



Algoritmos Bioinspirados

Unidad II : Cómputo Evolutivo - Algoritmos Genéticos

Programa: Ingeniería en Inteligencia Artificial

Introducción

La teoría de la evolución propuesta por Charles Darwin en el “Origen de las Especies” [1] introdujo el principio de la **supervivencia del más apto**.

- ❖ Entorno con recursos limitados e individuos que compite por ellos.
- ❖ Los mejor adaptados (aptos) tienen mayor chance de sobrevivir y reproducirse.
- ❖ Al reproducirse la combinación de las buenas características de los padres puede producir descendencia mejor adaptada que sus padres.
- ❖ Tras varias generaciones la población se vuelve mejor adaptada a su entorno.

[1]. Darwin, C. & Kebler, L. (1859) On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life. London: J. Murray.

Genética

La ciencia que estudia los mecanismos responsables de las similitudes y diferencias en una especie.

De la palabra griega *genēsis* que significa *relativo al origen*.

Los términos involucrados en el trasfondo biológico de las especies son las siguientes:

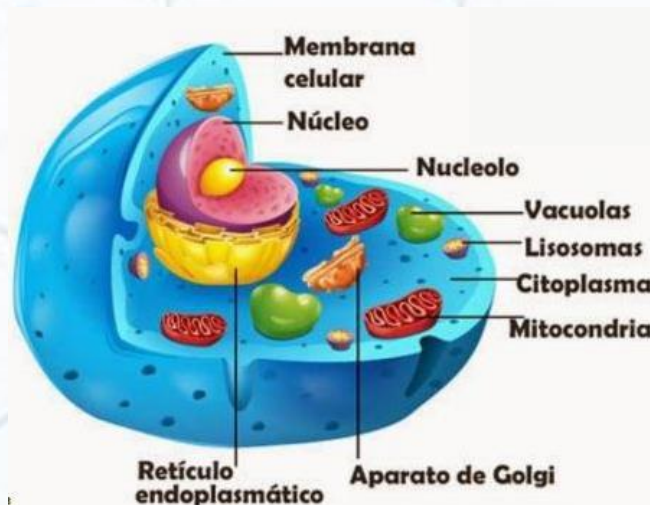
- ❖ La célula
- ❖ Genética
- ❖ Selección natural
- ❖ El cromosoma
- ❖ Reproducción

La célula

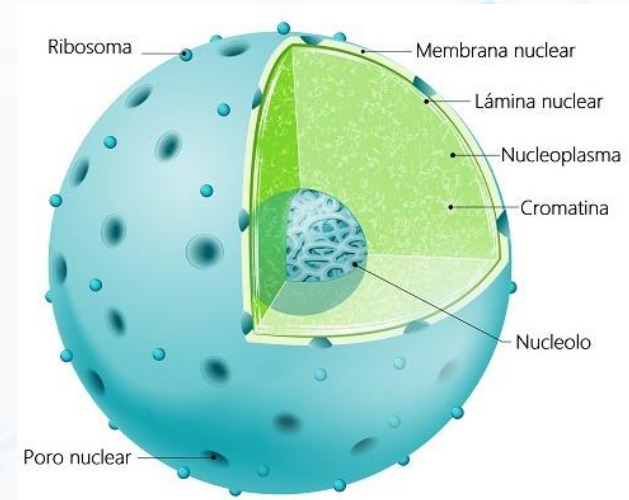
Unidad estructural y funcional de los organismos vivos.

El centro de todo esto es el núcleo celular. La información genética está contenida en el núcleo celular.

Anatomía de una célula

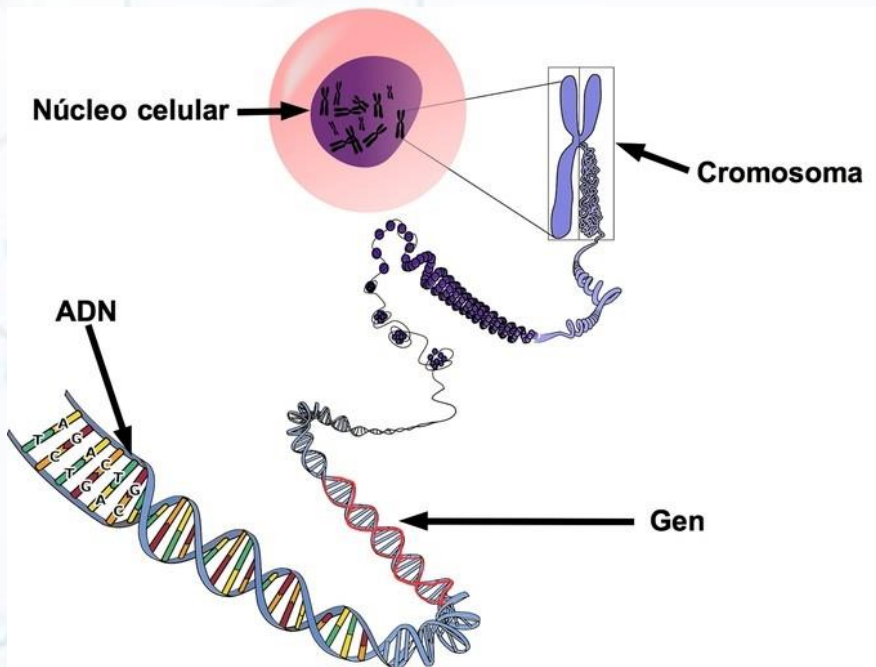


Núcleo de una célula



Cromosomas

Estructura formada por el ADN (ácido desoxirribonucleico) y proteínas que almacena la información genética de los seres vivos.



- ❖ Los cromosomas se dividen en varias partes llamadas genes.
- ❖ Los genes codifican las características de un individuo.

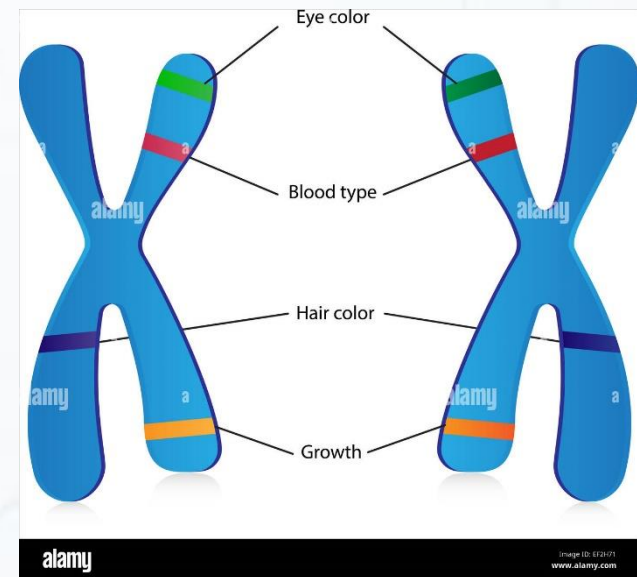
Genética

Genotipo: es la combinación de los genes de un individuo.

Fenotipo: describe los aspectos físicos observables del individuo.

Un alelo es uno de los posibles valores que puede tener un gen, por lo que su relación con un gen es como la de un valor específico con una variable en matemáticas.

Alelos



Genética

En formas de vida superiores, los cromosomas contienen dos conjuntos de genes. Esto se conoce como diploides.

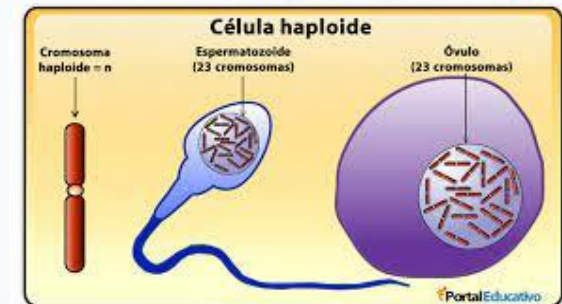


Seres humanos
46 cromosomas
organizados en 23
pares diploides

La fusión de dichos gametos forma una célula diploide, el cigoto. En el cigoto, cada par de cromosomas está formado por una mitad paterna y otra materna.

Gametos

Espermatozoides y
óvulos, son haploide



- ❖ En computación evolutiva, la combinación de características de dos individuos se le llama Crossover.
- ❖ Esto no es análogo al funcionamiento de los organismos diploides, donde el entrecruzamiento no se da durante el apareamiento y la fertilización, sino que ocurre durante la formación de gametos, un proceso llamado meiosis.
- ❖ Entre dos valores del mismo par de genes, el dominante determinará el fenotipo mientras que el otro, llamado recesivo, seguirá presente y podrá transmitirse a la descendencia .
- ❖ La mayoría de los Algoritmos Genéticos (AG) se concentran en los cromosomas haploides porque son mucho más simples de construir, donde solo se almacena un conjunto de cada gen, por lo que se evita el proceso de determinar qué alelo debe ser dominante y cuál debe ser recesivo.

Selección

Todas las variaciones (mutación y recombinación) ocurren a nivel genotípico, mientras que la selección se basa en el desempeño real en un ambiente dado, es decir, a nivel fenotípico.

- ❖ El genoma del nuevo individuo no es idéntico al de sus padres.
- ❖ Se crean variaciones de genotipo, que a su vez se traducen en variaciones de fenotipo y, por lo tanto, están sujetas a selección.
- ❖ Por tanto, los genes también se ven sujetos al juego de supervivencia y reproducción .
- ❖ Algunos genetistas argumentan que, en lugar de pensar en poblaciones de individuos, deberíamos pensar en una reserva de genes que contiene genes que compiten y se replican con el tiempo, siendo evaluados a medida que se repiten en diferentes individuos .



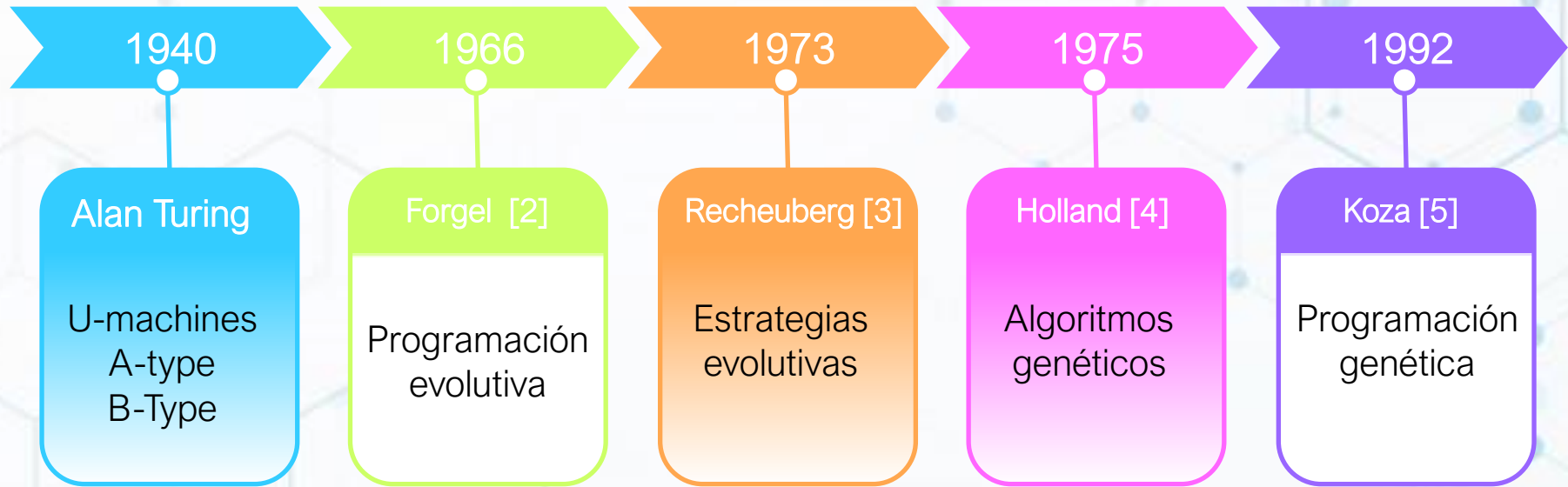
Evolución y AG

La necesidad de nuevas soluciones ha llevado al estudio de la naturaleza en busca de inspiración.

Dos grandes fuentes de inspiración han sido el cerebro humano y el proceso evolutivo.

Evolución Natural	Algoritmo Genético
Cromosoma	Cadena
Gen	Característica o rasgo
Alelo	Valor del rasgo
Locus	Posición en la cadena

Breve Reseña Histórica



Las diferencias básicas entre los diferentes paradigmas radican en los esquemas de representación, los operadores de reproducción y los métodos de selección. Hoy en su conjunto se conocen como **Computación Evolutiva**.

Computación Evolutiva (CE)

- ❖ Es un área de investigación dentro de la informática. Como sugiere el nombre, es un sabor especial de la informática, que se inspira en el proceso de evolución natural.
- ❖ Métodos que abstraen principios evolutivos en algoritmos que pueden usarse para buscar soluciones óptimas a un problema.

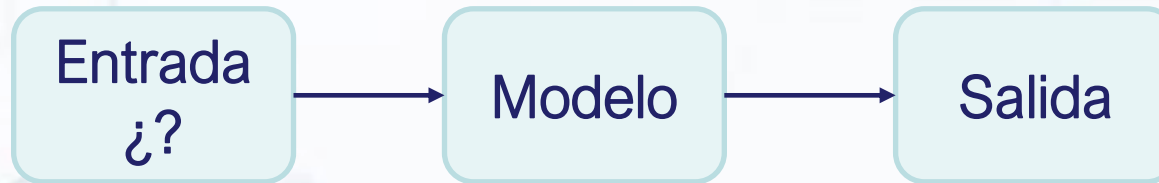
Analogías entre Evolución y CE2

Entorno	↔	Problema
Individuos	↔	Candidatos a solución
Aptitud	↔	Calidad de la solución

Espacio de búsqueda

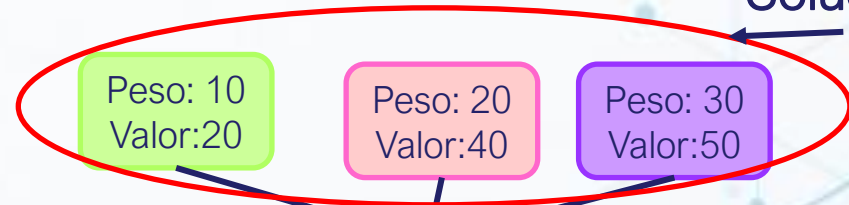
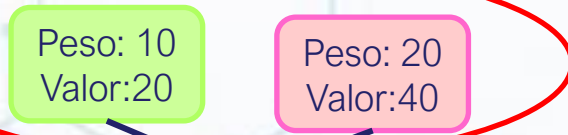
El conjunto de soluciones entre las que reside la solución deseada

Optimización



Solución

Solución



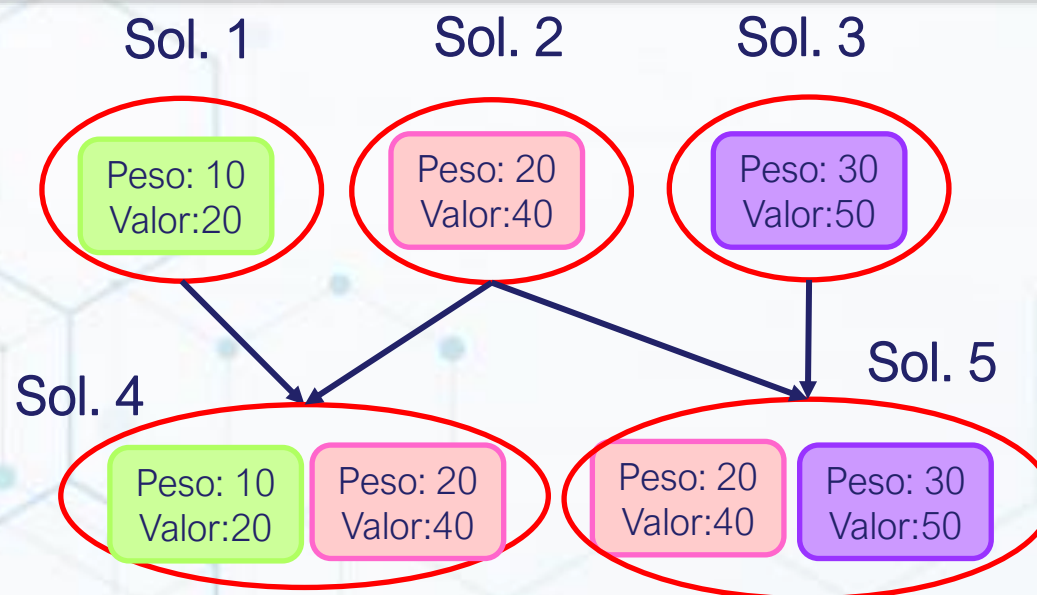
Espacio de Búsqueda

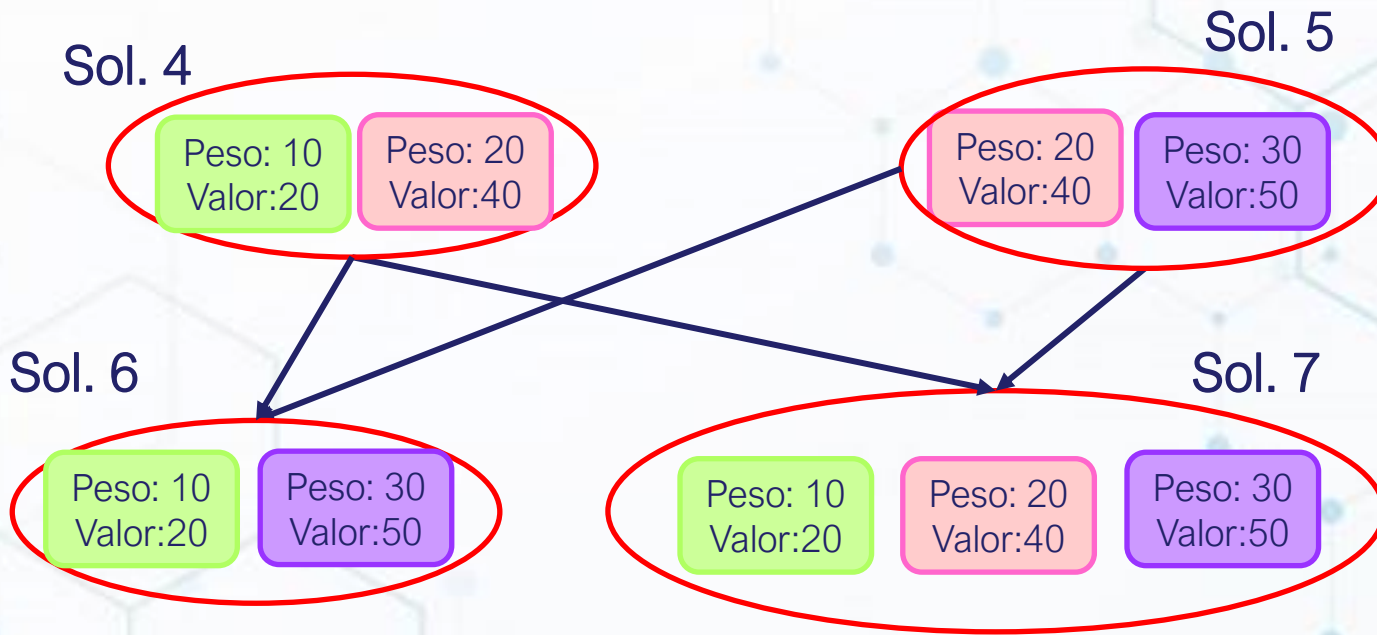
- ❖ Cada punto del espacio de búsqueda representa una posible solución.
- ❖ Cada posible solución puede ser "marcada" por su valor de aptitud, dependiendo de la definición del problema.
- ❖ Con Algoritmo Genético se busca la mejor solución entre un número de posibles soluciones representadas por un punto en el espacio de búsqueda



Espacio de Búsqueda

- ❖ La reserva genética de una población dada contiene una solución (mejor solución)
- ❖ La solución no está activa porque se encuentra distribuida entre varios sujetos de la población.
- ❖ La interacción entre varios genomas puede guiarnos a dicha solución.





Algoritmos Genéticos

- ❖ Algoritmos estocásticos: la aleatoriedad juega un papel importante.
- ❖ Los procesos de selección y reproducción necesitan de funciones aleatorias.
- ❖ Siempre se considera una población de soluciones.
- ❖ Robustez: capacidad de desempeñarse bien de forma consistente en un amplio rango de problemas

Un algoritmo genético es un método de resolución de problemas que utiliza la genética como modelo de resolución de problemas.

Es una técnica de búsqueda para encontrar soluciones aproximadas a problemas de optimización y búsqueda.



Esquema General de un AG



[1]. A. E. Eiben and James E. Smith. 2015. Introduction to Evolutionary Computing (2nd. ed.). Springer Publishing Company, Incorporated.

Algoritmos Genéticos – Principales componentes

- ❖ Representación del cromosoma
- ❖ Función de evaluación.
- ❖ Población
- ❖ Mecanismo de selección de los padres
- ❖ Mecanismo de los operadores de cruce y mutación
- ❖ Mecanismo de reemplazo

Individuos

Un individuo es una sola solución. Un individuo agrupa dos formas de soluciones como se indica a continuación:

1. El cromosoma, que es la información "genética" sin procesar (genotipo) que trata el AG.
2. El fenotipo, que es lo expresivo del cromosoma en los términos del modelo.

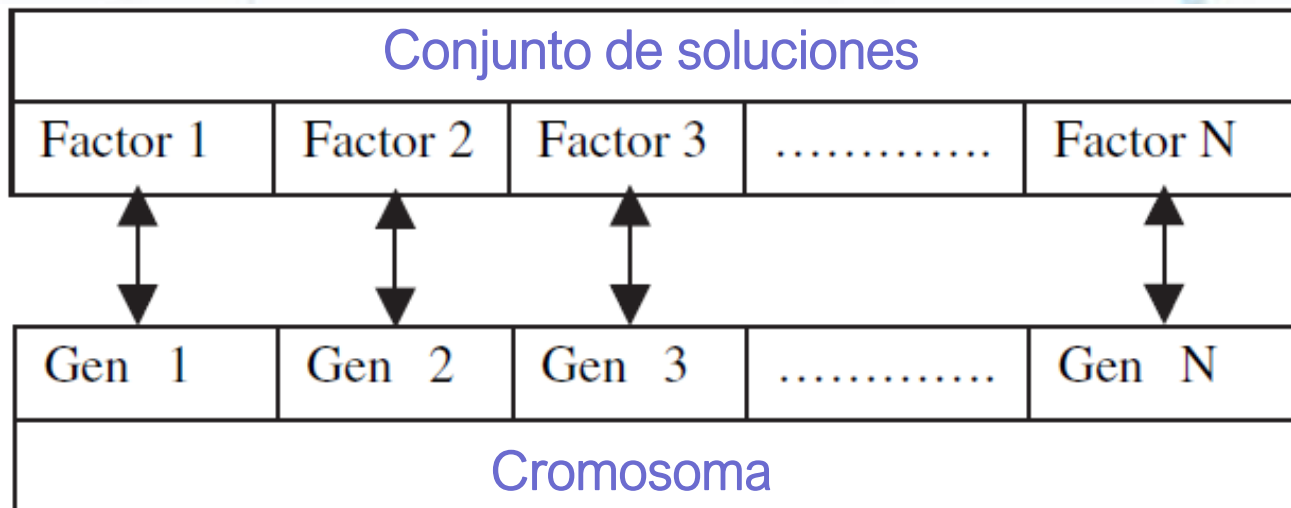
Problema de la mochila



Individuos

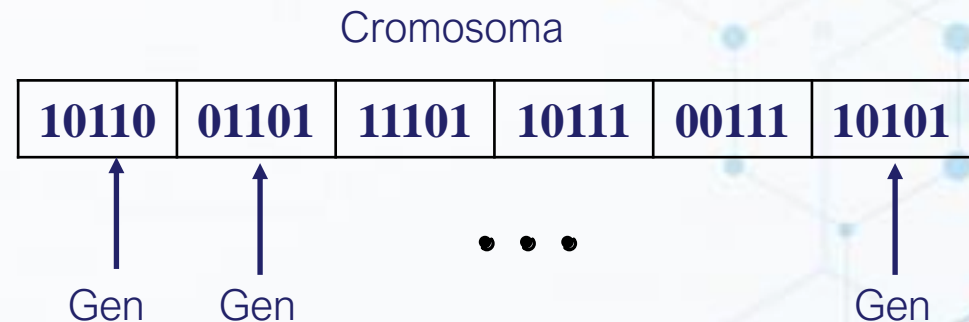
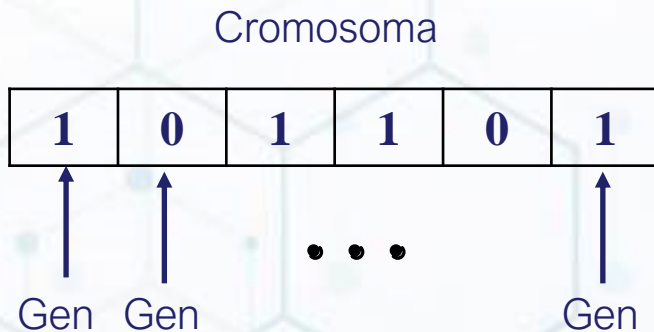
El primer paso para implementar un AG es la representación

Se deben especificar y almacenar las posibles soluciones de manera que puedan ser manipuladas por una computadora.



Gen

- ❖ Es la unidad más básica en un AG.
- ❖ Los cromosomas se encuentran conformados por genes.
- ❖ Cada gen representa una variable del problema



Codificación

Es el proceso para representar genes individuales. El proceso se puede realizar usando bits, números o caracteres.

La codificación depende mucho del problema a resolver.

Codificación Binaria

La más común

Cromosoma 1	1100101110
Cromosoma 2	1001010111

Codificación Enteros

Cromosoma 1	182369547
Cromosoma 2	271935468

Codificación Hexadecimal

Cromosoma 1	9A37D1
Cromosoma 2	5F1C6A

Codificación Valores Reales

Cromosoma 1	1.3	2.5	3.3	7.1
Cromosoma 2	2.1	6.4	2.7	1.4

Mapeo

Proceso entre el genotipo y el fenotipo que permite convertir las soluciones del modelo en un formato con el que el AG pueda trabajar, así como el proceso inverso

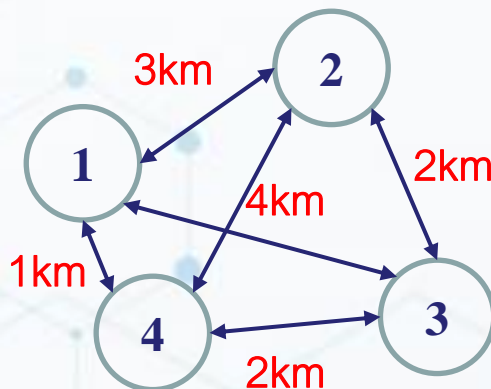
1	0	1	1	0	1
---	---	---	---	---	---

 $= 45$

$$1 * 2^5 + 0 * 2^4 + 1 * 2^3 + 1 * 2^2 + 0 * 2^1 + 1 * 2^0$$

$$= 32 + 8 + 4 + 1 = 45$$

1	4	2	3
---	---	---	---



Peso: 20
Valor:40

Peso: 10
Valor:20

Peso: 30
Valor:50

1	0	1
---	---	---

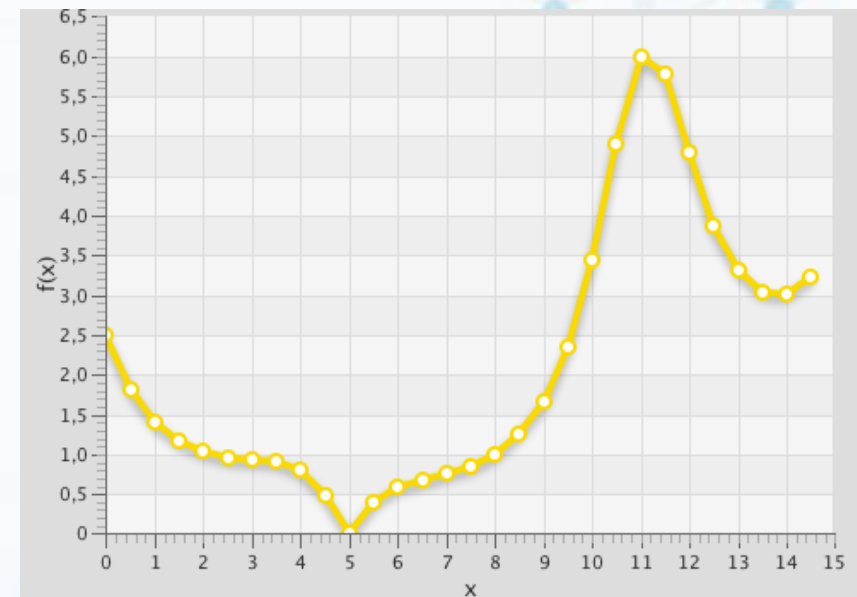
Función de aptitud (*fitness*)

La aptitud de un individuo en un algoritmo genético es el valor de la función objetivo para su fenotipo.

El cromosoma debe ser decodificado y evaluado en la función objetivo.

$$f(x) = \text{ABS} \left| \frac{x - 5}{2 + \text{Sen}(x)} \right|$$

	Genotipo	Fenotipo	$f(x)$
Individuo 1	0011	3	0.93
Individuo 2	1111	15	3.77
Individuo 3	0001	1	1.40
Individuo 4	1000	8	1



- ❖ Representa los requisitos a los que la población debe adaptarse.
- ❖ Constituye la base para la selección y, por lo tanto, facilita las mejoras.
- ❖ Una función o procedimiento que asigna una medida de calidad a los genotipos.
- ❖ Se compone de la representación inversa (decodificar para crear el fenotipo correspondiente) seguida de una medida de calidad en el espacio del fenotipo.



Población

- ❖ Es un multiconjunto de genotipos.
- ❖ Los individuos son objetos estáticos que no cambian ni se adaptan; es la población la que lo hace. La población forma la unidad de evolución.

Población	Individuo 1	0011
	Individuo 2	1111
	Individuo 3	0001
	Individuo 4	1000

En un AG los 2 aspectos más importantes a definir desde el punto de vista de la población:

1. La generación de la población inicial.
2. El tamaño de la población.

- ❖ Generalmente, la población inicial se genera de forma aleatoria.
- ❖ Alguna heurística podría ser usada para la población inicial lo que permitiría una convergencia más rápida, pero se corre el riesgo de falta de diversidad que no permita explorar todo el espacio de soluciones.
- ❖ Entre más grande el tamaño de la población más fácil explorar el espacio de soluciones.
- ❖ De acuerdo con [2], el tiempo requerido por un AG converge a $O(n \log n)$ donde n es el tamaño de la población. Por lo que una población grande puede ser útil, pero requiere mayor costo computacional.

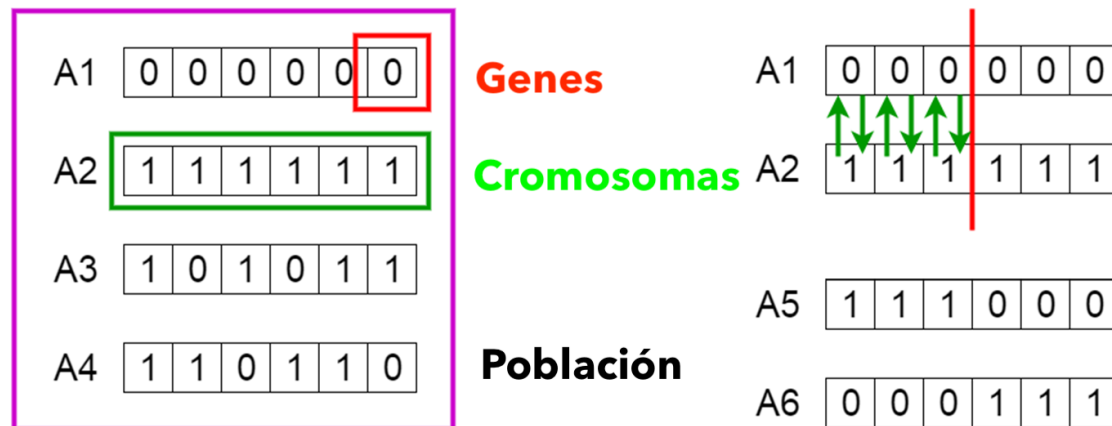
[2]. S. N. Sivanandam and S. N. Deepa. 2010. Introduction to Genetic Algorithms (1st. ed.). Springer Publishing Company, Incorporated.



- ❖ Una población inicial de 100 individuos es común, pero dependiendo de los recursos computacionales disponibles
- ❖ Generalmente el tamaño de la población es una constante y no cambia durante el proceso evolutivo.

Reproducción

- ❖ Selección de los padres
- ❖ Creación de nuevos individuos
- ❖ Reemplazo de individuos en la población.



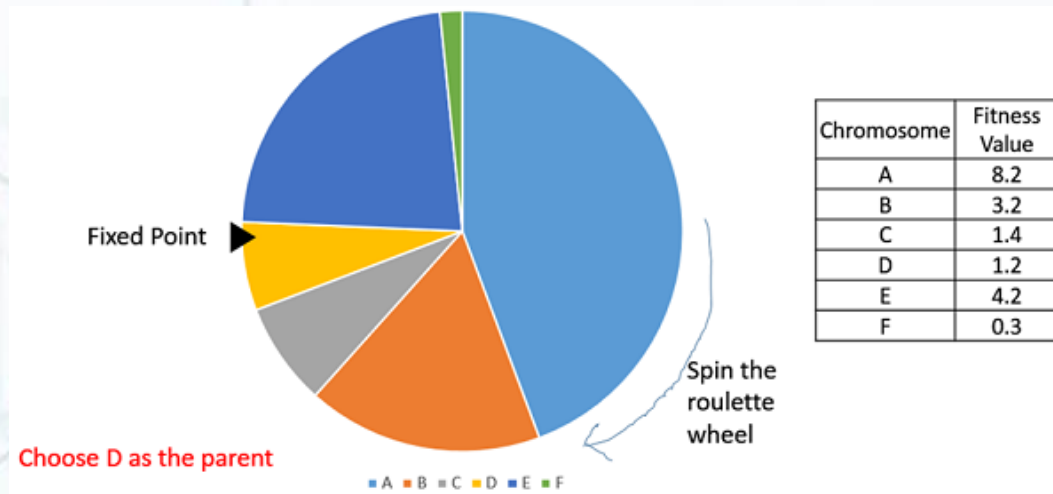
Selección

- ❖ Es el proceso de elegir dos padres de la población para cruzar y crear nuevos individuos
- ❖ La presión de selección se define como el grado al que se favorece los mejores individuos para ser seleccionados como padres.
- ❖ Por lo general, podemos distinguir dos tipos de esquema de selección:
 - ✓ Selección proporcional.
 - ✓ Selección basada en ordinales.

Selección Ruleta

- ❖ La selección de ruleta es una de las técnicas tradicionales de selección de AG.
- ❖ Un operador de selección proporcional donde se selecciona un individuo con una probabilidad proporcional a la aptitud.

Se llama de la ruleta porque imita el proceso de girar una ruleta y seleccionar un punto al azar en esta.



$$f(x) = 4x_1 + 5x_2 + 6x_3 + 3x_4$$

	x_1	x_2	x_3	x_4	$f(x)$	P_{sel}	P_{sel_acum}
Individuo 1	0	1	1	0	11	0.29730	0.29730
Individuo 2	1	0	1	0	10	0.27027	0.56757
Individuo 3	0	1	0	0	5	0.13514	0.70270
Individuo 4	0	1	1	0	11	0.29730	1.00000

$$P_{sel} = \frac{f(x)}{Total}$$

Total: 37

Suma total de las aptitudes de todos los individuos en la población

$$\text{Individuo 1} \rightarrow P_{sel} = \frac{11}{37} = 0.29730$$

$$\text{Individuo 2} \rightarrow P_{sel} = \frac{10}{37} = 0.27027$$

$$\text{Individuo 3} \rightarrow P_{sel} = \frac{5}{37} = 0.13514$$

La P_{sel_acum} es la suma acumulada de P_{sel} de cada individuo.

$$f(x) = 4x_1 + 5x_2 + 6x_3 + 3x_4$$

	x_1	x_2	x_3	x_4	$f(x)$	P_{sel}	P_{sel_acum}
Individuo 1	0	1	1	0	11	0.29730	0.29730
Individuo 2	1	0	1	0	10	0.27027	0.56757
Individuo 3	0	1	0	0	5	0.13514	0.70270
Individuo 4	0	1	1	0	11	0.29730	1.00000

Total: 37

Valor aleatorio 1: 0.4771

Padre 1 → Individuo 2

Valor aleatorio 2: 0.9341

Padre 2 → Individuo 4

Buscar en P_{sel_acum} el valor igual o mayor al número aleatorio generado

- ❖ Es sencillo de implementar.
- ❖ Los individuos con más aptitud tienden a dominar la población en pocas generaciones
- ❖ La tasa de evolución depende de la varianza de la aptitud en la población.
- ❖ Un individuo apto contribuirá más al valor objetivo, pero si no lo supera, el siguiente cromosoma en la línea tiene una oportunidad y puede ser débil.

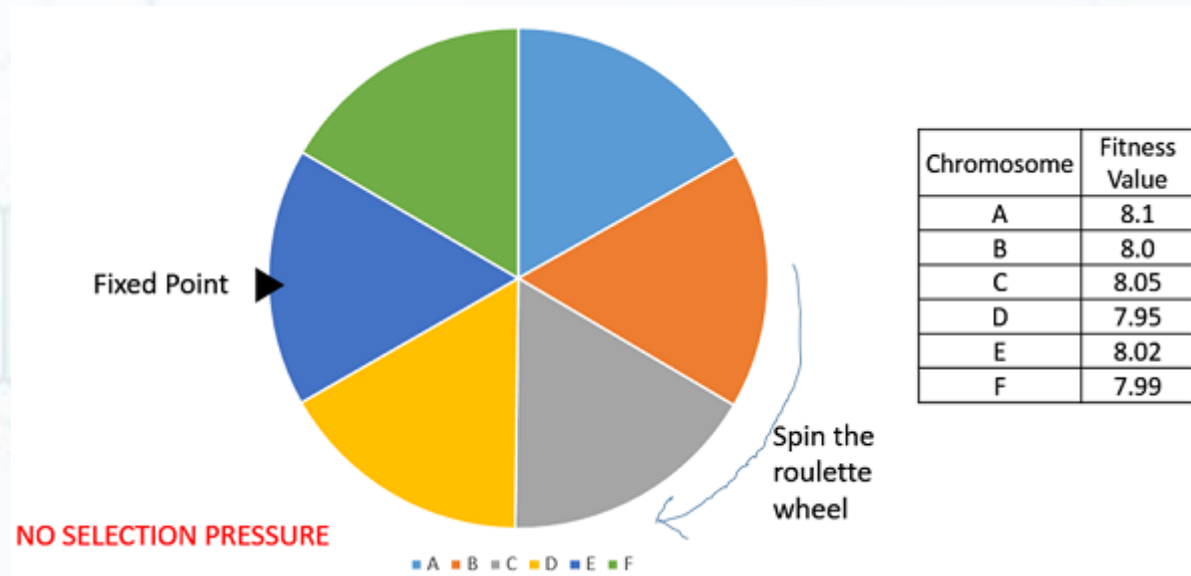
Selección Aleatoria

- ❖ Seleccionar al azar 2 individuos de la población para ser padres.
- ❖ No hay presión de selección hacia individuos más aptos y, por lo tanto, esta estrategia generalmente se evita.

Selección por Ranking

- ❖ La selección de rango ordena la población por su aptitud y cada individuo recibe una probabilidad de selección de acuerdo con su posición.
- ❖ El peor individuo tiene *ranking* 0, mientras que el mejor individuo tendrá *ranking* de $N - 1$ donde N es el número de individuos en la población.
- ❖ El mapeo del *ranking* a la probabilidad de selección se puede calcular de muchas maneras, por ejemplo, lineal o exponencialmente.

La selección por *ranking* es útil cuando el valor de aptitud de cada individuo es muy similar. Si en esta situación aplicáramos la selección de ruleta no existiría presión de selección.



Selección por Ranking

- ❖ Mantiene una presión de selección constante al ordenar la población con base en su aptitud.
- ❖ La probabilidad de selección lineal para cada individuo se calcula de acuerdo con la siguiente expresión:

$$P_{s_rank}(i) = \frac{(2 - s)}{N} + \frac{2i(s - 1)}{N(N - 1)}$$

Lineal.

Donde $1 < s \leq 2$ y N es el número de individuos en la población.

$$P_{s_rank}(i) = \frac{(2 - s)}{N} + \frac{2i(s - 1)}{N(N - 1)}$$

$$N = 4$$

	Aptitud $f(x)$	Ranking	P_{s_rank} ($s = 2$)	P_{s_rank} ($s = 1.5$)
Individuo 1	1	0	0	0.125
Individuo 3	2	1	0.17	0.208
Individuo 4	3	2	0.33	0.292
Individuo 2	5	3	0.50	0.375

Una vez calculada la probabilidad de selección (P_{s_rank}), se sigue el mismo procedimiento de selección que en la ruleta, es decir la diferencia de los métodos consiste en como se calcula la probabilidad de selección.

	Aptitud $f(x)$	Ranking	P_{s_rank} ($s = 2$)	$P_{s_rank_acum}$
Individuo 1	1	0	0	0
Individuo 3	2	1	0.17	0.17
Individuo 4	3	2	0.33	0.50
Individuo 2	5	3	0.50	1

Valor aleatorio 1: 0.15

Padre 1 → Individuo 3

Valor aleatorio 1: 0.57

Padre 1 → Individuo 2

Buscar en $P_{s_rank_acum}$ el valor igual o mayor al número aleatorio generado

Selección por Torneo

- ❖ Se seleccionan k individuos al azar y de estos se selecciona el que tenga mejor valor de aptitud para ser un padre.
- ❖ Se repite el proceso para seleccionar el otro padre.

	Aptitud ($f(x)$)
Individuo 1	7
Individuo 2	5
Individuo 3	2
Individuo 4	2
Individuo 5	9
Individuo 6	3
Individuo 7	6
Individuo 8	4

Padre 1

Seleccionar $k = 3$ individuos

Individuo 2

Individuo 5

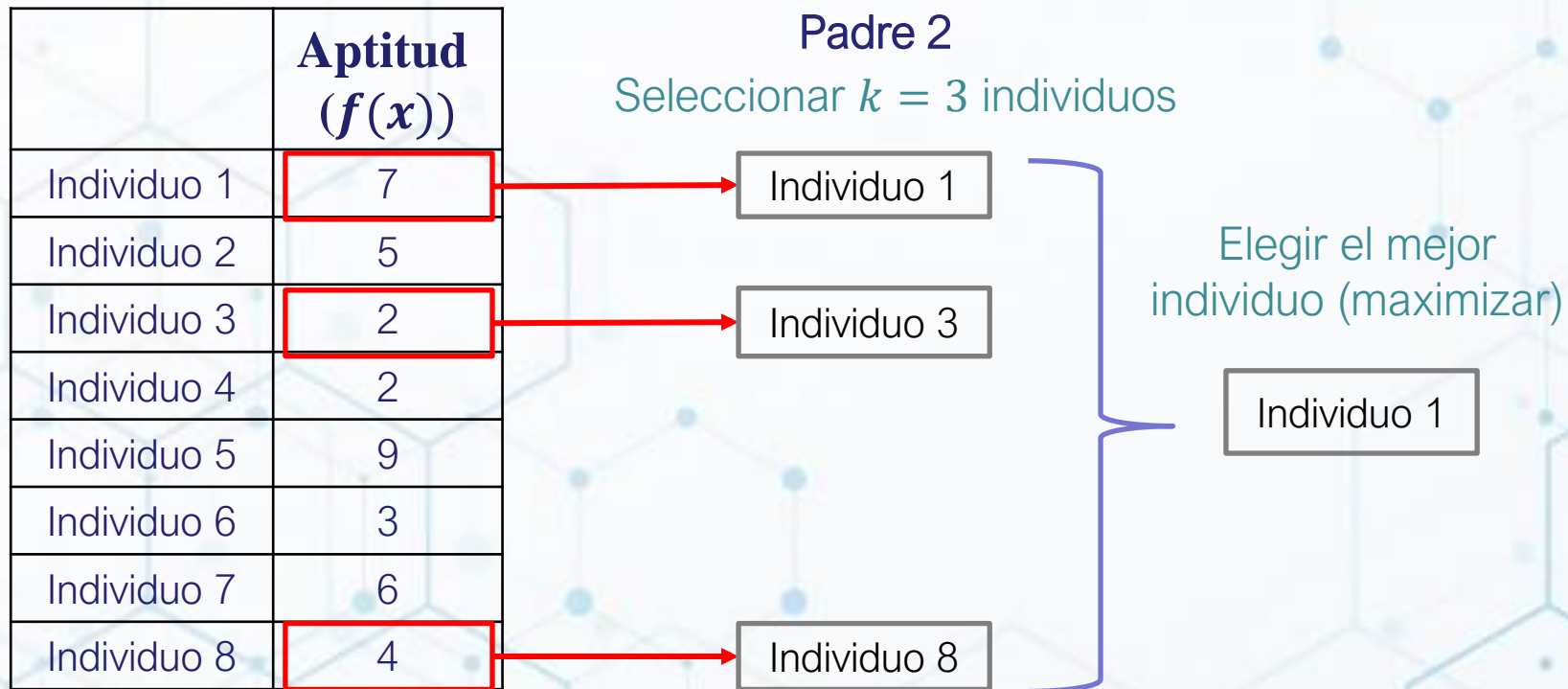
Individuo 7

Elegir el mejor
individuo (maximizar)

Individuo 5

Selección por Torneo

- ❖ Se seleccionan k individuos al azar y de estos se selecciona el que tenga mejor valor de aptitud para ser un padre.
- ❖ Se repite el proceso para seleccionar el otro padre.



Selección Elitista

- ❖ Método en el que un número pequeño de los mejores individuos pasan a la siguiente generación sin cambios.
- ❖ El número de individuos de élite que se seleccionan para la próxima generación se denomina *tasa de elitismo*.

	Aptitud ($f(x)$)
Individuo 1	7
Individuo 2	5
Individuo 3	8
Individuo 4	2
Individuo 5	9
Individuo 6	3
Individuo 7	8
Individuo 8	4

Pasan sin cambio a la siguiente generación

Individuo 3
Individuo 5
Individuo 7

El resto de la siguiente generación se genera usando cualquiera de los métodos de selección de padres.

Bibliografía

1. A. E. Eiben and James E. Smith. 2015. Introduction to Evolutionary Computing (2nd. ed.). Springer Publishing Company, Incorporated.
2. S. N. Sivanandam and S. N. Deepa. 2010. Introduction to Genetic Algorithms (1st. ed.). Springer Publishing Company, Incorporated.





¡ Gracias !

Thanks !

Obrigado

Xie xie ni

Domo arigatou

Спасибо

Merci

Grazie

Alfa Beta