategia de selección normalmente se incluye un elemento extra que sirve de “ancla”. Si sólo se hace selección forzando que sea más probable elegir al mejor individuo de la población pero sin asegurarlo, es posible que este individuo se pierda y no forme parte de la siguiente generación. Dentro de la literatura se encuentran diferentes mecanismos de muestreo, por ejemplo; ruleta, torneo, ranking, etc. Para el proyecto a desarrollar se implementan dos tipos de selección para escoger la más efectiva para dicho problema.

2.4.6.- Cruzamiento

Durante el cruzamiento ocurre el proceso de producción de descendientes (gametos). El código genético de los padres de un individuo se mezcla para producir gametos cuyo contenido genético es híbrido, es decir, una mezcla. De esta manera es posible que un individuo herede a sus descendientes las características mezcladas de sus propios padres. Si estas características le confirieron a sus ancestros una alta aptitud de sobrevivencia, entonces este individuo será, con alta probabilidad, un individuo exitoso en su manada. El cruzamiento de los códigos genéticos de individuos exitosos favorece la aparición de nuevos individuos que hereden de sus ancestros características deseables.

En el contexto de los AG's reproducirse significa que, dados dos individuos seleccionados en función de su grado de adaptación, éstos pasen a formar parte de la siguiente generación o, al menos, mezclen sus códigos genéticos para generar

hijos que posean un código híbrido. Es decir, los códigos genéticos de los individuos se *crusan*. Existen muchos mecanismos de cruzamiento, pero todos tienen por objeto que el código de un individuo *A* y el de uno *B*, previamente seleccionados, se mezclen para formar nuevos individuos con la esperanza de que éstos hereden de sus progenitores las características deseables. El mecanismo de cruzamiento más común es el llamado *cruzamiento de un punto*.

En la figura 2.19 se muestra un ejemplo de cruzamiento de dos padres denominados P1 y P2. Para saber a partir de qué posición se realizará el cruce, se genera un número aleatorio en base a ese número, dividimos en dos partes cada individuo.

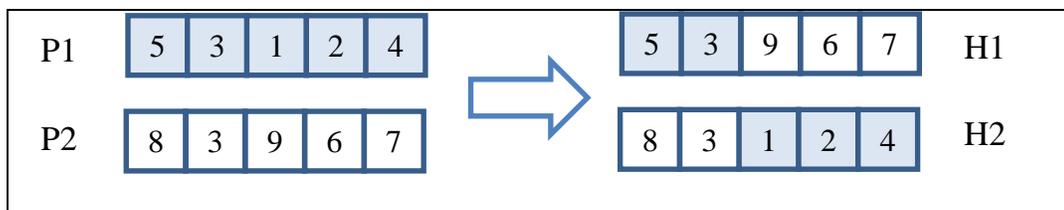


Figura 2.19.- Cruzamiento de P1 con P2.

2.4.7.- Mutación

Algunas veces cuando se copia la información genética a los descendientes, existe una alteración en este proceso y se produce una mutación, una alteración accidental en el código genético de los seres vivos. Ocasionalmente algunos elementos del código de ciertos individuos de un AG se alteran a propósito. Éstos se seleccionan aleatoriamente en lo que constituye el símil de una *mutación*. El objetivo es generar nuevos individuos, que exploren regiones del dominio del problema que probablemente no se han visitado aún.

Aleatoriamente se buscan nuevas soluciones posibles que quizá superen las encontradas hasta el momento. Esta es una de las características que hacen

aplicables los AG's a gran variedad de problemas: no presuponer conocimiento previo acerca del problema a resolver ni de su dominio, no sólo en la mutación sino en el proceso total.

Mutar consiste en alterar un gen en k posición del cromosoma de un individuo. No todos los individuos de la población mutaran, ya que hay que condicionar dicho proceso, por ello se define un grado de probabilidad, es necesario ya que respecto al grado de probabilidad se tomara la decisión de mutar o no a un individuo. En la figura 2.20 se muestra la mutación del individuo P1 (figura 2.19) donde el tercer gen cambia de valor de 1 a 7.

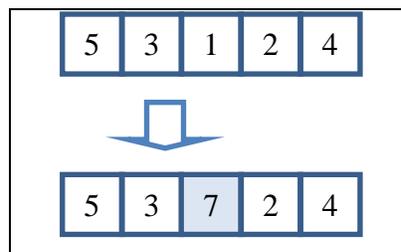


Figura 2.20.- Mutación del individuo

2.5.- Comentarios Finales

En este capítulo se describen los conceptos más importantes que se utilizan a lo largo del desarrollo de la tesis. Si bien existen varios métodos de balanceo de ecuaciones químicas que se ha aplicado tradicionalmente, se observa que en la actualidad es de interés científico el estudio de los modelos matemáticos para representar este problema.

En función a que el problema de balanceo de ecuaciones químicas se ha modelado como un problema de programación lineal entera, se describen en este capítulo las metaheurísticas emergidas del cómputo evolutivo que son una

herramienta muy poderosa para resolver problemas computacionalmente difíciles.

Si los algoritmos evolutivos son una meta-heurística aplicable diferentes problemas y los algoritmos genéticos pertenecen a los algoritmos evolutivos, es razonable aplicar un algoritmo genético para resolver el problema del balanceo de ecuaciones químicas. En esta investigación es de suma importancia definir correctamente la estructura del cromosoma del individuo.

CAPÍTULO III

DESARROLLO DEL MODELO

MATEMÁTICO PARA EL PROBLEMA DE

BALANCEO DE ECUACIONES QUÍMICAS Y

SU RESOLUCIÓN APLICANDO UN

ALGORITMO GENÉTICO

3.1.- Introducción

Es muy importante realizar el modelado matemático para la resolución de algún problema. El problema del balanceo de ecuaciones químicas se ha modelado en [Ris08] y [SAS06], ambos modelos generan una matriz de átomos, donde se tiene la relación de la cantidad del elemento x en el compuesto y .

El modelado que realiza Ristesky en [Ris08], se utiliza en varias publicaciones del tema, aplicando en cada una de ellas un diferente método de resolución, todos ellos basados en la manipulación matricial. La limitante con la que cuenta dicha formulación es que solo se puede aplicar a ecuaciones que tienen el mismo número de elementos que compuestos, caso contrario que la formulación propuesta por Sen en [SAS06]. Sin embargo, la formulación de Sen solo contempla el balanceo de la masa por compuesto y no por elementos. Es decir, que verifica que exista la misma cantidad de átomos en los reactivos como en los productos, dando la posibilidad a que no se encuentren balanceados los elementos.

3.2.- Formulación matemática

De acuerdo al tipo de problema la función objetivo tiene que minimizar los valores a

encontrar. Para una mejor explicación se describe paso a paso como se va obteniendo la función objetivo mediante la ecuación química de ejemplo (3.1)



Para que una ecuación esté balanceada tiene que cumplir con la ley de la conservación de las masas, por lo tanto, el número total de átomos de cada elemento químico de los reactantes debe ser igual al número total de átomos de dicho elemento en los productos. Para el problema del balanceo de ecuaciones químicas se buscan los coeficientes estequiométricos, por lo tanto, dichos coeficientes son incógnitas que se representaran con x_i . Anteponiendo x_i antes de cada compuesto de la ecuación (3.1) queda de la forma (3.2).



Colocando los subíndices de cada elemento en la ecuación (3.2) se forma la ecuación (3.3).



Se puede decir que los elementos químicos que no están en los compuestos tienen subíndices 0. Por ejemplo, en el primer compuesto el único elemento que no aparece es el C por lo tanto tiene subíndice 0. Sin alterar la ecuación, se agregan los elementos faltantes con subíndice 0 (3.4).



Se define que n representa el número de compuestos y m el número de elementos. Por lo tanto n es igual a 5 y m es igual a 4. Si se sustituye los nombres de los elementos por E, se tiene que $E^1=K$, $E^2=N$, $E^3=O$ y $E^4=C$ para la ecuación (3.4). La figura 3.1 muestra la sustitución de los elementos de la ecuación (3.4), separando cada compuesto que conforma la ecuación.

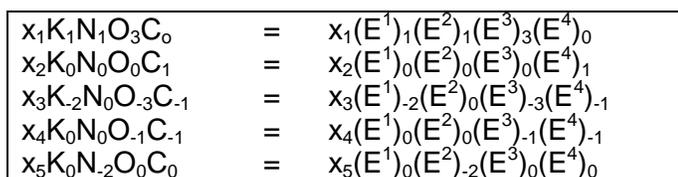


Figura 3.1.- Sustitución de símbolo químico por E

Aplicando la sustitución que se muestra en la figura 3.1 la ecuación química (3.4) se puede expresar como la ecuación (3.5)

$$x_1(E^1)_1(E^2)_1(E^3)_3(E^4)_0 + x_2(E^1)_0(E^2)_0(E^3)_0(E^4)_1 \rightarrow x_3(E^1)_{-2}(E^2)_0(E^3)_{-3}(E^4)_{-1} + x_4(E^1)_0(E^2)_0(E^3)_{-1}(E^4)_{-1} + x_5(E^1)_0(E^2)_{-2}(E^3)_0(E^4)_0 \quad (3.5)$$

Bajo esta representación, los subíndices de los elementos en los compuestos se puede representar como e_{ij} , para $i=1,\dots,n$ y $j=1,\dots,m$, y se colocan en una matriz. La tabla 3.1 muestra la distribución de la información en la matriz antes mencionada, se nota que los subíndices de los productos son negativos.

Tabla 3.1.- Matriz de subíndice

Compuestos		Coeficiente	Elementos = J			
			K	N	O	C
i	x_1	KNO_3	1	1	3	0
	x_2	C	0	0	0	1
	x_3	K_2CO_3	-2	0	-3	-1
	x_4	CO	0	0	-1	-1
	x_5	N_2	0	-2	0	0

En la figura 3.2 se muestra la sustitución los subíndices con la representación e_{ij} para los compuestos mostrados en la figura 3.1.

$x_1(E^1)_1(E^2)_1(E^3)_3(E^4)_0$	=	$x_1(E^1)e_{11}(E^2)e_{12}(E^3)e_{13}(E^4)e_{14}$
$x_2(E^1)_0(E^2)_0(E^3)_0(E^4)_1$	=	$x_2(E^1)e_{21}(E^2)e_{22}(E^3)e_{23}(E^4)e_{24}$
$x_3(E^1)_{-2}(E^2)_0(E^3)_{-3}(E^4)_{-1}$	=	$x_3(E^1)e_{31}(E^2)e_{32}(E^3)e_{33}(E^4)e_{34}$
$x_4(E^1)_0(E^2)_0(E^3)_{-1}(E^4)_{-1}$	=	$x_4(E^1)e_{41}(E^2)e_{42}(E^3)e_{43}(E^4)e_{44}$
$x_5(E^1)_0(E^2)_{-2}(E^3)_0(E^4)_0$	=	$x_5(E^1)e_{51}(E^2)e_{52}(E^3)e_{53}(E^4)e_{54}$

Figura 3.2.- Sustitución de símbolo químico por e_{ij}

Aplicando la sustitución que se muestra en la figura 3.2 la ecuación química (3.5) se puede expresar como la ecuación (3.6)

$$\begin{aligned}
 &x_1(E^1)e_{11}(E^2)e_{12}(E^3)e_{13}(E^4)e_{14} + x_2(E^1)e_{21}(E^2)e_{22}(E^3)e_{23}(E^4)e_{24} \rightarrow \quad (3.6) \\
 &x_3(E^1)e_{31}(E^2)e_{32}(E^3)e_{33}(E^4)e_{34} + x_4(E^1)e_{41}(E^2)e_{42}(E^3)e_{43}(E^4)e_{44} + \\
 &x_5(E^1)e_{51}(E^2)e_{52}(E^3)e_{53}(E^4)e_{54}
 \end{aligned}$$

Por lo tanto, el total de átomos de cada compuesto es $x_i(e_{i1} + e_{i2} + e_{i3} + e_{i4})$, para $i = 1, \dots, 5$. De acuerdo a lo anterior, el número total de átomos en la ecuación a balancear se rige por la fórmula 3.7.

$$\sum_{i=1}^n x_i \sum_{j=1}^m e_{ij} \quad (3.7)$$

No es suficiente que la sumatoria del total de los átomos de cada compuesto sea cero. Es necesario que el total de átomos de cada elemento en la ecuación química sea cero. Por ejemplo, K se encuentra presente en el compuesto 1 y compuesto 3, entonces la sumatoria de dichos átomos debe ser igual a cero. La figura 3.3 muestra la sumatoria para cada elemento.

$$\begin{array}{l}
 K (x_1 - 2x_3) = 0 \\
 N (x_1 - 2x_5) = 0 \\
 O (3x_1 - 3x_3 - x_4) = 0 \\
 C (x_2 - x_3 - x_4) = 0
 \end{array}$$

Figura 3.3.- Sumatoria de los elementos

Generalizando lo mostrado en la figura 3.3 queda de como se muestra en la figura 3.4.

$$\begin{array}{l}
 K (1x_1 + 0x_2 - 2x_3 - 0x_4 - 0x_5) = 0 \\
 N (1x_1 + 0x_2 - 0x_3 - 0x_4 - 2x_5) = 0 \\
 O (3x_1 + 0x_2 - 3x_3 - 1x_4 - 0x_5) = 0 \\
 C (0x_1 + 1x_2 - 1x_3 - 1x_4 - 0x_5) = 0
 \end{array}$$

Figura 3.4.- Generalización de la figura 3.3

Para que la ecuación química esté balanceada es necesario que todas estas sumas sean cero, lo que se puede garantizar sumando sus valores absolutos. Se observa en la matriz de subíndices (tabla 3.1) que estas sumas representan la suma del producto de cada coeficiente x_i por el correspondiente elemento e_{ij} de una columna j . Lo anterior descrito se muestra en la figura 3.5.

$$\begin{array}{l}
 K (e_{11}x_1 + e_{21}x_2 + e_{31}x_3 + e_{41}x_4 + e_{51}x_5) = 0 \\
 N (e_{12}x_1 + e_{22}x_2 + e_{32}x_3 + e_{42}x_4 + e_{52}x_5) = 0 \\
 O (e_{13}x_1 + e_{23}x_2 + e_{33}x_3 + e_{43}x_4 + e_{53}x_5) = 0 \\
 C (e_{14}x_1 + e_{24}x_2 + e_{34}x_3 + e_{44}x_4 + e_{54}x_5) = 0
 \end{array}$$

Figura 3.5.- Sustitución de símbolo químico por e_{ij}

Por lo que se podría representar el absoluto de cada suma como la ecuación 3.8

$$\left(\sum_{i=1}^n x_i e_{i1} \right) + \left(\sum_{i=1}^n x_i e_{i2} \right) + \left(\sum_{i=1}^n x_i e_{i3} \right) + \left(\sum_{i=1}^n x_i e_{i4} \right) \tag{3.8}$$

Simplificando la ecuación 3.8 queda como la ecuación 3.9

$$\sum_{j=1}^m \left(\sum_{i=1}^n x_i e_{ij} \right) \quad (3.9)$$

Entonces, la función a minimizar es (3.10)

$$f = \left(\sum_{i=1}^n x_i \sum_{j=1}^m e_{ij} \right) + \sum_{j=1}^m \left(\sum_{i=1}^n x_i e_{ij} \right) \quad (3.10)$$

Sin embargo el segundo término es condición suficiente para que la ecuación sea balanceada, por lo que la función a minimizar se reduciría a (3.11)

$$g = \sum_{j=1}^m \left(\sum_{i=1}^n x_i e_{ij} \right) \quad (3.11)$$

Para una mejor explicación del modelo matemático se probarán las dos funciones con los ejemplos ya estudiados, considere la ecuación (3.1). La ecuación 3.1 se puede formular como una matriz A del orden 5×4 como se muestra en la figura 3.6. La matriz se conforma de 5 filas y 4 columnas. La matriz contiene la cantidad de átomos de los elementos de cada compuesto que existe en la ecuación química. Las filas representan a los compuestos y las columnas a los elementos. Los elementos que no aparecen en un compuesto tienen cero átomos, por lo que en la matriz A se indican con un 0. Los átomos de los reactivos se identifican con el símbolo -.

$$\mathbf{A} = \begin{pmatrix} & \text{K} & \text{N} & \text{O} & \text{C} \\ \text{KNO}_3 & 1 & 1 & 3 & 0 \\ \text{C} & 0 & 0 & 0 & 1 \\ \text{K}_2\text{CO}_3 & -2 & 0 & -3 & -1 \\ \text{CO} & 0 & 0 & -1 & -1 \\ \text{N}_2 & 0 & -2 & 0 & 0 \end{pmatrix}$$

Figura 3.6.- Matriz A de la ecuación 3.1

De acuerdo al modelo matemático se desarrolla (3.10) queda de la forma (3.12).

$$f = |x_1(e_{11} + e_{12} + e_{13} + e_{14}) + x_2(e_{21} + e_{22} + e_{23} + e_{24}) + x_3(e_{31} + e_{32} + e_{32} + e_{34}) + x_4(e_{41} + e_{42} + e_{43} + e_{44}) + x_5(e_{51} + e_{52} + e_{53} + e_{54})| + g = 0 \quad (3.12)$$

$$g = |x_1e_{11} + x_2e_{21} + x_3e_{31} + x_4e_{41} + x_5e_{51}| + |x_1e_{12} + x_2e_{22} + x_3e_{32} + x_4e_{42} + x_5e_{52}| + |x_1e_{13} + x_2e_{23} + x_3e_{33} + x_4e_{43} + x_5e_{53}| + |x_1e_{14} + x_2e_{24} + x_3e_{34} + x_4e_{44} + x_5e_{54}| = 0 \quad (3.13)$$

Sustituyendo en 3.12 los valores de e_{ij} de acuerdo a la matriz A (figura 3.6), (3.12) queda de la forma (3.14)

$$f = |5x_1 + x_2 - 6x_3 - 2x_4 - 2x_5| + g \quad (3.14)$$

$$g = |x_1 - 2x_3| + |x_1 - 2x_5| + |3x_1 - 3x_3 - x_4| + |x_2 - x_3 - x_4| = 0 \quad (3.15)$$

Por lo tanto para evaluar las soluciones de la ecuación (3.1) se utiliza (3.14).

Ejemplo 1

Suponiendo que para la ecuación (3.1) se tiene $x = [2 \ 4 \ 1 \ 3 \ 2]$ se obtienen los valores mostrados en (3.17). Lo que demuestra que tanto f como g no son factibles.

$$g = |(1) - 2(1)| + |(2) - 2(2)| + |3(2) - 3(1) - (3)| + |(4) - (1) - (3)| = 2 \quad (3.16)$$

$$f = |5(2) + (4) - 6(1) - 2(3) - 2(2)| + g = 2 + 2 = 4 \quad (3.17)$$

Ejemplo 2

Suponiendo que para la ecuación (3.1) se tiene $x = [2 \ 4 \ 1 \ 3 \ 1]$ se obtienen los valores mostrados en (3.19). Lo que demuestra que tanto f como g son óptimas.

$$g = |(2) - 2(1)| + |(2) - 2(1)| + |3(2) - 3(1) - (3)| + |(4) - (1) - (3)| = 0 \quad (3.18)$$

$$f = |5(2) + (4) - 6(1) - 2(3) - 2(1)| + g = 0 + 0 = 0 \quad (3.19)$$

Ejemplo 3

Suponiendo que para la ecuación (3.1) se tiene $x = [1 \ 9 \ 1 \ 3 \ 1]$ se obtienen los valores mostrados en (3.20). Lo que demuestra que f no tiene impacto para determinar la factibilidad de la solución, demostrando que la función objetivo a minimizar es g (3.12).

$$g = |(1) - 2(1)| + |(1) - 2(1)| + |3(1) - 3(1) - (3)| + |(9) - (1) - (3)| = 10 \quad (3.20)$$

$$f = |5(1) + (9) - 6(1) - 2(3) - 2(1)| + g = 0 + 10 = 10 \quad (3.21)$$

Por lo tanto, en el algoritmo genético que utilizará la función g como función de aptitud.

3.3.- Descripción general del algoritmo

En este apartado se describe el algoritmo genético que se utilizó para balancear una ecuación química. Debido a que los métodos conocidos tienen sus limitantes en cuanto a las condiciones que deben cumplirse para que ellos sean aplicados (la relación entre el número de elementos y de compuestos de la ecuación química, por ejemplo) es prudente aplicar una metaheurística para encontrar los coeficientes estequiométricos de cualquier ecuación química. El uso de un algoritmo genético se soporta en el sentido que es una alternativa de búsqueda aleatoria “inteligente” en comparación con el método de inspección.

El algoritmo genético propuesto para el balanceo de ecuaciones químicas (AGBEQ) se muestra en la figura 3.7. La estructura de AGBEQ se basa en el algoritmo genético simple. El espacio de soluciones está formado por las diferentes

combinaciones de valores para los coeficientes estequiométricos, cada combinación representa a una posible solución del problema. Para el problema del balanceo de ecuaciones químicas la función objetivo es de minimización, por lo tanto, entre menor sea la adaptación del individuo tiene mayor probabilidad de sobrevivir.

0.	Codificación del problema
1.	Iniciar población actual aleatoriamente
2.	Evaluar población actual
3.	MIENTRAS no se alcance el número máximo de generaciones
4.	MIENTRAS población temporal no llena
5.	Seleccionar padres
6.	Cruzar padres con probabilidad P_c
7.	SI $f(\text{hijo})=0$
8.	Minimizar coeficientes
9.	Terminar algoritmo
10.	FIN SI
11.	SI se ha producido hubo cruzamiento
12.	Mutar descendientes con probabilidad P_m
13.	SI $f(\text{hijo})=0$
14.	Minimizar coeficientes
15.	Terminar algoritmo
16.	FIN SI
17.	FIN SI
18.	Añadir descendientes
19.	FIN MIENTRAS
20.	Actualizar población
21.	FIN MIENTRAS

Figura 3.7.- Algoritmo Genético para el balanceo de ecuaciones químicas

Se inicia el algoritmo desarrollando la codificación del problema. Posteriormente se genera la población inicial la cual es evaluada la función objetivo g (3.11). Se crean nuevas generaciones, aplicando los operadores genéticos (selección, cruzamiento y mutación). Cabe mencionar que después del cruzamiento y mutación se realiza la evaluación de los hijos, si uno de los hijos tiene aptitud de cero, quiere decir que ya se encontró la solución al problema. Por lo tanto se minimizan los coeficientes y se termina el algoritmo.

En caso de que el grado de adaptación del individuo evaluado sea mayor a cero se continúa ejecutando el algoritmo, aplicando los operadores genéticos a la población actual y generar la población temporal.

En cada generación, se crea una población temporal de nuevos individuos. La población temporal contiene los nuevos individuos generados mediante la aplicación de los operadores del AG. Por lo tanto, el número de iteraciones necesarias para llenar totalmente la población temporal dependerá de la cantidad de individuos que se creen en cada iteración de los operadores.

Cabe mencionar que es posible que no se encuentre una solución factible, entonces el criterio de paro es el número máximo de generaciones. Por lo tanto, el criterio de paro para la finalización del algoritmo contempla dos casos; alcanzar el número máximo de generaciones creadas y que un individuo de la población tenga aptitud igual con cero.

3.3.1.- Codificación

Un individuo es una posible solución al problema. Para el problema del balanceo de ecuaciones químicas la solución se denomina coeficientes estequiométricos. Por lo tanto, un individuo contendrá una combinación de valores para los coeficientes estequiométricos.

Ya que el número de coeficientes es igual al número de compuestos que integran una ecuación química. Para la ecuación 3.22 se tienen 5 coeficientes estequiométricos identificados como x_i .



El tamaño del individuo se determina de acuerdo al número de coeficientes de la ecuación química. Por lo tanto, para la ecuación 3.22 el cromosoma del individuo contiene 5 genes (figura 3.8), cada gen corresponde a un coeficiente de la ecuación 3.22. La figura 3.4 muestra la codificación directa de la ecuación 3.11, donde se incluye toda la información genética (La letra C indica *codificación* y la letra D indica *decodificación*). El primer renglón del cromosoma representa los elementos presentes en toda la ecuación en el orden en que se encuentran. El segundo renglón representa el subíndice del elemento, si el elemento pertenece a los reactivos el subíndice es positivo, en caso que sea producto el subíndice tiene valor negativo. El tercer renglón contiene el coeficiente correspondiente a cada compuesto.

$x_1\text{KNO}_3 + x_2\text{C} \rightarrow x_3\text{K}_2\text{CO}_3 + x_4\text{CO} + x_5\text{N}_2$									
C			↓	↑	D				
K	N	O	C	K	C	O	C	O	N
1	1	3	1	-2	-1	-3	-1	-1	-2
x_1			x_2	x_3			x_4		x_5

Figura 3.8.- Codificación directa de la ecuación 3.22.

Codificación Reducida

Analizando la estructura de las ecuaciones químicas, se observa que existe cierta dependencia entre los coeficientes de los reactivos con los productos. Dicha dependencia se determina para identificar los genes que no tienen dependencias, los cuales se modificaran durante el cruzamiento y mutación.

Para determinar las dependencias de los genes, se genera un sistema de ecuaciones de los elementos que aparecen una sola vez tanto en los reactivos como en los productos. La tabla 3.2 contiene aquellos elementos de la ecuación 3.22 que aparecen solo una vez tanto en los reactivos como en los productos y el número de átomos que existen en los reactivos y productos. Con dicha información se genera al última columna de la tabla que muestra la relación entre los coeficientes. El termino x_i indica el compuesto al q pertenece, para el K se encuentra presente en el compuesto x_1 y x_3 , en cambio el N se encuentra en el compuesto x_1 y x_5 .

Tabla 3.2.- Relaciones de los coeficientes de la ecuación 3.22

Elemento	# Átomos en Reactivos	# Átomos en Productos	Relación
K	1	2	$x_1 = 2x_3$
N	1	2	$x_1 = 2x_5$

Ya identificados los coeficientes con dependencias, se bloquean los genes del cromosomas del individuo que tienen dependencia para que no sean modificados durante el cruzamiento y mutación. Cabe mencionar que se bloquean los genes que forman parte de los productos y se determina su valor en función del valor del gen del reactivo con el que existe relación. La figura 3.9 muestra un el cromosoma de un individuo con genes bloqueados. Los genes bloqueados se encuentran sombreados.

K	N	O	C	K	C	O	C	O	N
1	1	3	1	-2	-1	-3	-1	-1	-2
X ₁		X ₂		X ₃			X ₄		X ₅

Figura 3.9.- Cromosoma de individuo con genes bloqueados

Los genes bloqueados al momento de realizar la evaluación del individuo, se determina su valor utilizando el sistema de ecuaciones. Donde se despeja x_i el cual pertenezca a los productos (figura 3.10).

$x_7 = 2x_3$ entonces $x_3 = x_7 \div 2$ $x_7 = 2x_5$ entonces $x_5 = x_7 \div 2$
--

Figura 3.10.- Despeje del sistema de ecuaciones.

Suponiendo que los valores de x_1 , x_2 y x_4 son 4,3 y 1 respectivamente, la estructura para evaluar el individuo quedará como se muestra en la figura 3.11. Como se muestra en la figura 3.10 el valor de x_3 y x_5 se determina con las ecuaciones de la figura 3.11.

K	N	O	C	K	C	O	C	O	N
1	1	3	1	-2	-1	-3	-1	-1	-2
4		3		2		1		2	

Figura 3.11.- Cromosoma completo.

Cabe mencionar que dicha representación depende del problema, entre más grande sea el problema mayor será el número de genes que contendrá el cromosoma del individuo. Sin embargo, se disminuye considerablemente el espacio de solución, cuando existen mayor cantidad de coeficientes dependientes. La descripción del algoritmo se realiza con la codificación reducida.

3.3.1.- Creación de la población inicial

Después de haber realizado la codificación del individuo, se genera la primera población de manera aleatoria. Los valores de los genes del cromosoma de cada

individuo de la población es generado aleatoriamente de acuerdo el rango establecido con las variables $limS$ y $limI$. Para el ejemplo $limS=10$ y $limI=1$.

Se utiliza un rango de posibles soluciones porque es necesario limitar el espacio de solución. Con dichos límites de realiza una cota al problema, los valores se fijan de acuerdo a la solución mostrada en la fuente de información de donde se tomó el ejemplo.

Una vez generada la población que se identifica como población actual, se realiza el cálculo de adaptabilidad de acuerdo a la función objetivo. La figura 3.12 muestra los genes de una población de 10 individuos con su aptitud después de ser evaluados. Para fines prácticos el tamaño de la población es de 10 individuos.

# Individuo	X ₁	X ₂	X ₃	X ₄	X ₅	Fitness
0	9	7	4	9	4	14
1	3	4	1	1	1	9
2	4	6	2	6	2	2
3	5	3	2	6	2	10
4	7	4	3	5	3	13
5	4	5	2	5	2	3
6	4	1	2	6	2	7
7	8	5	4	6	4	11
8	4	6	2	10	2	10
9	8	2	4	1	4	14

Figura 3.12.- Población inicial

Ya creada la población se evalúa la población actual (inicial). Se evalúa la población de acuerdo a la función g (3.11). El objetivo de evaluar la población es tener el grado de adaptación de cada individuo, que se utiliza para realizar la selección.

3.3.2.- Selección

La población del algoritmo genético se somete a un proceso de selección que debe tender a mejorar el grado de adaptación de los nuevos individuos en cada nueva generación. Se emplea la selección por torneo para el algoritmo genético propuesto. La selección por torneo consiste en realizar la selección en base a comparaciones directas entre individuos. Se selecciona aleatoriamente el $k\%$ de la población actual, de los cuales j ganan de acuerdo al criterio de selección, se lanza una moneda al aire para elegir si pasan los de mayor o menor grado de adaptación.

Para el algoritmo propuesto se elige el 40% del tamaño de la población para participar en la selección. Se utiliza un vector booleano para ubicar a los individuos que ya participaron en el torneo, así se evita elegir nuevamente al mismo individuo. La figura 3.13 muestra el vector booleano de los individuos ya elegidos aleatoriamente y la figura 3.14 el segmento de la población de los individuos elegidos de acuerdo al vector de la figura 3.13.

1	0	1	0	0	1	0	0	1	0
0	1	2	3	4	5	6	7	8	9

Figura 3.13.- Vector booleano de los individuos elegidos para la selección

# Individuo	x ₁	x ₂	x ₃	x ₄	x ₅	Fitness
0	9	7	4	9	4	14
2	4	6	2	6	2	2
5	4	5	2	5	2	3
8	4	6	2	10	2	10

Figura 3.14.- Individuos elegidos para la selección

De acuerdo a la figura 3.14 los individuos elegidos son el 0, 2, 5 y 8, dichos individuos se enfrentaran en parejas para elegir a los seleccionados para realizar cruce y mutación. Aleatoriamente se escogen al par de individuos a enfrentarse, y se determina si pasa el mejor o peor individuo de dicho par. Suponiendo que el primer par de números es 2,8 y que tiene que pasar el mejor individuo, se selecciona al individuo 2, como se muestra en la figura 3.15.

# Individuo	x ₁	x ₂	x ₃	x ₄	x ₅	Fitness
2	4	6	2	6	2	2
8	4	6	2	10	2	10

} 2 [4 6 2 6 2] 2

Figura 3.15.- Primer enfrentamiento.

Por lo tanto, el otro par de individuos son 0 y 5, donde se selecciona aleatoriamente el individuo 5 (figura 3.16). Es importante resaltar que no siempre pasa el mejor individuo, es por ello que se dice que la selección de torneo da posibilidad de que pase un individuo con bajo grado de adaptabilidad.

# Individuo	x ₁	x ₂	x ₃	x ₄	x ₅	Fitness
0	9	7	4	9	4	14
5	4	5	2	5	2	3

} 5 [4 5 2 5 2] 3

Figura 3.16.- Segundo enfrentamiento.

Por lo tanto, del porcentaje definido para realizar la selección pasa solo la mitad, ya que por cada enfrentamiento solo se selecciona un individuo.

3.3.3.- Cruzamiento

El cruzamiento consiste en intercambiar los genes entre los padres (individuos previamente seleccionados) para obtener nuevos individuos denominados hijos con

genes heredados por los padres. La figura 3.17 muestra los cromosomas de los padres seleccionados.

K	N	O	C	K	C	O	C	O	N
1	1	3	1	-2	-1	-3	-1	-1	-2
4		6		2		6		2	

Padre 1

K	N	O	C	K	C	O	C	O	N
1	1	3	1	-2	-1	-3	-1	-1	-2
4		5		2		5		2	

Padre 2

Figura 3.17.- Cromosoma de los individuos seleccionados como padres.

El cruzamiento tiene un valor de probabilidad para que se realice asignado a la variable de control p_C , para el ejemplo $p_C=80$ (equivale al 80%). Se determina aleatoriamente que par individuos se van a cruzar. Se utiliza un vector booleano para controlar que los individuos que ya participaron y no vuelvan a ser escogidos. Si el número aleatorio corresponde a un individuo que ya se cruzó se genera otro hasta encontrar un individuo que no haya sido escogido.

Primero se genera un valor aleatorio que oscila entre 0-1. Si el valor es menor a 80% o igual, se realiza el cruzamiento. La combinación de los genes se hace a partir de un punto que divide al cromosoma de los padres en dos partes, es el punto donde se combinan los genes para obtener a los descendientes. Cabe mencionar que existen genes que no son modificados por los operadores genéticos, para el ejemplo corresponden a x_3 y x_4 . La figura 3.18. muestra el cromosoma de los hijos heredados de los padres. El cruzamiento en dicho ejemplo es a partir de la segunda posición.

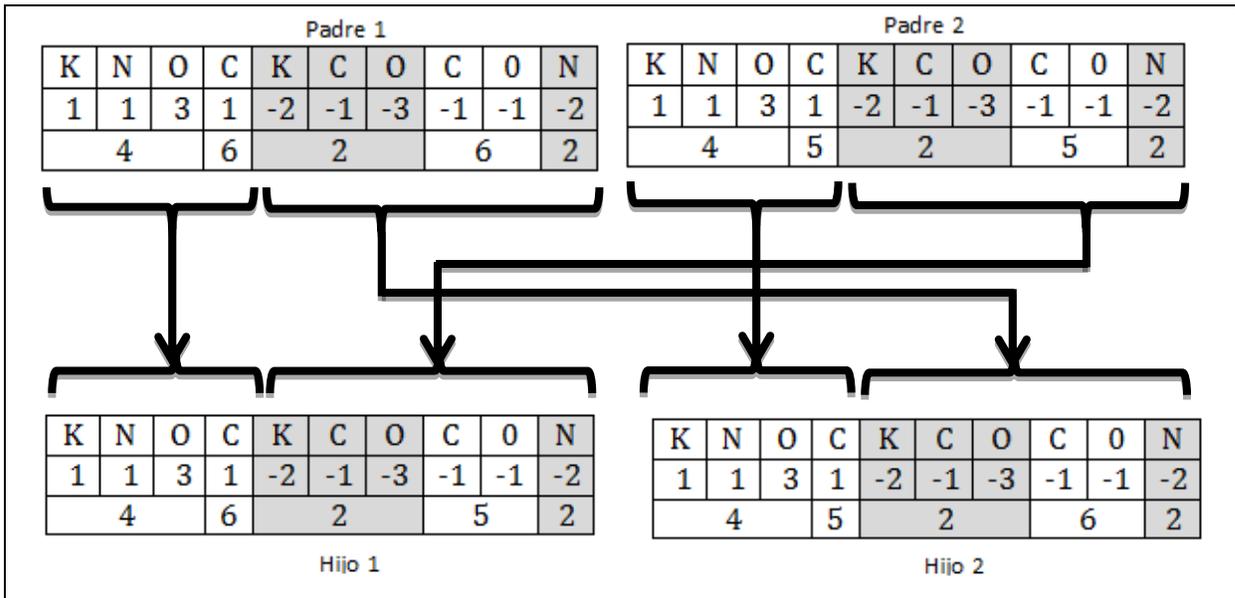


Figura 3.18.- Cromosoma de los hijos creados por el cruce.

En caso que el valor aleatorio sea mayor al 80% los hijos heredan el material genético idéntico a los padres, en otras palabras se copian los genes de padres a hijos, como se muestra en la figura 3.19.

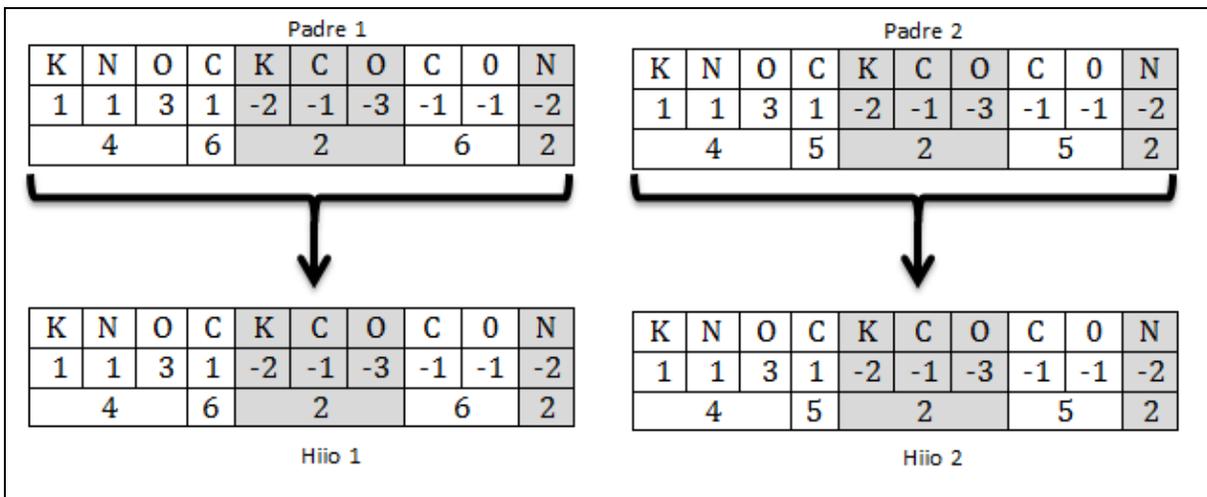


Figura 3.19.- Cromosoma de los hijos sin cruzamiento.

3.3.4.- Mutación

Si se realiza el cruzamiento se prosigue a mutar a los hijos. Sin embargo la mutación está sujeta al valor de pM (0.20 para este ejemplo). Cabe mencionar que cada hijo debe evaluarse para saber si va a ser mutado, es decir es probable que no mute ninguno, alguno(s) o todos.

En la mutación solo se altera el valor de un gen, la posición del gen a mutar se obtiene aleatoriamente (excepto los genes con dependencia) así como también el nuevo valor de dicho gen. La figura 3.20 se muestra la mutación del hijo 1 (figura 3.18) en diferentes genes.

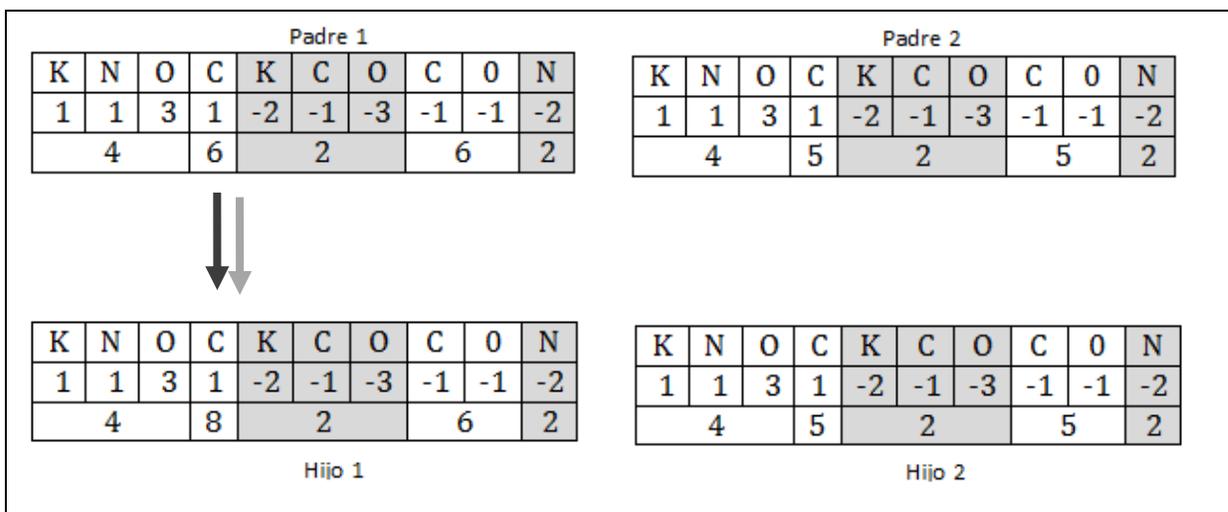


Figura 3.20.- Mutación del hijo 1

Una vez realizado los operadores se generan los hijos, estos individuos van a pertenecer a la población temporal. Después de mutar a un individuo se realiza la evaluación del grado de adaptación, si dicho individuo tiene adaptación igual a cero se termina de ejecutar el algoritmo, de lo contrario se prosigue con el procedimiento.

La población temporal se conforma de los hijos creados mediante los operadores genéticos aplicados a la población actual. Se aplican los operadores genéticos cuantas veces sea necesaria hasta tener completa la población temporal. Teniendo la población temporal completa sus individuos pasan a formar parte de la población actual (actualización de población), hasta este punto se tiene la primera generación. Por lo tanto la nueva población contiene individuos más adaptados.

3.3.5.- Evaluación

Los individuos que se generan al realizar el cruzamiento o la mutación se evalúan. Si el grado de adaptación del individuo es igual a cero significa que los valores de los genes del individuo balancean la ecuación, cumpliendo con la teoría de la materia¹. Luego se obtiene el valor mínimo común múltiplo para determinar el mínimo valor de los genes para que la ecuación esté balanceada, buscando minimizar solución factible óptima. Por lo tanto se termina de ejecutar el AG.

En los algoritmos genéticos simples, se realiza la evaluación de los nuevos individuos cuando se termina de crear una nueva generación. En el algoritmo propuesto considera evaluar a los individuos después de realizar el cruzamiento y la mutación. Cuando el grado de adaptación es igual a cero ya no es necesario seguir buscando soluciones. Por ello se plantea realizar la evaluación después de realizar cualquier cambio en los genes del individuo y no esperar a terminar de generar la nueva generación para realizar la evaluación de los individuos.

La aplicación de la evaluación después cualquier cambio en los genes del individuo, aumenta la eficiencia del algoritmo, ya que no se espera a terminar de

1 La materia no se crea ni se destruye solo se transforma.

generar la nueva población. La figura 3.21 muestra el cromosoma del hijo1 con grado de adaptación igual a cero. Se muestra en la figura 3.22 el mismo individuo después de haber minimizado los valores de los coeficientes. Se muestra en la figura 3.23 la decodificación de la solución del problema a la ecuación analizada.

K	N	O	C	K	C	O	C	O	N
1	1	3	1	-2	-1	-3	-1	-1	-2
4		8		2			6		2

Hijo 1

Figura 3.21.- Individuo con grado de adaptación igual con cero

K	N	O	C	K	C	O	C	O	N
1	1	3	1	-2	-1	-3	-1	-1	-2
2		4		1			3		1

Hijo 1

Figura 3.22.- Individuo con solución factible y óptima.

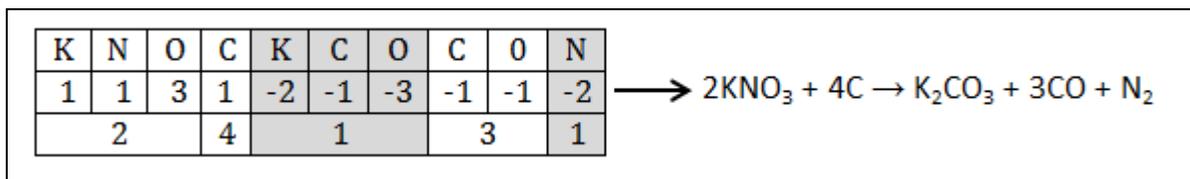


Figura 3.23.- Decodificación del cromosoma.

3.4- Comentarios finales

El algoritmo genético para el balanceo de ecuaciones químicas rompe el esquema de un algoritmo genético simple, ya que al realizar un cambio en los genes de los cromosomas del individuo se prosigue a evaluarlo. Otra aportación importante es el análisis realizado para determinar los genes dependientes de otros genes, mediante el sistema de ecuaciones que se determina.

El modelo matemático presentado es similar al propuesto en [SAS06]. Ambos modelos hacen uso de una matriz, la diferencia es el orden que ocupan los elementos y los compuestos en la matriz de átomos. Sin embargo, el modelo presentado en [SAS06] solo contempla que esté balanceada la ecuación en cantidad de reactivos y productos.

CAPÍTULO IV

IMPLEMENTACIÓN Y PRUEBAS

4.1.- Introducción

Las pruebas experimentales forman parte del ciclo de desarrollo de software dentro de la Ingeniería de software. Por lo que se realizan dichas pruebas mediante técnicas para descubrir que errores tiene.

En el ámbito de investigación la realización de pruebas se ocupa para evaluar la eficiencia y eficacia del algoritmo desarrollado, realizando varias ejecuciones del algoritmo para la misma instancias. El algoritmo se desarrolla en el paradigma de programación orientada a objetos, bajo la arquitectura MVC a través de la IDE Netbeans 6.9. Para llevar a cabo la experimentación del comportamiento del algoritmo genético para el balanceo de ecuaciones, se describe la plataforma de prueba. Las pruebas se realizaron en una computadora portátil marca Toshiba, modelo Satellite Pro U400 con procesador Intel Core2 Dou 2.13 GHz, memoria RAM de 3G, con sistema operativo Ubuntu 11.4 a 32 bits y compilador Java versión “jdk.1.6.0_21”

4.2.- Diseño e implementación de la lectura de la ecuación química desde un archivo de texto plano

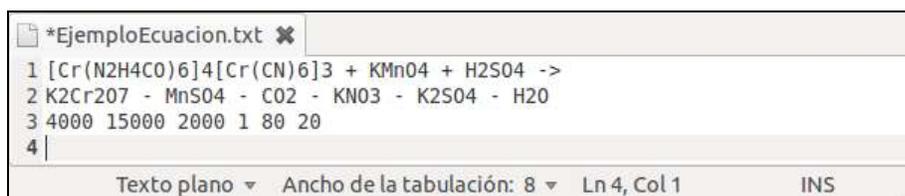
El diseño la lectura de la ecuación química desde un archivo de texto plano es para

generalizar el algoritmo a desarrollar para que resuelva diferentes problemas, cada problema con los datos particulares. El objetivo de ello es que el algoritmo desarrollado tome del archivo de entrada los valores de los datos requeridos para resolver el problema.

4.2.1 Diseño

Los datos que debe proporcionar cada problema para su resolución, corresponden a la ecuación química a balancear y las variables de control requeridas por el algoritmo genético. El formato del archivo es de texto plano (txt).

En la figura 4.1 se muestra el contenido del archivo *EjemploEcuacion.txt*. El archivo se estructura en 3 líneas, las primeras dos corresponden a la ecuación química [Pet98] y la última a las variables de control.



```
*EjemploEcuacion.txt
1 [Cr(N2H4CO)6]4[Cr(CN)6]3 + KMnO4 + H2SO4 ->
2 K2Cr2O7 - MnSO4 - CO2 - KNO3 - K2SO4 - H2O
3 4000 15000 2000 1 80 20
4 |
Texto plano Ancho de la tabulación: 8 Ln 4, Col 1 INS
```

Figura 4.1.- Estructura del archivo de un problema

Los reactivos se escriben en la primera línea como se muestra en la figura 4.2 separando los compuestos con el símbolo de suma '+' y entre los compuestos y el símbolo se pone un espacio en blanco. La línea termina con "->", para hacer referencia al resultado. El primer compuesto contiene dos tipos de símbolos de agrupación: paréntesis () y corchetes [], rigiéndose bajo las mismas reglas de matemáticas. Hasta el momento solo se reconoce los dos símbolos antes

mencionados, no se ha encontrado en la literatura alguna ecuación con otro símbolo de agrupación.

```
EjemploEcuacion.txt ✕
1 [Cr(N2H4CO)6]4[Cr(CN)6]3 + KMnO4 + H2SO4 ->
```

Figura 4.2.- Reactivos (EjemploEcuacion)

Los productos de la ecuación se escriben en la segunda línea (figura 4.3) los cuales se separan por el símbolo de menos '-'. Cabe resaltar que se deja un espacio en blanco entre los compuestos y el símbolo (+ ó -).

```
EjemploEcuacion.txt ✕
1 [Cr(N2H4CO)6]4[Cr(CN)6]3 + KMnO4 + H2SO4 ->
2 K2Cr2O7 - MnSO4 - CO2 - KN03 - K2SO4 - H2O
```

Figura 4.3.- Productos (EjemploEcuacion)

En la tercera línea de la figura 4.4 contiene los valores de las variables de control, utilizando como delimitador entre cada valor un espacio en blanco.

```
EjemploEcuacion.txt ✕
1 [Cr(N2H4CO)6]4[Cr(CN)6]3 + KMnO4 + H2SO4 ->
2 K2Cr2O7 - MnSO4 - CO2 - KN03 - K2SO4 - H2O
3 4000 15000 2000 1 80 20
```

Figura 4.4.- Variables de Control (EjemploEcuacion)

4.2.2 Implementación

Para balancear una ecuación química con cualquier método de balanceo se identifican el número de átomos de los elementos presentes en los compuestos, dicho paso se realiza en la lectura del archivo. En la figura 4.5 se muestra el procedimiento que se implementa para la lectura del archivo.

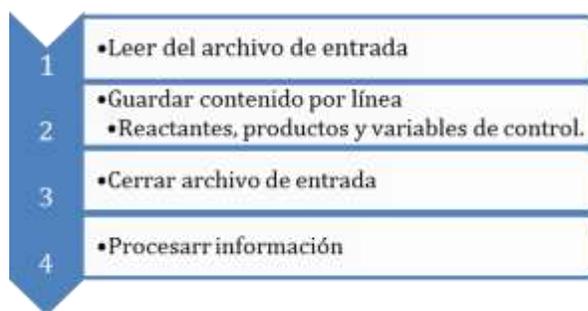


Figura 4.5.- Procedimiento para lectura de archivo de entrada.

En el cuarto paso del procedimiento primero se procesa la información de la ecuación y luego las variables de control. Referente a la ecuación se determina el número de compuestos, reactivos, productos, elementos y átomos del elemento 'x' en el compuesto 'y'. En otras palabras, se va transformando la información a su mínima expresión. Los valores de las variables de control solo se van asignando a los atributos correspondientes para utilizarse posteriormente.

Después de ser procesada la información se crea un archivo de salida que contiene el resultado del procesamiento del archivo de entrada. Dentro de todo el procesamiento se crea una matriz de átomos que contendrá el número de átomos del elemento 'x' en el compuesto 'y', dicha matriz corresponde a la función objetivo desarrollada en la formulación matemática, descrita en el 1er reporte parcial.

En la figura 4.6 [Pet98] y 4.7 [CFV07] se muestra el archivo de salida para dos problemas diferentes. Se presenta primero la ecuación a balancear, las variables de control, los reactivos (mostrando por filas los elementos en cada compuesto con la cantidad de átomos de dicho elementos), los productos y la matriz de átomos.

```
1 Ecuacion Quimica:
2 [Cr(N2H4CO)6]4[Cr(CN)6]3 + KMnO4 + H2SO4 ->
3 K2Cr2O7 + MnSO4 + CO2 + KNO3 + K2SO4 + H2O
4
5 Variables de Control:
6 Tamaño poblacion: 40000
7 Numero generaciones: 100
8 Limite superior: 2000
9 Limite inferior: 1
10 Probabilidad de cruzar: 80
11 Probabilidad de mutar: 20
12
13 :::Reactivos:::
14 Cr: 7   N: 66  H: 96  C: 42  O: 24
15 K: 1   Mn: 1  O: 4
16 H: 2   S: 1  O: 4
17
18 :::Productos:::
19 K: 2   Cr: 2  O: 7
20 Mn: 1  S: 1  O: 4
21 C: 1   O: 2
22 K: 1   N: 1  O: 3
23 K: 2   S: 1  O: 4
24 H: 2   O: 1
25
26 Matriz de subindices
27 Cr   N   H   C   O   K   Mn   S
28 7    66  96  42  24  0   0   0
29 0    0   0   0   4   1   1   0
30 0    0   2   0   4   0   0   1
31 -2   0   0   0   -7  -2  0   0
32 0    0   0   0   -4  0   -1  -1
33 0    0   0   -1  -2  0   0   0
34 0    -1  0   0   -3  -1  0   0
35 0    0   0   0   -4  -2  0   -1
36 0    0   -2  0   -1  0   0   0
```

Figura 4.6.- Archivo de salida

En la figura 4.7 se muestra una ecuación donde se ocupan signos de agrupación, por lo tanto, el programa realiza los cálculos necesarios para determinar el número de átomos de cada elemento. Por ejemplo, en el primer compuesto aparece dos veces el Nitrógeno. El N (en su primera aparición) tiene un subíndice de valor 2, dicho elemento se encuentra dentro de paréntesis y también de corchetes, se realiza

la multiplicación correspondiente teniendo 48 átomos y en la segunda aparición tiene 18 átomos en total son 66 átomos de N para el primer compuesto.

```
1 Ecuacion Quimica:
2 CH3CH2OH + Cr2O7 + H ->
3 CH3CO2H + Cr + H2O
4
5 Variables de Control:
6 Tamaño poblacion: 1600
7 Numero generaciones: 2000
8 Limite superior: 20
9 Limite inferior: 1
10 Probabilidad de cruzar: 80
11 Probabilidad de mutar: 20
12
13 :::Reactivos:::
14 C: 2 H: 6 O: 1
15 Cr: 2 O: 7
16 H: 1
17
18 :::Productos:::
19 C: 2 H: 4 O: 2
20 Cr: 1
21 H: 2 O: 1
22
23 Matriz de subindices
24 C H O Cr
25 2 6 1 0
26 0 0 7 2
27 0 1 0 0
28 -2 -4 -2 0
29 0 0 0 -1
30 0 -2 -1 0
```

Figura 4.7.- Archivo de salida 2

4.3.- Variables de control

Las variables de control son aquellas variables que al cambiar su valor afecta en el comportamiento del algoritmo. Las variables de control propias de un algoritmo genético son el porcentaje de tamaño de la población, número máximo de generaciones, cruzamiento y de mutación. Para el problema del balanceo de ecuaciones químicas se establece como variables de control las utilizadas para

definir el rango en que oscila los valores de los coeficientes estequiométricos. En base a lo anterior, en la tabla 4.1, se definen las variables de control del problema

Tabla 4.1.- Variables de control

Variable de Control	Descripción
tamañoPoblacion	Define el tamaño de la población.
maxGeneraciones:	Número máximo de generaciones a crear.
limS y limI:	Límites entre los cuales oscila el valor de los genes.
pC	Porcentaje de probabilidad para realizar el cruce.
pM	Porcentaje de probabilidad para hacer mutación.

Las pruebas experimentales consisten en ejecutar 30 veces el algoritmo para la misma instancia. De cada prueba se contabiliza el número de ocasiones que el algoritmo ha encontrado la solución factible óptima, también el menor y mayor tiempo requerido en la ejecución del algoritmo, así como el tiempo promedio para cada ejecución.

4.5.- Selección de problemas

De la literatura consultada, se escogen algunos problemas (ecuaciones), considerando que cada problema tenga condiciones diferentes, como el número de elementos y el número de compuestos químicos en la ecuación, así como el rango en el que oscilan los valores de los coeficientes de cada compuesto. Cada problema tiene diferentes características, con el objetivo de probar el algoritmo genético para el balanceo de ecuaciones en diferentes casos y observar su comportamiento.

La tabla 4.2 contiene los problemas encontrados dentro de la literatura resueltos por métodos exactos y por heurísticas. El método algebraico se utiliza en el problema 1 [Osl97], el método oxido-reducción en el problema 2 [CFV07],

mediante el método redox los problema 3-5 [Pet98] y por matriz los problemas 6-7 [Cam95] y problema 8 [Ris09]. #P se ocupara para identificar el número de problema a resolver. Los problemas 9-21 fueron resueltos con NGHS propuesto en [Zou10]. El artículo [Zou10] menciona que los problemas son fáciles de resolver excepto 9, 12,13. Así como también que los problemas 19 y 21 tienen solución con coeficientes 0's. Todos los problemas se ejecutaran con el algoritmo genético desarrollado.

Tabla 4.2.- Problemas resueltos por métodos exactos.

#P	ECUACIÓN
1	$\text{ClO}_3 + \text{As}_2\text{S}_3 + \text{H}_2\text{O} \rightarrow \text{H}_2\text{AsO}_4 + \text{SO}_4 + \text{Cl} + \text{H}$
2	$\text{CH}_3\text{CH}_2\text{OH} + \text{Cr}_2\text{O}_7 + \text{H} \rightarrow \text{CH}_3\text{CO}_2\text{H} + \text{Cr} + \text{H}_2\text{O}$
3	$\text{CuSCN} + \text{KIO}_3 + \text{HCl} \rightarrow \text{CuSO}_4 + \text{KCl} + \text{HCN} + \text{ICl} + \text{H}_2\text{O}$
4	$\text{Ag}_3\text{AsO}_4 + \text{Zn} + \text{H}_2\text{SO}_4 \rightarrow \text{AsH}_3 + \text{Ag} + \text{ZnSO}_4 + \text{H}_2\text{O}$
5	$\text{PbN}_6 + \text{CrMn}_2\text{O}_8 \rightarrow \text{Pb}_3\text{O}_4 + \text{NO} + \text{Cr}_2\text{O}_3 + \text{MnO}_2$
6	$\text{K}_4\text{Fe}(\text{CN})_6 + \text{SO}_4\text{H}_2 + \text{H}_2\text{O} \rightarrow \text{SO}_4\text{K}_2 + \text{SO}_4\text{Fe} + (\text{NH}_4)_2\text{SO}_4 + \text{CO}$
7	$\text{Cu}(\text{NH}_3)_4\text{Cl}_2 + \text{KCN} + \text{H}_2\text{O} \rightarrow \text{NH}_3 + \text{NH}_4\text{Cl} + \text{K}_2\text{Cu}(\text{CN})_3 + \text{KCNO} + \text{ClK}$
8	$\text{KNO}_3 + \text{C} + \text{S} \rightarrow \text{K}_2\text{CO}_3 + \text{K}_2\text{SO}_4 + \text{K}_2\text{S}_2 + \text{CO}_2 + \text{CO} + \text{N}_2$
9	$\text{H}_3\text{PO}_4 + \text{MgSiO}_3 + \text{CF}_2\text{Cl}_2 + \text{NaAlF}_4 + \text{KI} + \text{PbCrO}_4 + \text{FeSO}_4 + \text{BrCl} + \text{Ca}(\text{CN})_2 + \text{SO}_2 + \text{H}_2 \rightarrow \text{PI}_3 + \text{MgCO}_3 + \text{Na}_2\text{SiO}_3 + \text{PbBr}_2 + \text{CrCl}_3 + \text{KAl}(\text{OH})_4 + \text{Fe}(\text{SCN})_3 + \text{CaF}_2 + \text{H}_2\text{O}$
10	$\text{As}_2\text{S}_3 + \text{H}_2\text{O} + \text{HNO}_3 \rightarrow \text{NO} + \text{H}_3\text{AsO}_4 + \text{H}_2\text{SO}_4$
11	$\text{KClO}_3 + \text{HCl} \rightarrow \text{KCl} + \text{Cl}_2 + \text{H}_2\text{O}$
12	$\text{C}_2\text{H}_5\text{NO}_2 + \text{C}_3\text{H}_7\text{NO}_3 + \text{C}_6\text{H}_{14}\text{N}_4\text{O}_2 + \text{C}_5\text{H}_9\text{NO}_2 + \text{C}_9\text{H}_{11}\text{NO}_2 \rightarrow \text{H}_2\text{O} + \text{C}_{50}\text{H}_{73}\text{N}_{15}\text{O}_{11}$
13	$\text{C}_5\text{H}_9\text{NO}_2 + \text{C}_5\text{H}_{11}\text{NO}_2 + \text{C}_6\text{H}_{13}\text{NO}_2 + \text{C}_9\text{H}_{11}\text{NO}_2 + \text{C}_5\text{H}_{12}\text{N}_2\text{O}_2 \rightarrow \text{H}_2\text{O} + \text{C}_{60}\text{H}_{92}\text{N}_{12}\text{O}_{10}$
14	$\text{KNO}_3 + \text{C}_2 \rightarrow \text{K}_2\text{CO}_3 + \text{CO} + \text{N}_2$
15	$\text{NH}_3 + \text{O}_2 \rightarrow \text{NO} + \text{H}_2\text{O}$
16	$\text{C}_3\text{H}_6 + \text{O}_2 \rightarrow \text{CO}_2 + \text{H}_2\text{O}$
17	$\text{CaCO}_3 + \text{C} \rightarrow \text{CaC}_2 + \text{CO}_2$
18	$\text{FeCrO}_4 + \text{C} \rightarrow \text{Cr} + \text{Fe} + \text{CO}$
19	$\text{NO}_2 + \text{HClO} \rightarrow \text{HNO}_3 + \text{HCl}$
20	$\text{KClO}_3 + \text{HCl} \rightarrow \text{KCl} + \text{ClO}_2 + \text{Cl}_2 + \text{H}_2\text{O}$
21	$(\text{NH}_4)_2\text{SO}_4 \rightarrow \text{NH}_4\text{OH} + \text{SO}_2$

En la tabla 4.3 presenta las características de cada problema utilizado en la prueba. Se tiene que **#P** es el número que identifica al problema de acuerdo a la tabla 4.2. **#E** es el número de elementos diferentes y **#C** indica el número de compuestos que presentes en el problema. Por otro lado, **LimS** y **LimI** definen el rango en el que se encuentran los valores de los coeficientes para la solución de cada problema. Todos los valores son tomados de [Osl97],[CFV07],[Pet98],[Cam95],[Ris09] y [Zou10] . Para el problema 19 y 21 se tienen **LimS** y **LimI** con valor de cero ya que no se pudieron resolver con NGHS [Zou10] ni con Linsolver [SAS06].

Tabla 4.3.- Características de los problemas

#P	#E	# C	LimS	LimI
1	5	7	25	3
2	4	6	16	2
3	9	6	14	4
4	6	7	11	2
5	5	6	90	5
6	7	7	6	1
7	7	8	7	1
8	5	9	19768	1850
9	19	20	88	2
10	5	6	28	3
11	3	5	6	1
12	4	7	8	1
13	4	7	10	1
14	4	5	4	1
15	3	4	6	4
16	3	4	9	2
17	4	4	5	2
18	4	5	4	1
19	4	4	0	0
20	4	6	4	2
21	4	3	0	0

4.6.- Convergencia del algoritmo implementando Codificación Directa y Codificación Reducida.

Durante el desarrollo del presente trabajo, se realizaron diferentes versiones del algoritmo genético. La primera versión del algoritmo genético para el balanceo de ecuaciones químicas realiza la Codificación Directa (CD) del individuo como se muestra en la figura 3.8 y la segunda versión implementa la Codificación Reducida (CR) que se muestra en la figura 3.9.

Al realizar algunas pruebas experimentales y al comparar los resultados obtenidos se determinó trabajar con la CR debido a que tardaba menos tiempo en encontrar una solución. Las figuras 4.8 y 4.9 muestran la convergencia del algoritmo para los problemas 1 y 14 respectivamente, el eje x contempla el número de generaciones y el eje y corresponde el valor fitness del mejor individuo en la generación n . Los valores de las variables de control con las cuales se realizaron las pruebas se muestran en la tabla 4.4.

Tabla 4.4.- Valores de variables de control para pruebas de convergencia.

Problema	VARIABLES DE CONTROL					
	tamPob	maxGen	limS	limI	pC	pM
1	500	100	10	1	80	20
14	500	100	5	1	90	60

En la figura 4.8 se muestran la mejor y peor convergencia del algoritmo para ambas codificaciones. De color azul se distingue a la codificación directa y de color rojo la codificación reducida. Con línea continua se distingue el mejor resultado y con línea punteada el peor resultado. Para el problema 1 se muestra que con la codificación directa no se logró balancear la ecuación. Por otro lado el mejor resultado de la CR

encontró el óptimo global en la 1 generación y en el peor caso el algoritmo converge en la generación 24.

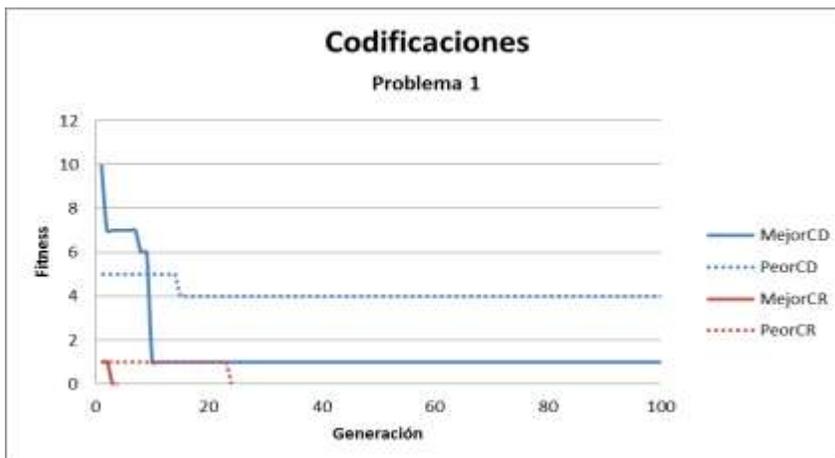


Figura 4.8.- Convergencia de la CD y CR para el problema 1.

En la figura 4.9 se muestra la convergencia del algoritmo implementando la CD y CR, en ambas codificaciones se logra resolver el problema. Sin embargo la codificación reducida tarda menos tiempo en resolver el problema. Comparando el peor caso de la codificación reducida contra el mejor caso de la codificación directa, el mejor resultado de ambos es la codificación reducida.

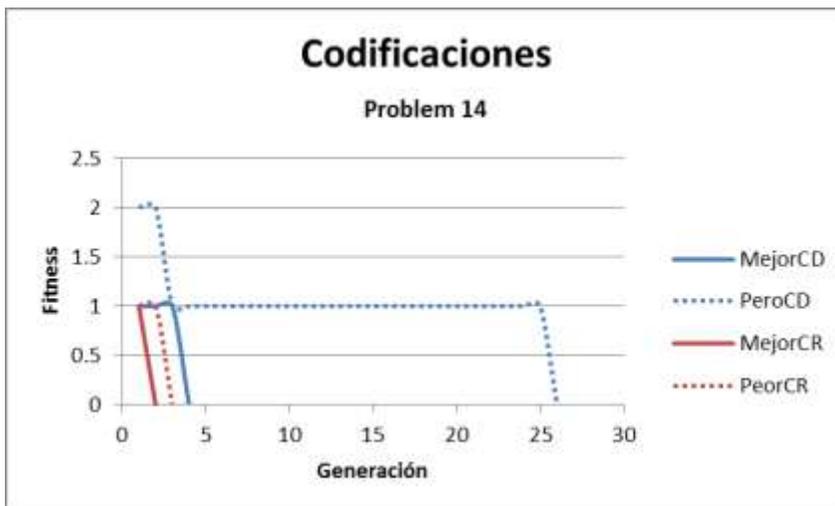


Figura 4.9.- Convergencia de la CD y CR para el problema 14.

El comportamiento mostrado tanto en la figura 4.8 como en la figura 4.9 se observa para la mayoría de los problemas. Todo depende de la cantidad de genes que se bloqueen en la codificación reducida, ya que reduce el espacio de soluciones.

4.7.- Convergencia del algoritmo implementando evaluación del individuo con f y g .

Otro análisis realizado para medir la eficiencia y eficacia del algoritmo se definió implementando dos tipos de función objetivo. Dicho análisis contempla una versión del algoritmo implementando la función F (3.10) y la otra versión correspondo a la función G (3.11). Estas nuevas versiones de algoritmo utiliza la codificación reducida del algoritmo.

Las figuras 4.10 y 4.11 muestra la convergencia de del algoritmo implementando como función objetivo la ecuación 3.10 y la ecuación 3.11. De color azul se distingue el comportamiento del algoritmo implementando la función F y de color rojo corresponde a la función G . Se emplea el mismo criterio para distinguir entre el mejor y peor caso, al igual que el criterio del eje x y eje y . Los valores de las variables de control con cuales se realizaron las pruebas se muestran en la tabla 4.5.

Tabla 4.5.- Valores de variables de control para pruebas de convergencia.

Problema	VARIABLES DE CONTROL					
	tamPob	maxGen	limS	limI	pC	pM
1	500	100	10	1	80	20
8	10000	300	95	1	80	20

En la figura 4.10 se nota que en el peor caso el algoritmo encontró la solución hasta la generación 90 y el mejor caso converge en la generación 3. A diferencia que el

mejor caso para ambas funciones convergen en la misma generación, pero el mejor caso con FG inicia con una solución muy alejada del fitness cero.

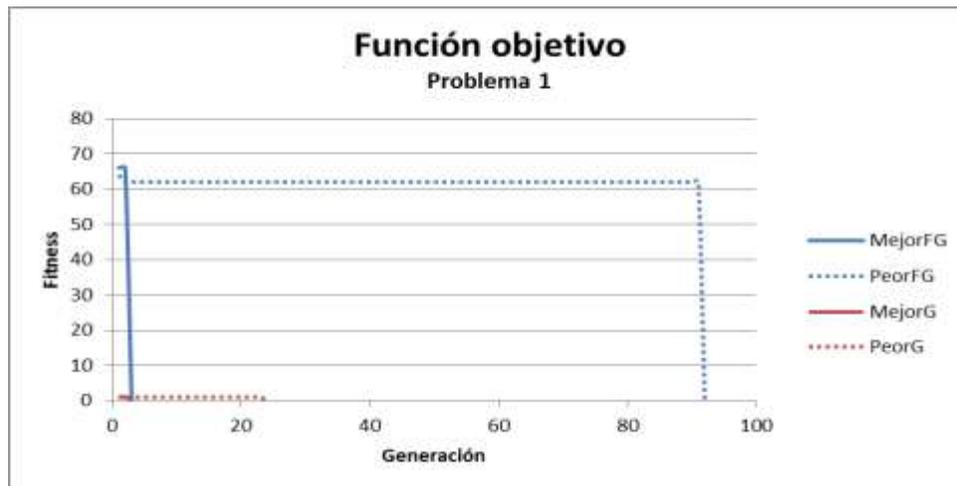


Figura 4.10.- Convergencia de función objetivo GF y G para el problema 1.

En la figura 4.11 se observa que la convergencia para el mejor caso para ambas funciones es el mismo, ambas convergen a la solución óptima en la generación 1. Lo que hace que para el problema 8 no se note gran diferencia en la implementación de la función FG o de la función G.

Observando el comportamiento para ambas gráficas, se observa que es más factible utilizar como función objetivo a G (3.11), además que en capítulo anterior se demuestra que es suficiente realizar la evaluación de la población teniendo como función objetivo a G.

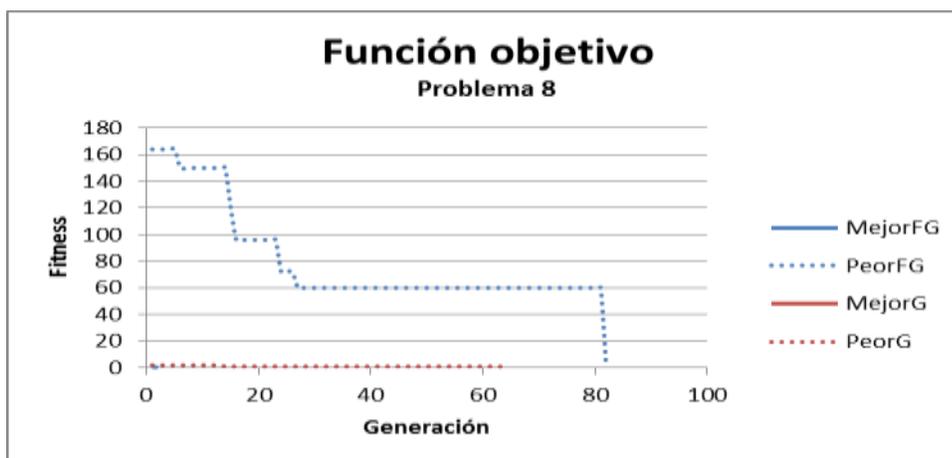


Figura 4.11.- Convergencia de función objetivo FG y G para el problema 8.

4.7.- Análisis de resultados

El espacio de soluciones de cualquier problema de balanceo de ecuaciones químicas se determina en función de dos valores: el máximo valor que pueden tener un gen de cada individuo (LimS) y el número de genes que contenga cada individuo (numGen). Los valores de LimS se calculan en base a la tabla 4.3. Cabe mencionar que existen genes que su valor depende de otro, por lo tanto dichos genes no pertenecen al espacio de soluciones. Por lo tanto para calcular el espacio de solución se contempla el número de genes independientes.

La función para calcular el espacio de solución es $f(x) = \text{LimS}^{\text{numGen} - \text{numGenBloq}}$, donde al número total de genes (numGen) se le resta el número de genes bloqueados (numGenBloq). La tabla 4.6 muestra el espacio de soluciones para cada problema.

Tabla 4.6.- Espacio de Solución para cada problema

Problema	limS	#numGen	#numGenBloq	EspSol
1	10	7	4	10^4
2	20	6	4	20^4

Tabla 4.6.- Espacio de Solución para cada problema

Problema	limS	#numGen	#numGenBloq	EspSol
3	15	6	4	15 ⁴
4	15	7	4	15 ⁴
5	90	6	2	90 ²
6	10	7	3	10 ³
7	20	8	6	20 ⁶
8	15	9	8	15 ⁸
9	90	20	12	90 ¹²
10	30	6	3	30 ³
11	10	5	3	10 ³
12	15	7	7	15 ⁷
13	15	7	7	15 ⁷
14	10	5	3	10 ³
15	10	4	2	10 ²
16	10	4	2	10 ²
17	10	4	2	10 ²
18	10	5	2	10 ²
19	20	4	2	20 ²
20	10	6	2	10 ⁴
21	10	3	0	10 ³

De acuerdo a las características presentadas en la tabla 4.3, se realiza una estimación para las variables de control limS y limI. Cabe mencionar que aún no se ha definido un patrón para la asignación de los valores de las variables de control. La tabla 4.7 muestra los valores asignados para cada problema después de haber realizado la sintonización de variables.

Tabla 4.7.- Valores de variables de control para cada problema.

Problema	VARIABLES DE CONTROL					
	tamPob	maxGen	limS	limI	pC	pM
1	500	100	10	1	80	20
2	500	500	20	1	80	20
3	5000	500	15	1	80	20
4	2000	300	15	1	80	20
5	500	100	90	1	80	20
6	500	300	10	1	80	20
7	50000	1000	20	1	80	20
8	100000	300	15	1	80	20

Tabla 4.7.- Valores de variables de control para cada problema.

Problema	VARIABLES DE CONTROL					
	tamPob	maxGen	limS	limI	pC	pM
9	1000000	500	90	1	80	50
10	10000	300	30	1	80	50
11	1000	100	10	1	80	20
12	1000000	300	15	1	80	20
13	210000	300	15	1	80	20
14	500	100	5	1	90	60
15	1000	500	10	1	80	20
16	500	100	10	1	80	20
17	500	100	10	1	80	20
18	200	50	10	1	80	20
19	100000	1000	20	1	80	50
20	1000	500	10	1	80	50
21	10000	1000	10	1	80	50

Se realizaron 30 ejecuciones experimentales por cada uno de los problemas (tabla 4.2) con los valores de las variables de control mostradas en la tabla 4.7. La tabla 4.8 muestra los resultados obtenidos para cada problema, excepto los problema 9, 19 y 21 debido a que fallo el algoritmo al buscar una solución. De acuerdo a [Zou10] y [SAS06] el problema 19 y 21 se considera inviable.

Tabla 4.8.- Soluciones del AGBEQ

#P	ECUACIÓN
1	$5\text{ClO}_3 + \text{As}_2\text{S}_3 + 5\text{H}_2\text{O} \rightarrow 2\text{H}_2\text{AsO}_4 + 3\text{SO}_4 + 5\text{Cl} + 6\text{H}$
2	$2\text{CH}_3\text{CH}_2\text{OH} + 2\text{Cr}_2\text{O}_7 + 20\text{H} \rightarrow 2\text{CH}_3\text{CO}_2\text{H} + 4\text{Cr} + 12\text{H}_2\text{O}$
3	$4\text{CuSCN} + 7\text{KIO}_3 + 14\text{HCl} \rightarrow 4\text{CuSO}_4 + 7\text{KCl} + 4\text{HCN} + 7\text{I}_2 + 5\text{H}_2\text{O}$
4	$2\text{Ag}_3\text{AsO}_4 + 11\text{Zn} + 11\text{H}_2\text{SO}_4 \rightarrow 2\text{AsH}_3 + 6\text{Ag} + 22\text{ZnSO}_4 + 8\text{H}_2\text{O}$
5	$15\text{PbN}_6 + 44\text{CrMn}_2\text{O}_8 \rightarrow 5\text{Pb}_3\text{O}_4 + 90\text{NO} + 22\text{Cr}_2\text{O}_3 + 88\text{MnO}_2$
6	$\text{K}_4\text{Fe}(\text{CN})_6 + 6\text{SO}_4\text{H}_2 + 6\text{H}_2\text{O} \rightarrow$ $2\text{SO}_4\text{K}_2 + \text{SO}_4\text{Fe} + 3(\text{NH}_4)_2\text{SO}_4 + 6\text{CO}$
7	$2\text{Cu}(\text{NH}_3)_4\text{Cl}_2 + 7\text{KCN} + \text{H}_2\text{O} \rightarrow$ $6\text{NH}_3 + 2\text{NH}_4\text{Cl} + 2\text{K}_2\text{Cu}(\text{CN})_3 + \text{KCNO} + 2\text{ClK}$
8	$6\text{KNO}_3 + 7\text{C} + 3\text{S} \rightarrow \text{K}_2\text{CO}_3 + \text{K}_2\text{SO}_4 + \text{K}_2\text{S}_2 + 5\text{CO}_2 + \text{CO} + 3\text{N}_2$
10	$3\text{As}_2\text{S}_3 + 4\text{H}_2\text{O} + 28\text{HNO}_3 \rightarrow 28\text{NO} + 6\text{H}_3\text{AsO}_4 + 9\text{H}_2\text{SO}_4$

Tabla 4.8.- Soluciones del AGBEQ

#P	ECUACIÓN
11	$KClO_3 + 6HCl \rightarrow KCl + 3Cl_2 + 3H_2O$
12	$C_2H_5NO_2 + C_3H_7NO_3 + 2C_6H_{14}N_4O_2 + 3C_5H_9NO_2 + 2C_9H_{11}NO_2 \rightarrow 8H_2O + C_{50}H_{73}N_{15}O_{11}$
13	$C_5H_9NO_2 + 3C_5H_{11}NO_2 + 2C_6H_{13}NO_2 + 3C_9H_{11}NO_2 + 2C_5H_{12}N_2O_2 \rightarrow H_2O + 10C_{60}H_{92}N_{12}O_{10}$
14	$2KNO_3 + 4C_2 \rightarrow K_2CO_3 + 3CO + N_2$
15	$4NH_3 + 5O_2 \rightarrow 4NO + 6H_2O$
16	$2C_3H_6 + 9O_2 \rightarrow 6CO_2 + 6H_2O$
17	$2CaCO_3 + 5C \rightarrow 2CaC_2 + 3CO_2$
18	$FeCrO_4 + 4C \rightarrow Cr + Fe + 4CO$
20	$2KClO_3 + 4HCl \rightarrow 2KCl + 2ClO_2 + Cl_2 + 2H_2O$

La tabla 4.9 muestra el número de veces que se encontró la solución de las 30 ejecuciones, así como el menor, mayor y tiempo promedio de ejecución.

Tabla 4.9.- Resultados de las pruebas experimentales

Problema	# de Balanceos	Tiempo de ejecución (mseg)		
		Menor	Mayor	Promedio
1	26	87	4118	1443
2	30	3	4780	1134
3	30	70	11121	1971
4	30	139	12420	3346
5	28	7	3235	1160
6	30	5	2016	263
7	15	15866	538944	728500
8	19	1008	357016	274198
10	30	74	2641	675
11	30	4	471	66
12	9	24500	235988	210938
13	30	78373	1725794	517972
14	30	2	42	10
15	30	3	5	3
16	30	2	8	2
17	30	1	3	2
18	30	1	4	1
20	30	4	121	27

En base a los resultados mostrados en la tabla 4.9, se nota que el algoritmo genético para el balanceo de ecuaciones químicas no muestra dificultad en resolver la mayoría de los problemas. El algoritmo fue 100% eficaz en el 66% de los problemas analizados.

De los 21 problemas analizados no se logra encontrar resultado para el problema 9, 19 y 21. Campanario en [Cam95] explica algunos casos de porque no es posible balancear algunas ecuaciones. Para el problema 19 y 21 Campanario menciona que es imposible realizar el balanceo simultáneo para todos los elementos. El algoritmo genético no encontró alguna solución, el grado de adaptación de los mejores individuos para cada problema es de 1 y 2 respectivamente. También dicho artículo deduce que el programa que desarrollaron es posible que balanceen una ecuación cuando los valores en lo que oscilan los coeficientes son muy grandes. Sin embargo el algoritmo genético no tiene complicación en dicho punto para el problema 5, pero para el problema 9 falló. Por lo tanto, se concluye que tiene mayor complejidad un problema cuando el rango en que oscilan los coeficientes en grandes (mayor de 60) y si la ecuación está formada por más de 10 compuestos.

El problema 9 de acuerdo a [Ris09] los coeficientes para dicha ecuación inicialmente son valores con decimales, luego realizan corrimiento del punto decimal logrando obtener valores mayores a 1000, dichos valores se trataron de minimizar, pero no se logró. Sin embargo los resultado del algoritmo genético los coeficientes no rebasan el valor de 7, teniendo una mejor solución.

Los problemas 15-18, el tiempo de ejecución promedio está por debajo de los 5 segundos, dichos problemas de acuerdo a la tabla 3 el número de elementos que tienen las ecuaciones es menor a 5 al igual que el número de compuestos. Por ello se pueden clasificar como problemas pequeños y fáciles de resolver.

El problema 13 es el que mayor tiempo de ejecución consume (excluyendo el problema 9 que no fue resuelto). Cabe mencionar que el problema 13 y 12 comparte características en la estructura de la ecuación. A pesar de que el tamaño de la población para la prueba del problema 13 equivale aproximadamente a un 20% del tamaño de la población de la prueba del problema 12, el problema 13 consumió más tiempo en las 30 ejecuciones para la prueba realizada.

4.7.- Comentarios finales

Después de los resultados obtenidos se concluye que el algoritmo genético desarrollado funciona adecuadamente para la mayoría de las ecuaciones analizadas, implementando la codificación reducida y teniendo como función objetivo a G. Por otro lado, se plantea seguir realizando pruebas hasta definir los valores adecuados de las variables de control, para obtener el resultado deseado.

CONCLUSIONES

Como se puede observar en este trabajo, aunque existen diferentes enfoques para balancear ecuaciones químicas, todos están limitados a ciertos tipos de ecuaciones químicas. En el presente artículo se presenta un algoritmo genético que, en base a una formulación matemática, puede balancear cualquier tipo de ecuación química, siempre y cuando está sea factible.

En [SAS06], la solución que se propone es para ecuaciones donde el número de elementos y de compuestos es diferente. En [Ris08] Risteski plantea la formulación matemática para el balanceo de ecuaciones que contienen el mismo número de elementos como de compuestos. La formulación matemática desarrollada dentro de la investigación contempla que las ecuaciones donde el número de elementos presentes en la ecuación es diferente al número de compuestos.

Existen diferentes factores que influyen en el comportamiento del algoritmo, como el rango en que oscilan los valores de los coeficientes y el número de compuestos. En base a las pruebas realizadas es difícil concluir el factor que determina la complejidad del problema. En primera estancia se concluye que el rango en el que oscila los valores de los coeficientes de la solución de un problema impacta en el tiempo de ejecución del algoritmo, debido a que aumenta el espacio de soluciones, aumentando exponencialmente la combinación de valores que contienen los genes de los individuos de la población.

En segundo plano se contempla el tamaño del cromosomas del individuo definido para el problema, entre mayor sea el número de compuestos del problema más genes contendrá el individuo. Sin embargo, se desarrolló un mecanismo que determina las dependencias que existen entre los coeficientes de los elementos que están presentes tanto de los reactivos como de los productos una sola vez. Con dicho mecanismo de codificación se reduce el espacio de solución del problema, al determinar la función de los coeficientes tanto de los reactivos como de los productos que tienen relación.

También se notó cierta dificultad de solución en el problemas con mayor número de elementos que de compuestos. Por lo tanto, al evaluar a los individuos de la población y como existe mayor número de elementos, los coeficientes tienen que satisfacer a todos los elementos que afectan.

Es importante realizar un análisis de sensibilidad de cada problema para mejorar la eficacia del algoritmo. Se plantea seguir realizando pruebas con otros problemas con características diferentes para tratar de determinar los diferentes factores que influyen en la complejidad del problema a resolver.

Hasta el momento se sigue siendo pionero en dicho campo de estudio, ya que no se ha encontrado en la literatura científica alguna investigación de implementación de algoritmos genéticos para el balanceo de ecuaciones.

REFERENCIAS

REFERENCIAS

- [AC09]** Araujo, L. y Cervigón, C: Algoritmos evolutivos. Un enfoque práctico, Alfaomega, 1ra ed, México (2009).
- [ACA00]** F. Ascolano, M.A. Carzola, M. I. Alfonso, O. Colombina, M. A. Lozano: Inteligencia Artificial. Modelos, Técnicas y Áreas de Aplicación (Madrid, España: Ed. Parainfo, 2000).
- [Ang60]** A. Angiolani: Introducción a la Química Industrial: Fundamentos químicos y tecnológicos, Ed. Andrés Bello, Chile (1960).
- [AWT09]** D. Aljanati, E. Wolovelsky y C. Tambussi: Biología 3: Los códigos de la vida (Buenos Aires, Argentina: Ed. Colihue, 2009).
- [Berg01]** F. van den Bergh, An Analysis of Particle Swarm Optimizers, Tesis Doctoral, University of Pretoria, Sudáfrica, 2001.
- [BHS91]** T. Back, F. Hoffmeister y H. P. Schwefel, A Survey of Evolution Strategies, en Proc. of the Fourth Int. Conf. on Genetic Algorithms, San Mateo, E.U., 1991, 2-9.
- [BHS97]** T. Bäck, U. Hammel y H. P. Schwefel, Evolutionary Computation: Comments on the History and Current State, in IEEE Trans. on Evolutionary Computation, 1 (1), 1997, 3-17.

- [BoI03]** V. Bolaños Chombo: Química Analítica Cualitativa: reacciones en solución. Editorial UAEMex (2003).
- [BOP06]** E.J. Boottani, H.S. Odetti, O. H. Pliego, E.R. Villareal: Química General, Ed. Universidad Nacional del Litoral, Colombia (2006).
- [BW77]** R. S. Becker y W. E. Wentworth: Química general, Ed. Reverté, España, 1977.
- [Cam95]** Campanario, J. M.: Automatic Balancing of Chemical Equation, en Elsevier Science: Computer Chem, Vol. 19, No. 2, pp 85-99 (1995).
- [CFV07]** Croteau, J., Fox, W. P. y Varazo, K.: Mathematical modeling of chemical stoichiometry, en Primus: Problem, Resources, and Issues in Mathematics Undergraduate Studies, Vol. 17, No. 4, pp 301-315 (2007).
- [Chu97]** Chunsi, C.: A new Inspection Method for Balancing Redox Equations, en Journal of Chemical Education, Vol. 74, No. 11, pp 1365-1366 (1997).
- [Cor02]** C. A. Correa Maya: Fenómenos Químicos. Fondo Editorial Universidad EAFIT, Colombia (2002).
- [Cot03]** P., Moscato y C. Cotta: An Introduction to Memetic Algorithms, en Revista Iberoamericana de Inteligencia Artificial, Vol. 19, pp 131-148 (2003).

- [Dep10]** Departamento de matemáticas ITESM, Matrices Invertibles y Elementos de Álgebra Matricial, URL: <http://www.mty.itesm.mx/> (2010).
- [DLJ00]** D. Dumitrescu, B. Lazzerini, L.C. Jain y A. Dimitrescu, Evolutionary Computation (Boca Ratón: CRC Press, 2000).
- [DSC86]** Das, S. C., 1986, "A Mathematical Method of Balancing a Chemical Equation," Int. J. Math. Educ. Sci. Technol., 17, pp. 191-200.
- [Fog98]** D.B. Fogel, Evolutionary Computation. The Fossil Record (NY: IEEE Press, 1998).
- [Fog00]** D. B. Fogel, Evolutionary Computation (Piscataway: IEEE Press, 2000).
- [Ges10]** Gestal, P. M.: Introducción a los Algoritmos Genéticos, Depto. Tecnologías de la Información y las Comunicaciones Universidades Coruña, URL: <http://sabia.tic.udc.es/mgestal>
- [Gut86]** E. Gutiérrez Ríos: Química. Ed. Reverté, España (1986).
- [Holl75]** J. H. Holland, Adaptation in Natural and Artificial Systems (Ann Arbor: The University of Michigan Press, 1975).
- [KG99]** Kuri, A. y Galaviz, J.: Algoritmos Genéticos, URL: <http://pateame.fcencias.unam.mx/libag/> (1999)

- [Koz89]** J. R. Koza, Hierarchical Genetic Algorithms Operating on Populations of Computers Programs, en [Fog98], 1989.
- [Kum01]** Kumar, D. D.: Computer Applications in Balancing Chemical Equations, Journal of Science Education and Tecnology, Vol. 10, No.4, (2001)
- [Lir11]** Lira, J. D.: 3. Estequiometría, Instituto Tecnológico Superior de Calkini en el Estado de Campeche, URL:
<http://www.itescam.edu.mx/principal/sylabus/fpdb/recursos/r54618.DOCX>
(2011)
- [McL02]** D. McLaughlin: Chemical Principles, en
(<http://www.unm.edu/~dmclaugh/PrinciplesPDF.html>)
- [MM97]** Z. Michalewicz y M. Michalewicz, Evolutionary Computation, techniques and their Applications, en Proc. of the 1997 IEEE Int. Conf. on Intelligent Processing Systems, Beijing, China, 1997, 14-25.
- [Ols97]** Olson, J. A.: An Analysis of the Algebraic Method for Balancing Chemical Reactions, en Journal of Chemical Education, Vol. 74, No. 5, pp 538-542 (1997).
- [Pet98]** Petrusevki, V. M. : A fast solution to the problem of balancing redox equation, en Bulletin of the Chemist and Technologists of Macenodia, Vol. 17, No. 2, pp 141-145 (1998).

- [Rec65]** I. Rechenberg. Cybernetic Solution Path of an Experimental Problem, en [Fog98], 1965.
- [Ris08]** Risteski, I. B.: A New Generalized Matrix Inverse Balancing Chemical Equations and Their Stability, en Sociedad Química de México, Vol. 2, No. 3, pp 104-115 (2008).
- [Ris09]** Risteski, I. B.: A New Singular Matrix Method for Balancing Chemical Equations and Their Stability, en Journal of the Chinese Chemical Society, Vol. 56, pp 65-79 (2009).
- [SAS06]** Sen, S., K., Agarwal, H. y Sen S.: Chemical Equation Balancing: An Integer Programming Approach, en Elsevier: Mathematical and Computer Modelling, Vol. 44, pp 678-691 (2006).
- [SGC95]** Subramaniam, R, Goh, N K, Chia, L S.: The relationship between the number of elements and the number of independent equations of elemental balance in inorganic chemical equations, en Journal of Chemical Education, Vol. 72, No.. 10, pg. 894 (1995).
- [Zou10]** Dexuan Zou, LiqunGao, Yanfeng Ge, Peifeng Wu: A Novel Global Harmony Search Algorithm for Chemical Equation Balancing, 2010 International Conference On Computer Design And Applications (ICCCA 2010)

