#### CAPITULO I

#### Introducción: Los algoritmos genéticos

* 1. **Introducción**

En este primer capítulo se relata la utilidad de los algoritmos genéticos, así como su uso en problemas de optimización y de simulación. Además se explica un poco de la historia de los algoritmos genéticos, y se nombra diversas aplicaciones de los mismos, de las cuales he podido desarrollar unos cuantos ejemplos para la comprensión de los mismos.

Se conoce que muchos problemas pueden ser resueltos de una forma computacional determinística, mientras que otros no tienen un método de resolución exacta y utilizan métodos numéricos o técnicas computacionales; pero todos estos métodos de resolución son complejos en su implantación y uso. Para la solución de éstos puede utilizarse métodos heurísticos y metaheurísticos como los algoritmos genéticos.

Los algoritmos genéticos forman parte de la inteligencia artificial de los sistemas inspirados en la naturaleza y la evolución, y son métodos que simulan los procesos naturales y los aplican a la solución de problemas reales de búsqueda, optimización, diseño y simulación, aplicando la idea darwiniana de la selección natural de acuerdo a la aptitud y la disposición, y la combinan con otros operadores genéticos para producir métodos de gran robustez y de aplicación óptima.

Los algoritmos genéticos es una técnica de búsqueda basada en la teoría de la evolución de Darwin, que ha cobrado tremenda popularidad en todo el mundo durante los últimos años. Se presentarán aquí los conceptos básicos que se requieren para abordarla, así como unos sencillos ejemplos que permitan a los lectores comprender cómo aplicarla al problema de su elección.

En los últimos años, la comunidad científica internacional ha mostrado un creciente interés en una nueva técnica de búsqueda basada en la teoría de la evolución y que se conoce como el algoritmo genético. Esta técnica se basa en los mecanismos de selección que utiliza la naturaleza, de acuerdo a los cuales los individuos más aptos de una población son los que sobreviven, al adaptarse más fácilmente a los cambios que se producen en su entorno. Hoy en día se sabe que estos cambios se efectúan en los genes de un individuo que es la unidad básica de codificación de cada uno de los atributos de un ser vivo, y que sus atributos más deseables, es decir los que le permiten adaptarse mejor a su entorno, se transmiten a sus descendientes cuando éste se reproduce sexualmente.

Una definición bastante completa de un algoritmo genético es la propuesta por John Koza: "Es un algoritmo matemático altamente paralelo que transforma un conjunto de objetos matemáticos individuales con respecto al tiempo usando operaciones modeladas de acuerdo al principio Darwiniano de reproducción y supervivencia del más apto, y tras haberse presentado de forma natural una serie de operaciones genéticas de entre las que destaca la recombinación sexual. Cada uno de estos objetos matemáticos suele ser una cadena de caracteres (letras o números) de longitud fija que se ajusta al modelo de las cadenas de cromosomas, y se les asocia con una cierta función matemática que refleja su aptitud."

Un aspecto por demás importante de ellos es su aplicación como técnica de optimización que se basa en el azar, pero que aprovecha criterios que la naturaleza ha desarrollado, tales como la selección de los cromosomas más aptos, el cruce de genes en los cromosomas y la mutación. Por esto, no es de extrañarse que en algoritmos genéticos se utilicen términos tomados de la genética natural.

Estos algoritmos están basados en los procesos genéticos de los organismos biológicos, codificando una posible solución a un problema de cromosomas compuesto por cadena de bits y caracteres. Estos cromosomas representan individuos qué son llevados a lo largo de varias generaciones, en forma similar a los problemas naturales, evolucionando de acuerdo con los principios de selección natural y de supervivencia de los más aptos, descritos por primera vez por Charles Darwin en su libro "El Origen de las Especies". Simulando estos procesos, los algoritmos genéticos son capaces de evolucionar soluciones de problemas del mundo real.

En la naturaleza, los individuos compiten entre sí por recursos, tales como comida, agua y refugio. Adicionalmente, entre los animales de una misma especie, aquellos que no obtienen acierto tienden a tener un número reducido de descendientes, teniendo por tanto, menor probabilidad de que sus genes sean propagados a lo largo de sucesivas generaciones. La combinación entre los genes de los individuos que perduran en una especie, puede producir un nuevo individuo mucho mejor adaptado a las características de su medio ambiente.

Los algoritmos genéticos utilizan una analogía directa de este fenómeno de evolución en la naturaleza, donde cada individuo representa una posible solución para un problema dado. A cada individuo se atribuye una puntuación de adaptación, dependiendo de la respuesta dada al problema por este individuo. A los más adaptados se les da la oportunidad de reproducirse mediante cruces con otros individuos de la población, produciendo descendientes con características de ambas las partes.

Si un algoritmo genético es desarrollado correctamente, la población (conjunto de posibles respuestas) convergerá a una buena solución para el problema propuesto, tal vez hasta óptima. Los procesos que más contribuyen para la evolución son el cruce de genes en cromosomas y la adaptación basada en una selección y de reproducción. La mutación también tiene un papel significativo, en tanto, su importancia continúa siendo un asunto en que los investigadores aún no entran en consenso, ya que al igual que en la naturaleza, debe darse con una posibilidad muy pequeña.

Un algoritmo genético puede converger en una búsqueda de azar, pero su utilización asegura qué ningún punto del espacio de búsqueda tiene probabilidad cero de ser examinado. Toda tarea de búsqueda y optimización posee varios componentes, entre ellos se encuentran:

* El espacio de búsqueda, donde son consideradas todas las posibilidades de solución de un determinado problema, y
* La función de validación (o función de coste), es una manera de validar los miembros del espacio de búsqueda.

Como técnicas de búsqueda y optimización tradicionales comienzan con un único candidato que, iterativamente, es manipulado utilizando algunas técnicas heurísticas (estáticas) directamente asociadas al problema a ser solucionado y por otro lado, las técnicas de computación y evolución operan sobre una población de candidatos en paralelo. Así, ellos pueden satisfacer la búsqueda en diferentes áreas del espacio de solución, alocando un número de miembros apropiado para la búsqueda en varias regiones. Los algoritmos genéticos difieren de los métodos tradicionales de búsqueda y optimización, principalmente en cuatro aspectos:

1. Los algoritmos genéticos trabajan con una codificación del conjunto de parámetros y no con los propios parámetros;
2. Los algoritmos genéticos trabajan con una población de varios puntos y no con un único punto;
3. Los algoritmos genéticos utilizan la misma función y no derivadas u otro conocimiento auxiliar para adaptarse al problema;
4. Los algoritmos genéticos utilizan reglas de transición probabilísticas y no determinísticas.

Los Algoritmos Genéticos son muy eficientes para búsqueda de soluciones óptimas, o aproximadamente óptimas (buenas soluciones), en una gran variedad de problemas, porque no imponen muchas de las limitaciones encontradas en los métodos de búsqueda tradicionales. Los investigadores se refieren a los algoritmos genéticos o a un algoritmo genético y no al algoritmo genético, porque los algoritmos genéticos son una clase de procedimientos con muchos pasos separados, y cada uno de estos pasos posee muchas variaciones posibles.

Antes de continuar ahondando en la técnica de los Algoritmos Genéticos he pensado que sería interesante dejarla situada dentro de un marco más amplio. Me refiero con esto a la rama de la Inteligencia Artificial que se ha denominado Computación Evolutiva.

El término Computación Evolutiva se refiere al estudio de los fundamentos y aplicaciones de ciertas técnicas heurísticas de búsqueda basadas en los principios naturales de la evolución. Una gran variedad de algoritmos evolutivos ha sido propuesta; pero principalmente pueden clasificarse en: Algoritmos Genéticos, Programación Evolutiva, Estrategias Evolutivas, Sistemas Clasificadores y Programación Genética.

Esta clasificación se basa sobre todo en detalles de desarrollo histórico más que en el hecho de un funcionamiento realmente diferente, de hecho las bases biológicas en las que se apoyan son esencialmente las mismas. Las diferencias entre ellos se centra en los operadores que se usan en cada caso y en general en la forma de implementar la selección, reproducción y sustitución de individuos en una población.

Aunque los detalles de la evolución no han sido completamente comprendidos, incluso hoy en día, existen algunos puntos en los que se fundamentan:

1. La evolución es un proceso que opera a nivel de cromosomas, y no a nivel de individuos. Cada individuo es codificado como un conjunto de cromosomas.
2. La selección natural es el mecanismo mediante el cual los individuos mejor adaptados son los que tienen mayores posibilidades de reproducirse.
3. El proceso evolutivo tiene lugar en la etapa de la reproducción. Es en esta etapa donde se producen la mutación, que es la causante de que los cromosomas de los hijos puedan ser diferentes a los de los padres, y el cruce, que combina los cromosomas de los padres para que los hijos tengan cromosomas diferentes.

Los Algoritmos Genéticos resuelven los problemas generando poblaciones sucesivas a las que se aplican los operadores de mutación y cruce. Cada individuo representa una solución al problema, y se trata de encontrar al individuo que represente a la mejor solución.

La programación genética funciona igual que la técnica anterior pero se basa en el estudio de problemas cuya solución es un programa. De manera que los individuos de la población son programas que se acercan más o menos a realizar una tarea que es la solución.

La Programación Evolutiva es otro enfoque de los algoritmos genéticos, en este caso el estudio se centra en conseguir operadores genéticos que imiten lo mejor posible a la naturaleza, en cada caso, más que en la relación de los padres con su descendencia. En este caso no se utiliza el operador de cruce, tomando la máxima importancia el operador de mutación.

* 1. **Funcionamiento de los algoritmos genéticos**

El funcionamiento de un algoritmo genético se basa en la creación de un conjunto de posibles soluciones (población), representadas cada una de ellas como una cadena de símbolos que pueden ser caracteres, valores, o direcciones de memoria. Luego se califica a cada solución de acuerdo con una función de aptitud definida para el problema en cuestión. Con base en tal calificación se seleccionan elementos de ese conjunto para que se les apliquen los operadores genéticos. Dichos operadores tienen como objetivo crear nuevos individuos a partir de los anteriores, conservando de esta forma las características que hicieron buenos a estos últimos.

Cada representación de las posibles soluciones depende del tipo de problema sobre el que se aplicará un algoritmo genético, es decir que para obtener un buen algoritmo genético debe implantarse una correcta representación, y luego una buena función de aptitud. Sólo de este modo, tendremos un algoritmo genético que convergerá a la solución deseada.

* 1. **Ventajas y desventajas de los algoritmos genéticos**

El hecho tener una población de posibles soluciones hace que el algoritmo genético esté explorando siempre varias soluciones en paralelo, y además se analizará cada una de las posibles soluciones, ya que cada punto tendrá igual probabilidad de ser seleccionado inicialmente. Es de aquí de donde aparece la dificultad de que un algoritmo genético se quede atrapado en un óptimo local. Esta es posiblemente una de las mayores ventajas de los algoritmos genéticos cuando se les ve como una alternativa en la optimización. Esta característica hace natural pensar en implementar algoritmos genéticos en sistemas de cómputo paralelo. Además, se debe considerar que la implantación de los algoritmos genéticos en algún lenguaje de programación es relativamente simple, y aunque antes se pensaba que sólo se los podía implantar en lenguajes que utilicen el concepto de predicados; pero actualmente se han desarrollado muchos algoritmos genéticos en lenguajes hasta de cuarta generación, dándoles de esta manera un ambiente más apropiado a la actualidad.

Por otro lado, entre las desventajas de los algoritmos genéticos tenemos el hecho de que hay que diseñar funciones para representar las soluciones mediante cadenas y para calificar a los individuos de la población. Como eso depende del problema, pueden darse casos en que no sea tan fácil encontrarlas, o de aplicarlas en el problema. Además, los algoritmos genéticos son una herramienta de propósito general que normalmente son superados cuando existen algoritmos especializados para un problema particular.

En problemas donde obtener un resultado exacto es importante, el algoritmo genético puede servir simplemente como base para que luego otro algoritmo encuentre el óptimo, con una buena aproximación. En este caso el algoritmo genético lo que hace es encontrar un punto muy cercano a la solución, lo que posiblemente a un algoritmo tradicional no le hubiera sido posible si la función analizada fuera muy compleja.

Los Sistemas de adaptación tratan de resolver problemas acumulando conocimiento sobre el mismo y utilizando estas informaciones para generar soluciones aceptables. Estos problemas, típicamente, se encuentran en las áreas de configuración de sistemas complejos, colocación de tarifas, selección de rutas, máximos, mínimos, búsquedas y otros problemas de optimización y aprendizaje de máquinas.

#### Aplicaciones que pueden realizarse con algoritmos genéticos.

La aplicación más común de los algoritmos genéticos ha sido la solución de problemas de optimización, en donde han mostrado ser muy eficientes y confiables. Sin embargo, no todos los problemas pudieran ser apropiados para la técnica, y se recomienda en general tomar en cuenta las siguientes características del mismo antes de intentar usarla:

1. Su espacio de búsqueda, es decir sus posibles soluciones, debe estar delimitado dentro de un cierto rango.
2. Debe poderse definir una función de aptitud que nos indique qué tan buena o mala es una cierta respuesta.
3. Las soluciones deben codificarse de una forma que resulte relativamente fácil de implementar en la computadora.

El primer punto es muy importante, y lo más recomendable es intentar resolver problemas que tengan espacios de búsqueda discretos, aunque éstos sean muy grandes; sin embargo, también podrá intentarse usar la técnica con espacios de búsqueda continuos, pero preferentemente cuando exista un rango de soluciones relativamente pequeño para cubrir una mayor parte del mismo.

* + 1. **Ejemplos de aplicaciones realizadas:**

Los siguientes son ejemplos de aplicaciones que he realizado mientras comprendía el funcionamiento de los algoritmos genéticos, y planteaba como utilizarlos en mi aplicación dirigida a los modelos de crecimiento poblacional.

1. **El problema del vendedor viajero (TSP)**

Es una pequeña aplicación realizada en Visual C++ 6.0 que optimiza rutas. Esta aplicación lee un archivo de texto ubicado en el mismo directorio que la aplicación se encuentra. En esta aplicación, se utilizó una representación de los cromosomas por medio de permutaciones. Además debe configurarse como parámetros del algoritmo, el tamaño de la población, el tamaño del torneo, el número de iteraciones y la semilla aleatoria. Estos son explicados en detalle en el capítulo II, para una mejor comprensión de los algoritmos genéticos.

1. **Punto máximo en una función normal estándar**

Es una pequeña aplicación que me permitió comprender el algoritmo genético básico representando los cromosomas mediante cadenas binarias. Está desarrollado completamente en Visual C++ 6.0 en el modo ventana de diálogo con MFC, y pide el ingreso del número de iteraciones, tamaño de la población, intervalo de acción, probabilidad de crossover, probabilidad de mutación, y probabilidad de convergencia.

1. **Algoritmos genéticos – modelo poblacional**

Es la principal aplicación que realicé, con la que determino el crecimiento de una población heterogénea, basada en el número de hijos, edad límite de vida, intervalo de edades de reproducción, y la configuración del algoritmo genérico que presenta algunas opciones que luego son explicadas detalladamente.

Este es ámbito al que nos vamos a dedicar en este proyecto, en este caso se aprovecha la selección natural, el cruce y la mutación para simular a través de generaciones, y observar las medidas de desempeño de los índices poblacionales que nos interesan, tales como tasa de natalidad, tasa de mortalidad y las tasas de migraciones, así como la edad promedio de la población y la distribución por sexos.

* + 1. **Ejemplos de otras aplicaciones que pueden realizarse:**
1. **Planificación de multitarea**

Una aplicación de los algoritmos genéticos en un problema de asociación óptima de procesos y procesadores. Un objetivo es disminuir el costo que se deriva de la comunicación entre procesos en un ordenador paralelo de memoria distribuida. La Planificación de multitarea también puede ser relacionada con problemas de robótica.

1. **Biología molecular y física o química**

Existe una hipótesis de trabajo que sustentaría el uso de técnicas basadas en poblaciones debido a una estructura de función objetivo. En problemas relacionados con estructura molecular existen experimentos que dan indicios claros sobre la validez de esa hipótesis. Así que es fácil observar que son cada vez más numerosas las aplicaciones de algoritmos genéticos en este campo.

###### **Ingeniería en construcciones**

###### Los Algoritmos genéticos se han ganado la aceptación en un gran número de problemas de ingeniería, estos problemas están basados en la optimización de recursos. Una aplicación es la optimización discreta de estructuras.

1. **Búsqueda en bases de dados**

Así como existen los métodos de búsqueda convencionales en bases de datos, los algoritmos genéticos pueden facilitar dichas búsquedas, dando resultados muy eficientes.

# Geofísica

# Para un problema llamado de Inversión de la forma de la ola sísmica, generalmente se utiliza la estadística Bayesiana, a dicho problema se asocia una información que es apriori sobre un modelo. Surgen así funciones que requieren métodos de optimización global, donde es de mucho provecho el uso de los algoritmos genéticos.

1. **Compresión de Dados**

La compresión de dados en general, es la compresión de imágenes sólidas en particular. Esta aplicación consiste en encontrar un método que utiliza los Algoritmos genéticos para encontrar un sistema de funciones locales iteradas (LIFS) para la codificación de imágenes. Produciendo como resultado final una imagen con calidad similar a la utilización de un método convencional de compresión fractal, con un tiempo 30% menor.

1. **Optimización de rutas**

Cuando queremos optimizar rutas, existen muchas técnicas de programación que nos llevan a una respuesta adecuada. Podemos aplicar también los algoritmos genéticos para hallar una ruta óptima.

Además existen aplicaciones de los algoritmos genéticos muy generales, que pueden ser adaptadas a cualquier necesidad, siempre que se tenga cuidado al considerar la correcta representación de sus elementos.

Entre estos algoritmos encontramos la aplicación GaUI 1.0, la cual está desarrollada completamente en un lenguaje de cuarto nivel, y presenta muchas etapas de los algoritmos genéticos para la resolución de diferentes problemas comunes.

* 1. **Historia**

Para comprender mejor la historia de los algoritmos genéticos, se ha dividido en dos partes muy claras como son la evolución y la informática evolutiva, en la primera se relatan los hechos evolutivos más notables, mientras en la segunda se comentan los hechos más sobresalientes de la informática evolutiva.

Se desea estudiar qué es lo que inspira dichos algoritmos, la naturaleza, ese fenómeno natural denominado evolución, y si de veras optimiza o no. Además es importante conocer un poco más sobre los personajes que destacan en esta reseña de la historia de los algoritmos genéticos.

* + 1. **Personajes**



***Carl Linnaeus***

***Figura 1.1***

***Carl Linnaeus*** (1707-1778), Conocido también como Carolus Linnaeus, es conocido como el padre de la Taxonomía, su sistema de nombramiento, discriminación y clasificación de organismos es usado en la actualidad con algunos pequeños cambios. Sus ideas de clasificación han influenciado a algunas generaciones de biólogos, incluso aquellos que se opusieron a las raíces filosóficas y teológicas de su trabajo.

******

***Thomas Robert Malthus***

***Figura 1.2***

***Thomas Robert Malthus***(1766 - 1834), Economista británico, discípulo de Adam Smith. Expuso sus teorías en la obra "Primer Ensayo sobre la población" (1798).

Se señala que de acuerdo a Malthus la población suele aumentar en una proporción geométrica y la producción de alimentos sólo puede aumentar en una proporción aritmética.

En este cuadro las diversas medidas de control de natalidad se convierten en un factor clave en la lucha por el desarrollo, aun cuando no se llega a asegurar que controlado el crecimiento de la población el progreso será realmente posible.

La aparición de Malthus en este trabajo es importante, ya que es un personaje que está involucrado tanto en la historia de los algoritmos genéticos como en la de los modelos poblacionales.



***Eramus Darwin***

***Figura 1.3***

**Eramus Darwin** (1731 – 1802), Era el abuelo de Charles Darwin, fue uno de los principales intelectuales de decimoctavo siglo Inglaterra, un hombre con una serie notable de intereses y persecuciones.

Como un naturalista, él formuló uno de las primeras teorías formales en evolución en Zoonomía, o, Las Leyes de Vida Orgánica (1794-1796). Él también presentó sus ideas evolutivas en verso, en particular en el poema póstumamente publicado El Templo de Naturaleza.



***Charles Darwin***

***Figura 1.4***

***Charles Darwin***(1809 –1882), nació el 12 de febrero de 1809 en Shrewsbury, Inglaterra. Él era el británico naturalista que tomó fama por sus teorías de evolución y la selección natural. Como varios científicos antes que él, Darwin creyó toda la vida en tierra evolucionada (desarrollado gradualmente) sobre de millones de años de unos antepasados comunes.



***Alfred Russel Wallace***

***Figura 1.5***

**Alfred Russel Wallace** (1823 - 1913) es uno de los padres olvidados de ciencia moderna. Él nació en el pueblo de Usk en Monmouthshire, Inglaterra. Su padre se murió cuando Alfred era joven.

Estuvo en Singapur desde 1853 y durante los próximos ocho años que Alfred Russel Wallace hizo el gran viaje que llevó a su formulación de la teoría de Selección Natural. Los nombres de Charles Robert Darwin (1809-1882) y Alfred Russel Wallace (1823-1913) tienen gran sentido con el concepto moderno de evolución y la teoría de selección natural.

En 1900, la teoría moderna de evolución combina la genética y las ideas de Darwin y Wallace sobre la selección natural, creando el principio básico de Genética Poblacional: la variabilidad entre individuos en una población de organismos que se reproducen sexualmente es producida por la mutación y por la recombinación genética.



*John H. Holland*

***Figura 1.6***

**John Holland**, Se le conoce como el padre de los algoritmos genéticos. John Holland desde pequeño, se preguntaba cómo logra la naturaleza, crear seres cada vez más perfectos.

Fue a principios de los 60, en la Universidad de Michigan en Ann Arbor,donde, dentro del grupo lógica de computadoras, sus ideas comenzaron a desarrollarse y a dar frutos.

En esa universidad, Holland impartía un curso titulado: “Teoría de sistemas adaptativos”, dentro de este curso, y con una participación activa por parte de sus estudiantes, fue donde se crearon las ideas que más tarde se convertirían en los algoritmos genéticos.



*David Golberg*

***Figura 1.7***

David Golberg, en la actualidad es el maestro de los algoritmos genéticos, conoció a Holland, y se convirtió en su estudiante y discípulo, aprendió su teoría y la aplicó mediante la computación evolutiva.

# La evolución

La teoría de la evolución (que no es tal teoría, sino una serie de hechos probados), fue descrita por Charles Darwin veinte años después de su viaje por las islas Galápagos en el Beagle en el libro Sobre el Origen de las Especies por medio de la Selección Natural. Este libro fue bastante polémico en su tiempo, y en cualquier caso es una descripción incompleta de la evolución. La hipótesis de Darwin, presentada junto con Wallace, que llegó a las mismas conclusiones independientemente, es que pequeños cambios heredables en los seres vivos y la selección son los dos hechos que provocan el cambio en la naturaleza y la generación de nuevas especies; pero Darwin desconocía cuál es la base de la herencia, pensaba que los rasgos de un ser vivo eran como un fluido, y que los fluidos de los dos padres se mezclaban en la descendencia. Esta hipótesis tenía el problema de que al cabo de cierto tiempo, una población tendría los mismos rasgos intermedios.

Fue Lendel quien descubrió que los caracteres se heredaban de forma discreta, y que se tomaban del padre o de la madre, dependiendo de su carácter dominante o recesivo. A estos caracteres que podían tomar diferentes valores se les llamaron genes, y a los valores que podían tomar se les denominó aleles. En realidad, las teorías de Lendel, que trabajó en total aislamiento, se olvidaron y no se volvieron a redes cubrir hasta principios del siglo XX.

Además, hasta 1930 el geneticista inglés Robert Aylmer no relacionó ambas teorías, demostrando que los genes mendelianos eran los que proporcionaban el mecanismo necesario para la evolución.

Por la misma época, el biólogo alemán Walther Flemming describió los cromosomas, como ciertos filamentos en los que se agregaba la cromatina del núcleo celular durante la división; poco más adelante se descubrió que las células de cada especie viviente tenían un número fijo y característico de cromosomas. Y no fue hasta los años 50, cuando Watson y Crick descubrieron que la base molecular de los genes está en el ADN, ácido desoxirribonucleico. Los cromosomas están compuestos de ADN, y por tanto los genes están en los cromosomas.

La macromolécula de ADN está compuesta por bases púricas y pirimidínicas, la adenina, citosina, guanina y timina. La combinación y la secuencia de estas bases forma el código genético, único para cada ser vivo. Grupos de tres bases forman un codon, y cada codon codifica un aminoácido; el código genético codifica todas las proteínas que forman parte de un ser vivo. Mientras que al código genético se le llama genotipo, al cuerpo que construyen esas proteínas, modificado por la presión ambiental, la historia vital, y otros mecanismos dentro del cromosoma, se llama gen.

No toda la cadena de ADN codifica proteínas, es decir, no todos son genes; las zonas que codifican proteínas se llaman intrones, las zonas que no lo hacen, exones. La cantidad de ADN basura aumenta desde los seres vivos más simples, como las bacterias, donde no hay nada, hasta los seres humanos, donde gran cantidad del ADN no codifica.

Un gen comienza con el sitio tres o aceptor y termina con el sitio cinco o donante. Proyectos como el del Genoma Humano tratan de identificar cuáles son estos genes, sus posiciones, y sus posibles alteraciones, que habitualmente conducen a enfermedades.

Todos estos hechos forman hoy en día la teoría del neo-darwinismo, que afirma que la historia de la mayoría de la vida está causada por una serie de procesos que actúan en las poblaciones y dentro de las poblaciones: reproducción, mutación, competición y selección. La evolución se puede definir entonces como cambios en el conjunto genético de una población.

Un tema polémico, con opiniones variadas dependiendo de que se trate de informáticos evolutivos o de biólogos o geneticistas, es si la evolución optimiza o no. Según los informáticos evolutivos, la evolución optimiza, puesto que va creando seres cada vez más perfectos, cuya culminación es el hombre; además, indicios de esta optimización se encuentran en el organismo de los animales, desde el tamaño y tasa de ramificación de las arterias, diseñada para maximizar flujo, hasta el metabolismo, que optimiza la cantidad de energía extraída de los alimentos.

Sin embargo, los geneticistas y biólogos evolutivos afirman que la evolución no optimiza, sino que adapta y optimiza localmente en el espacio y el tiempo; evolución no significa progreso. Un organismo más evolucionado puede estar en desventaja competitiva con uno de sus antepasados, si se colocan en el ambiente del último.

##### La informática evolutiva

Las primeras ideas, incluso antes del descubrimiento del ADN, vinieron de Von Neumann, uno de los mayores científicos de este siglo. Von Neumann afirmó que la vida debía de estar apoyada por un código que a la vez describiera como se puede construir un ser vivo, y tal que ese ser creado fuera capaz de autorreproducirse, por tanto, un autómata o máquina autorreproductiva tendría que ser capaz, aparte de contener las instrucciones para hacerlo, de copiar tales instrucciones a su descendencia, y así continuar con este proceso.

 Sin embargo, no fue hasta mediados de los años cincuenta, cuando el rompecabezas de la evolución se había prácticamente completado, cuando Box comenzó a pensar en imitarla, para mejorar procesos industriales. La técnica de Box, denominada EVOP (Operación Evolucionaria), consistía en elegir una serie de variables que regían un proceso industrial. Sobre esas variables se creaban pequeñas variaciones que formaban un hipercubo, variando el valor de las variables una cantidad fija.

Se probaba entonces con cada una de las esquinas del hipercubo durante un tiempo, y al final del periodo de pruebas, un comité humano decidía sobre la calidad del resultado. Es decir, se estaba aplicando mutación y selección a los valores de las variables, con el objeto de mejorar la calidad del proceso. Este procedimiento se aplicó con éxito a algunas industrias químicas.

Un poco más adelante, en 1958, Friedberg y sus colaboradores pensaron en mejorar usando técnicas evolutivas la operación de un programa. Para ello diseñaron un código máquina de 14 bits (2 para el código de operación, y 6 para los datos y/o instrucciones); cada programa, tenía 64 instrucciones. Un programa llamado Herman, ejecutaba los programas creados, y otro programa, el profesor, le mandaba a Herman ejecutar otros programas y ver si los programas ejecutados habían realizado su tarea o no. La tarea consistía en leer unas entradas, situadas en una posición de memoria, y debían depositar el resultado en otra posición de memoria, que era examinada al terminarse de ejecutar la última instrucción.

Para hacer evolucionar los programas, Friedberg hizo que en cada posición de memoria hubiera dos alternativas; para cambiar un programa, alternaba las dos instrucciones (que eran una especie de aleles), o bien reemplazaba una de las dos instrucciones con una totalmente aleatoria.

En realidad, lo que estaba haciendo es usar mutación para generar nuevos programas; al parecer, no tuvo más éxito que si hubiera buscado aleatoriamente un programa que hiciera la misma tarea. El problema es que la mutación sola, sin ayuda de la selección, hace que la búsqueda sea prácticamente una búsqueda aleatoria.

Más o menos simultáneamente, Bremmerman trató de usar la evolución para entender los procesos de pensamiento creativo y aprendizaje, y empezó a considerar la evolución como un proceso de aprendizaje. Para resolver un problema, codificaba las variables del problema en una cadena binaria de ceros y unos, y sometía la cadena a mutación, cambiando un bit de cada vez; de esta forma, estableció que la tasa ideal de mutación debía de ser tal que se cambiara un bit cada vez. Bremmerman trató de resolver problemas de minimización de funciones, aunque no está muy claro qué tipo de selección usó, y el tamaño y tipo de la población. En todo caso, se llegaba a un punto, la trampa de Bremmerman, en el cual la solución no mejoraba, y luego en intentos sucesivos trató de añadir entrecruzamiento entre soluciones, pero tampoco obtuvo buenos resultados.

Una vez más, el simple uso de operadores que creen diversidad no es suficiente para dirigir la búsqueda genética hacia la solución correcta, y corresponde a un concepto de la evolución darwiniano clásico: por mutación, se puede mejorar a un individuo; en realidad, la evolución actúa al nivel de población, es decir que actúa de manera global.

El primer uso de procedimientos evolutivos en inteligencia artificial se debe a Reed, Toombs y Baricelli, que trataron de hacer evolucionar un tahúr que jugaba a un juego de cartas simplificado. Las estrategias de juego consistían en una serie de 4 probabilidades de apuesta alta o baja con una mano alta o baja, con cuatro parámetros de mutación asociados. Se mantenía una población de 50 individuos, y aparte de la mutación, había intercambio de probabilidades entre dos padres. Es de suponer que los perdedores se eliminaban de la población (tirándolos por la borda). Aparte de, probablemente, crear buenas estrategias, llegaron a la conclusión de que el entrecruzamiento no aportaba mucho a la búsqueda.

Los intentos posteriores, ya realizados en los años 60, ya corresponden a los algoritmos evolutivos modernos, y se han seguido investigando hasta nuestros días. Algunos de ellos son simultáneos a los algoritmos genéticos, pero se desarrollaron sin conocimiento unos de otros. Uno de ellos, la programación evolutiva de Fogel, se inició como un intento de usar la evolución para crear máquinas inteligentes, que pudieran prever su entorno y reaccionar adecuadamente a él. Para simular una máquina pensante, se utilizó un autómata celular. Un autómata celular es un conjunto de estados y reglas de transición entre ellos, de forma que, al recibir una entrada, cambia o no de estado y produce una salida.

Fogel trataba de hacer aprender a estos autómatas a encontrar regularidades en los símbolos que se le iban enviando. Como método de aprendizaje, usó un algoritmo evolutivo: una población de diferentes autómatas competía para hallar la mejor solución, es decir, predecir cual iba a ser el siguiente símbolo de la secuencia con un mínimo de errores; los peores 50% eran eliminados cada generación, y sustituidos por otros autómatas resultantes de una mutación de los existentes.

De esta forma, se lograron hacer evolucionar autómatas que predecían algunos números primos (por ejemplo, uno, cuando se le daban los números más altos, respondía siempre que no era primo; la mayoría de los números mayores de 100 son no primos). En cualquier caso, estos primeros experimentos demostraron el potencial de la evolución como método de búsqueda de soluciones novedosas.

Más o menos a mediados de los años 60, Rechenberg y Schwefel describieron las estrategias de evolución. Las estrategias de evolución son métodos de optimización paramétricos, que trabajan sobre poblaciones de cromosomas compuestos por números reales. Hay diversos tipos de estrategias de evolución, que se verán más adelante; pero en la más común, se crean nuevos individuos de la población añadiendo un vector de mutación a los cromosomas existentes en la población, en cada generación, se elimina un porcentaje de la población (especificado por los parámetros), y los restantes generan la población total, mediante mutación y crossover. La magnitud del vector mutación se calcula adaptativamente.

John Holland desde pequeño, se preguntaba cómo logra la naturaleza, crear seres cada vez más perfectos. Lo curioso era que todo se lleva a cabo basándose en interacciones locales entre individuos, y entre estos y lo que les rodea. No sabía la respuesta, pero tenía una cierta idea de como hallarla: tratando de hacer pequeños modelos de la naturaleza, que tuvieran alguna de sus características, y ver cómo funcionaban, para luego extrapolar sus conclusiones a la totalidad. De hecho, ya de pequeño hacía simulaciones de batallas célebres con todos sus elementos: copiaba mapas y los cubría luego de pequeños ejércitos que se enfrentaban entre sí.

En los años 50 entró en contacto con los primeros ordenadores, donde pudo llevar a cabo algunas de sus ideas, aunque no se encontró con un ambiente intelectual fértil para propagarlas. Fue a principios de los 60, en la Universidad de Michigan en Ann Arbor,donde, dentro del grupo lógica de computadoras, sus ideas comenzaron a desarrollarse y a dar frutos. Además, fue leyendo un libro escrito por un biólogo evolucionista llamado, R. A. Fisher, titulado “La teoría genética de la selección natural”, como comenzó a descubrir los medios de llevar a cabo sus propósitos de comprensión de la naturaleza.

De ese libro aprendió que la evolución era una forma de adaptación más potente que el simple aprendizaje, y tomó la decisión de aplicar estas ideas para desarrollar programas bien adaptados para un fin determinado.

En esa universidad, Holland impartía un curso titulado: “Teoría de sistemas adaptativos”, dentro de este curso, y con una participación activa por parte de sus estudiantes, fue donde se crearon las ideas que más tarde se convertirían en los algoritmos genéticos.

Por tanto, cuando Holland se enfrentó a los algoritmos genéticos, los objetivos de su investigación fueron dos:

* La imitación de los procesos adaptativos de los sistemas naturales, y
* El diseño sistemas artificiales (normalmente programas) que retengan los mecanismos importantes de los sistemas naturales.

Unos 15 años más adelante, David Goldberg, que en la actualidad es el maestro de los algoritmos genéticos, conoció a Holland, y se convirtió en su estudiante. Goldberg era un ingeniero industrial trabajando en diseño de pipelines, y fue uno de los primeros que trató de aplicar los algoritmos genéticos a problemas industriales. Aunque Holland trató de disuadirle, porque pensaba que el problema era excesivamente complicado como para aplicarle algoritmos genéticos, Goldberg consiguió lo que quería, escribiendo un algoritmo genético en un ordenador personal Apple II. Estas y otras aplicaciones creadas por estudiantes de Holland convirtieron a los algoritmos genéticos en un campo con base suficiente aceptado para celebrar la primera conferencia en 1985, ICGA´85, la cual se sigue celebrando bianualmente.

* 1. **Ubicación del problema**

El problema consiste en desarrollar un algoritmo genético para la simulación del crecimiento de una población, basándose en los índices de natalidad, mortalidad, y las migraciones, y luego observar su comportamiento a través del tiempo.

Además se ha de optimizar el crecimiento de una población, a través del tiempo, mediante algoritmos genéticos, y finalmente se realizará una selección de un modelo más complejo, basándose en el paradigma de los algoritmos genéticos.

* 1. **Limitaciones y alcance del tema.**

Las técnicas de simulación convencionales, son de gran utilidad; pero aplicando los algoritmos genéticos, se puede realizar la simulación del crecimiento de una población demográfica, aplicando el proceso de selección natural y otros operadores genéticos, donde cada individuo forma parte de la población, y por lo tanto, pertenece a la solución.

* 1. **Objetivos: generales y específicos.**

El objetivo general es desarrollar un modelo de crecimiento poblacional, utilizando el paradigma de los algoritmos genéticos como una herramienta robusta para la simulación y la optimización y se tratará de estudiar el comportamiento de una población mediante un modelo basado en el proceso de la evolución.

Como objetivos específicos se tienen los siguientes:

* Observar el comportamiento de los índices poblacionales, así como la posibilidad de una tendencia, o convergencia en un modelo desarrollado con algoritmos genéticos.
* Realizar variaciones en los algoritmos genéticos de para adaptarlos al problema planteado.
* Implantar una aplicación robusta que aplique los algoritmos genéticos para modelar el crecimiento de una población, y sus interacciones.
* Comparar los modelos de crecimiento poblacional convencionales, con un modelo basado en algoritmos genéticos.