



# Programación Evolutiva

## Tema 1: Introducción a la Computación Evolutiva

Carlos Cervigón. 2023-2024

- ❑ Orígenes
- ❑ Antecedentes Biológicos
- ❑ Esquema General
- ❑ Modelos sobre el Esquema General
- ❑ Relación con otras Áreas de Conocimiento
  - Inteligencia Artificial
  - Optimización Numérica

- ❑ Los Algoritmos Evolutivos (AEs) recogen un conjunto de modelos basados en la evolución de los seres vivos. Inspirados en la Naturaleza.
- ❑ La Computación Evolutiva surge a finales de los años 60 (John Holland).
- ❑ La aparición de computadores de grandes prestaciones y bajo coste a mediados de los 80 permite aplicar los AEs a la resolución de ciertos problemas que antes eran inabordables.
- ❑ En la actualidad funcionan con éxito en aplicaciones industriales de diversas áreas: Diseño de circuitos, Planificación de tareas, Cortado de patrones, etc.

## Antecedentes biológicos

- ❑ Los Algoritmos Evolutivos se basan en un modelo de evolución biológica natural que fue propuesto por primera vez por Charles Darwin.
- ❑ Teoría que explica el cambio adaptativo de las especies por el principio de la selección natural:
  - Favorece la supervivencia y evolución de aquellas especies que están mejor adaptadas a las condiciones de su entorno.
  - Aparición de variaciones pequeñas, aparentemente aleatorias y sin dirección en los fenotipos (características físicas de los individuos).



## Antecedentes biológicos

- ❑ **ADN** es una macromolécula. Las secuencias de las cuatro bases (Adenina, Timina, Citosina y Guanina) en la molécula de ADN determinan las características de cualquier organismo.
- ❑ **Gen** es una sección de ADN que codifica una cierta función bioquímica definida, normalmente la producción de una proteína.
- ❑ El ADN de un organismo puede contener desde una docena de genes (como un virus), hasta decenas de miles (humanos).
- ❑ **Cromosoma** cada una de las cadenas de ADN que se encuentra en el núcleo de las células. Son los responsables de la transmisión de información genética.
- ❑ **Locus** Región del cromosoma que ocupa el gen
- ❑ En cada determinado lugar pueden existir formas alternativas del gen. A estas formas alternativas se les llama **alelos**.

# Antecedentes biológicos



- Se denomina **cromosoma** a una de las cadenas de ADN que se encuentra en el núcleo de las células
- Un **gen** es una sección de ADN
- Alelo** es cada una de las formas del gen

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

Gen 1

Gen 2

Gen 3

3.1423	34.789	34.678	78.999
--------	--------	--------	--------

Gen 1

Gen 2

Gen 3

Gen 4

## Antecedentes biológicos

- ❑ El **genoma** es el conjunto total de genes-cromosomas que posee un organismo.
- ❑ Se denomina **individuo** (cromosoma) a un solo miembro de una población.
- ❑ Se denomina **población** a un grupo de individuos que pueden interactuar juntos.
- ❑ Se denomina **fenotipo** a los rasgos (observables) específicos de un individuo.
- ❑ Se denomina **genotipo** a la composición genética de un organismo (la información contenida en el genoma).

01101
11000
01000
10011

Genotype



Phenotype





## Estructura de un algoritmo evolutivo

- ❑ Un AE trabaja con una **población de individuos**, que representan soluciones candidatas a un problema.
- ❑ Esta población se somete a ciertas transformaciones y después a un **proceso de selección**, que favorece a los mejores, según su aptitud.
- ❑ Cada ciclo de selección y transformación constituye una **generación**.
- ❑ Se espera que después de cierto número de generaciones el mejor individuo de la población esté cerca de la solución buscada.

¡ convergencia !

# Estructura de un algoritmo evolutivo

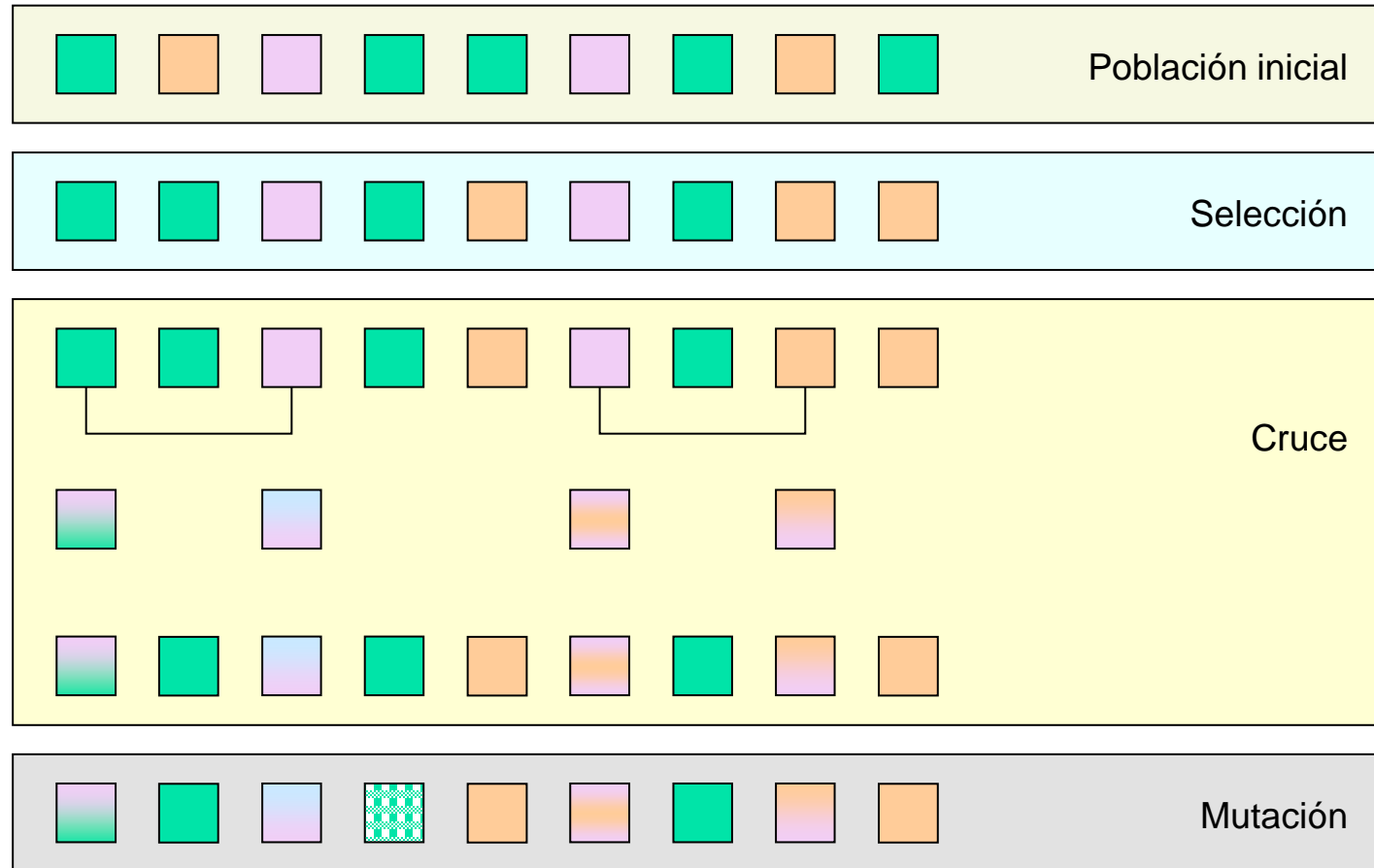
## □ Bases:

- **Población de individuos**, que son una representación de posibles soluciones.
- **Procedimiento de selección** basado en la aptitud de los individuos para resolver el problema. Se puede considerar que dirige la búsqueda.
- **Procedimiento de transformación** para construir nuevos individuos a partir de los anteriores. Componente aleatoria

## Esquema general



# Estructura de un algoritmo evolutivo



## Estructura de un algoritmo evolutivo

- ❑ Una **población** de individuos coexiste en un determinado **entorno** con recursos limitados
- ❑ Se aplica **selección** para ir obteniendo los individuos que están **mejor adaptados** al entorno
- ❑ Estos individuos se convierten en los progenitores de nuevos individuos a través de procesos de **mutación** y **cruce**
- ❑ Los nuevos individuos pasan a competir por su supervivencia
- ❑ Con el paso del tiempo, esta **selección natural** provoca el incremento en la *calidad* de los individuos de la población
- ❑ Los operadores de mutación y cruce generan **diversidad** facilitando la aparición de **novedad** en la población

# Esquema Algoritmo Genético Simple

Inicializar de forma aleatoria una población  $P$  con soluciones candidatas

Evaluar cada candidato de  $P$

Mientras no se cumpla condición de parada

1. Seleccionar según criterio de calidad  $S$  de  $P$
2. cruzar o recombinar progenitores seleccionados obteniendo descendencia  $D$
3. Mutar descendencia  $D$
4. Evaluar nuevos candidatos

## Modelos sobre el esquema general

- ❑ **Algoritmos Genéticos:** trabajan con una población de cadenas binarias.
- ❑ **Estrategias Evolutivas:** trabajan con números reales que codifican las posibles soluciones de problemas numéricos.
- ❑ **Programas o algoritmos Evolutivos:** los individuos pueden ser cualquier estructura de datos.
- ❑ **Programación Genética:** los individuos son programas o autómatas, generalmente representados en forma de árbol.
- ❑ Lo mejor es elegir la representación más adecuada para nuestro problema y elegir los operadores de variación más adecuados para dicha representación

## Modelos sobre el esquema general

- ❑ **De Jong** dice que todas las variantes definidas en la diapositiva anterior son realmente instancias concretas de un **Sistema Evolutivo General** que incluye:
  - Una o más poblaciones de individuos compitiendo por recursos limitados.
  - Poblaciones cambiantes por el nacimiento y muerte de individuos
  - El concepto de fitness que refleja la capacidad de un individuo de sobrevivir y reproducirse
  - El concepto de herencia modificada: los descendientes se parecen a sus padres pero no son idénticos



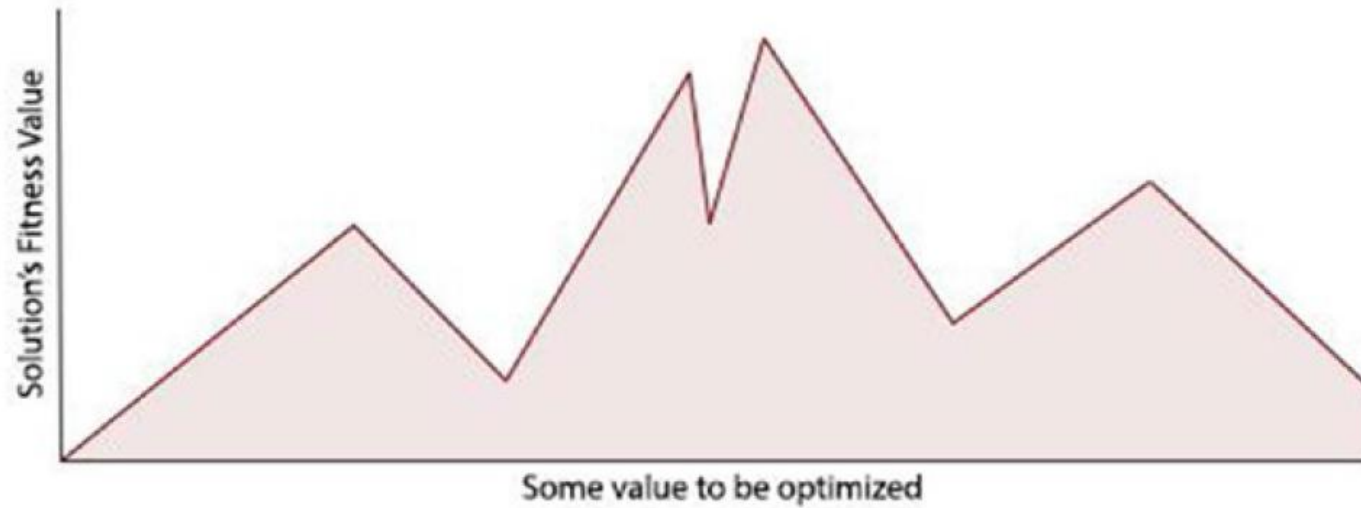
- ❑ Técnica de aprendizaje no supervisado (aprendizaje inductivo por observación y descubrimiento):
  - No hay ningún maestro que presente ejemplos, contraejemplos, ni siquiera conocimiento al sistema de aprendizaje (estos algoritmos generan ejemplos por si mismos).
  - La creación de nuevos ejemplos (puntos de búsqueda) por el algoritmo es una apuesta inductiva sobre la base del conocimiento existente.
  - Si esta apuesta se demuestra valiosa, se mantiene en la base de conocimiento (la población), en otro caso se **desecha** por medio de la selección.
- ❑ Método de búsqueda

- ❑ Tenemos un modelo de nuestro sistema y buscamos las entradas que nos conduzcan a un determinado objetivo



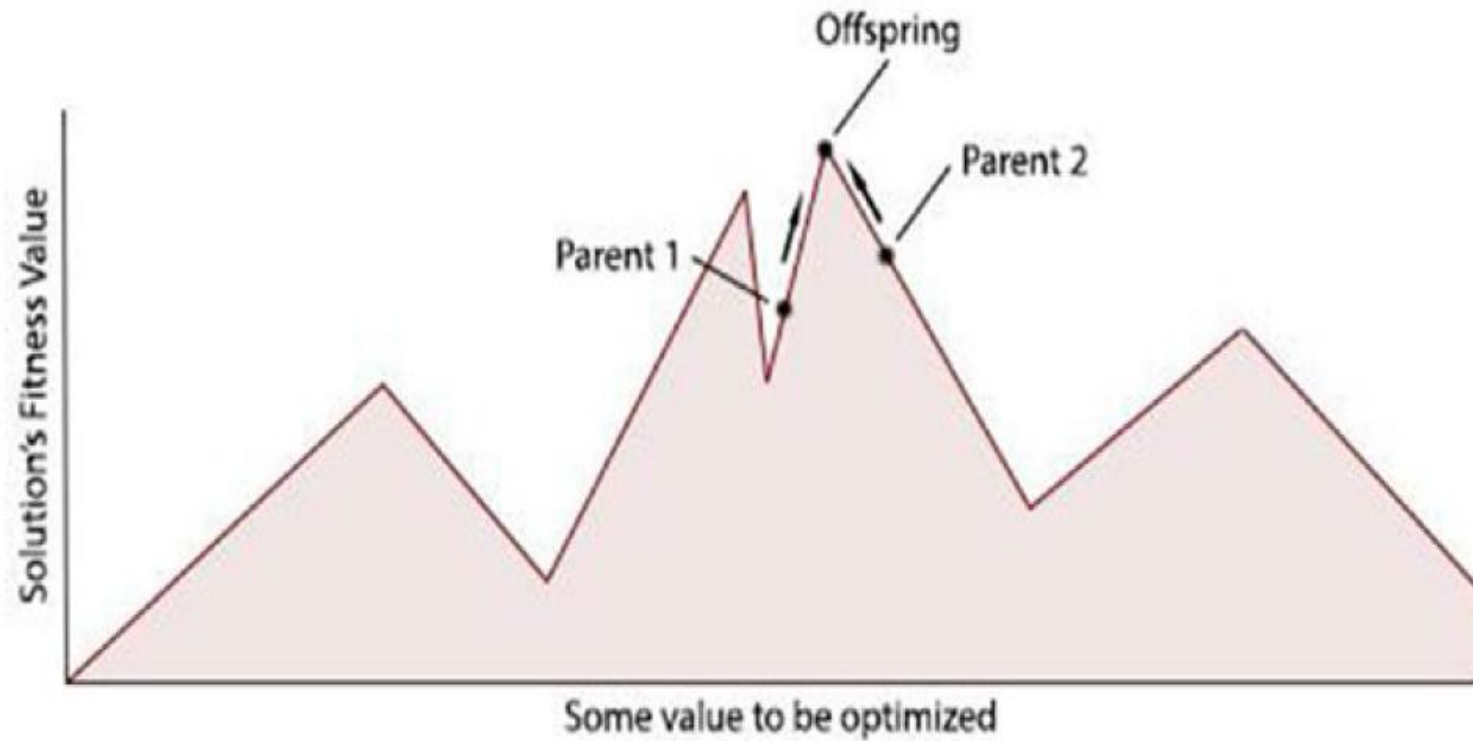
- ❑ El objetivo es encontrar un conjunto de parámetros tales que maximicen o minimicen cierto criterio de calidad. Por ejemplo, minimizar un error
- ❑ Durante la evolución, podemos comprobar que tiene lugar un proceso de **optimización** de la adaptación o aptitud
- ❑ Búsqueda en un paisaje de aptitudes (Fitness landscape)

# Paisaje de aptitud



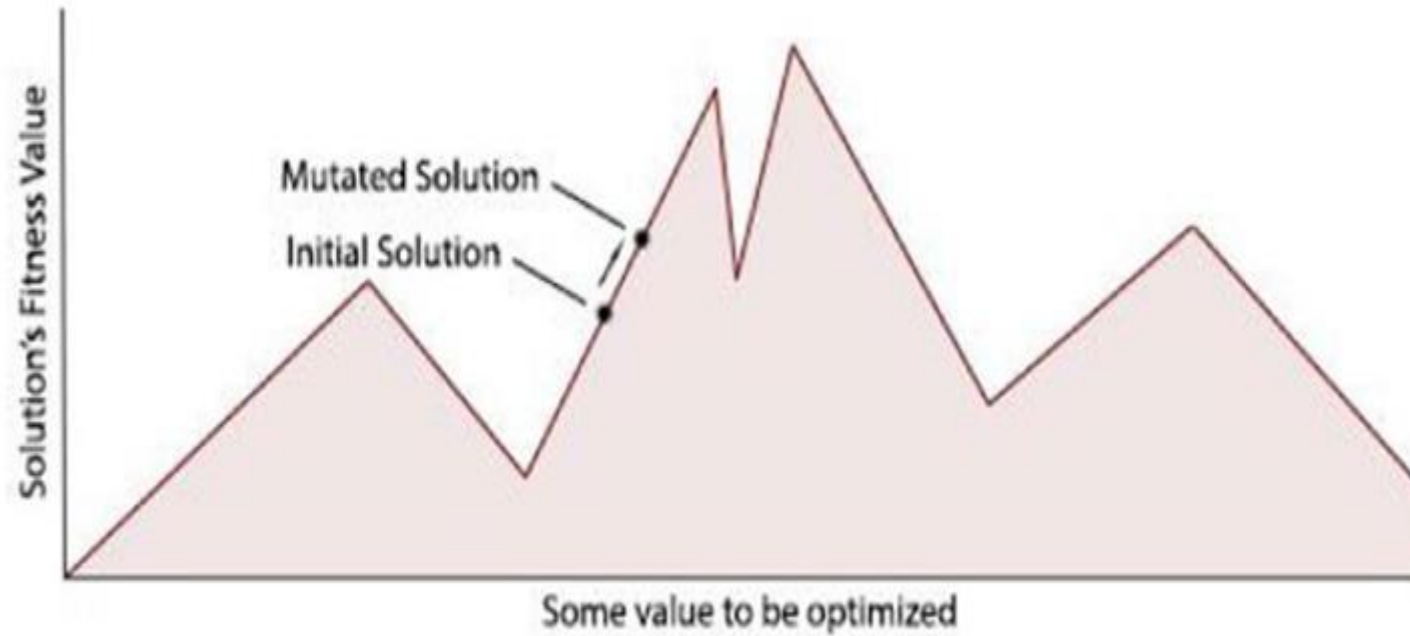
Fuente: Genetic Algorithms in Java Basics. Lee Jacobson. Burak Kanber. APress

# Paisaje de aptitud



Fuente: Genetic Algorithms in Java Basics. Lee Jacobson. Burak Kanber. APress

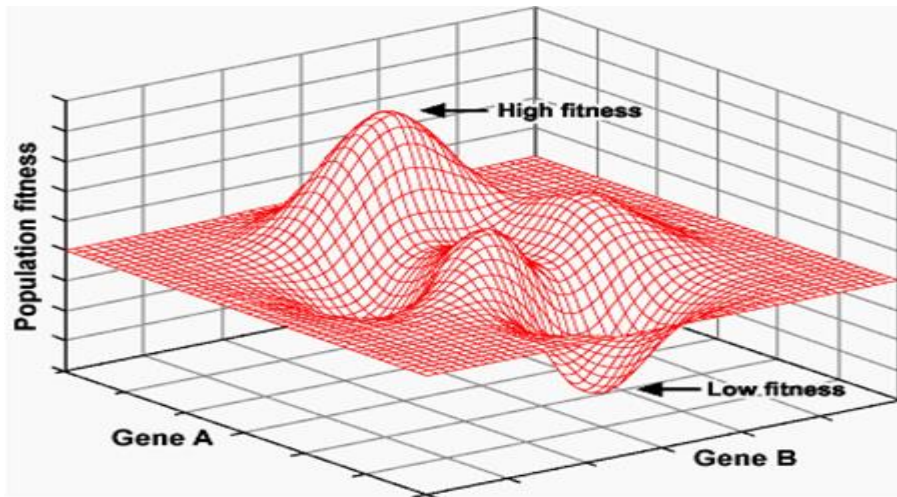
# Paisaje de aptitud



Fuente: Genetic Algorithms in Java Basics. Lee Jacobson. Burak Kanber. APress

## Paisaje de aptitud

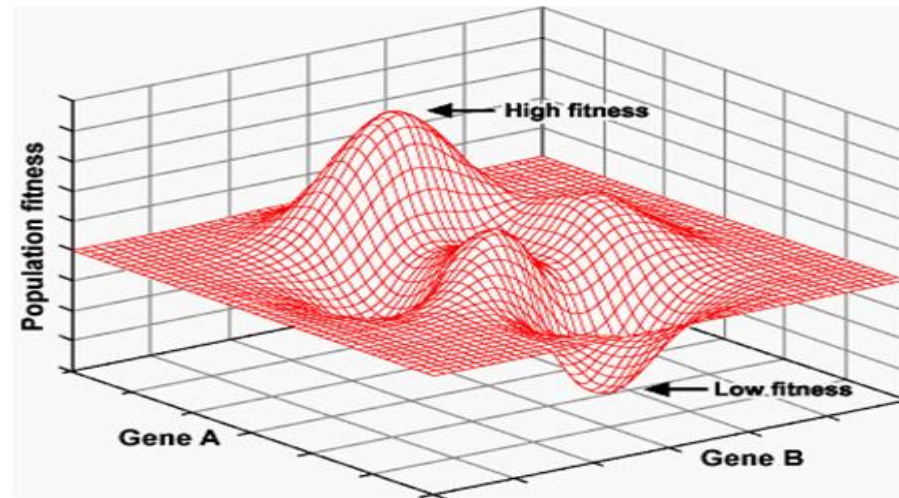
- ❑ Las posibles combinaciones de rasgos biológicos en una población de individuos definen puntos en un espacio multidimensional, donde cada eje de coordenadas corresponde a uno de estos rasgos.
- ❑ Se usa una dimensión adicional para dibujar los valores de adaptación de cada punto del espacio.



Se llama paisaje de aptitud (fitness landscape) a la hipersuperficie que se obtiene al aplicar la función de aptitud a cada punto del espacio de búsqueda.

## Paisaje de aptitud

- De esta forma se define la superficie adaptativa (topografía), que en su forma simplificada tridimensional (dos dimensiones de rasgos, y una de aptitud) parece una región montañosa, con valles, picos y puntos de silla.
- Ejemplos:
  - Optimización de funciones, Diseño de estructuras, Optimización de planificaciones (horarios, recursos ...)

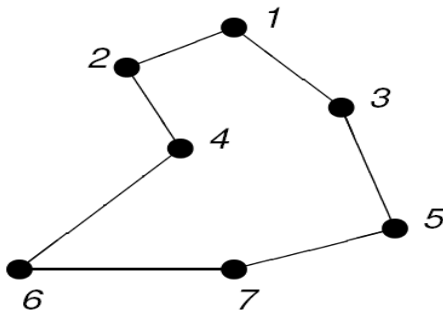


- Búsqueda y optimización son dos facetas de un mismo concepto:
  - Búsqueda se refiere al proceso
  - Optimización al resultado del proceso.
- El objetivo de un proceso de optimización es resolver

$$\mathbf{x}^{opt} = \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argopt}} \mathbf{f}(\mathbf{x})$$



- Clasificación de los procesos de optimización según la naturaleza de las soluciones:
  - **Numéricas**: Si la solución queda completamente especificada en términos de un conjunto de  $m$  parámetros o atributos.
  - **Combinatorias**: Si para especificar la solución hay que especificar un conjunto de  $m$  parámetros y el orden (total o parcial) con que estos se combinan para dar dicha solución. Ejemplo TSP



- Según el grado de aleatoriedad del proceso de búsqueda:
  - **Deterministas o dirigidas**: El procedimiento de búsqueda proporciona siempre el mismo resultado para las mismas condiciones de partida.
  - **Aleatorias o al azar**: El procedimiento de búsqueda es completamente aleatorio. Habitualmente, se delimita una región de búsqueda y se toman puntos al azar dentro de ellas. Después mediante argumentos estadísticos se hace una selección del valor del óptimo.
  - **Estocásticas u orientadas**: Se combinan la búsqueda determinista con la búsqueda aleatoria. La componente determinista orienta la dirección de búsqueda y la aleatoria se encarga de la búsqueda local.

- ❑ Según la dirección preferente de búsqueda:
  - **En profundidad**: Se da prioridad a la *explotación* de las soluciones disponibles antes que a la exploración de otras nuevas.
  - **En anchura**: La búsqueda da prioridad a la *exploración* de nuevas soluciones antes que a la explotación de las disponibles.
- ❑ Según el número de candidatos a solución que se mantienen simultáneamente:
  - **Simples**: Se mantiene un sólo candidato a solución que se va actualizando sucesivamente para proporcionar soluciones mejores al problema.
  - **Múltiples**: Se mantienen simultáneamente varios candidatos a solución con los cuales se va acotando cada vez con más precisión la región (o regiones) donde se encuentra el/los óptimo/s (más apropiadas para explotar paralelismo).

- Según la información disponible sobre la función a optimizar:
  - **Ciegas**: No se dispone de ninguna información explícita sobre la función.
  - **Heurísticas**: Se dispone de cierta información explícita sobre el proceso a optimizar (conocimiento específico), que se puede aprovechar para guiar la búsqueda.
- Los Algoritmos Genéticos (AGs) son métodos estocásticos de búsqueda ciega de soluciones cuasi-óptimas.