UNIVERSIDAD DE GUANAJUATO

Inteligencia Artificial

Enero-Junio 2021

Martes – Jueves

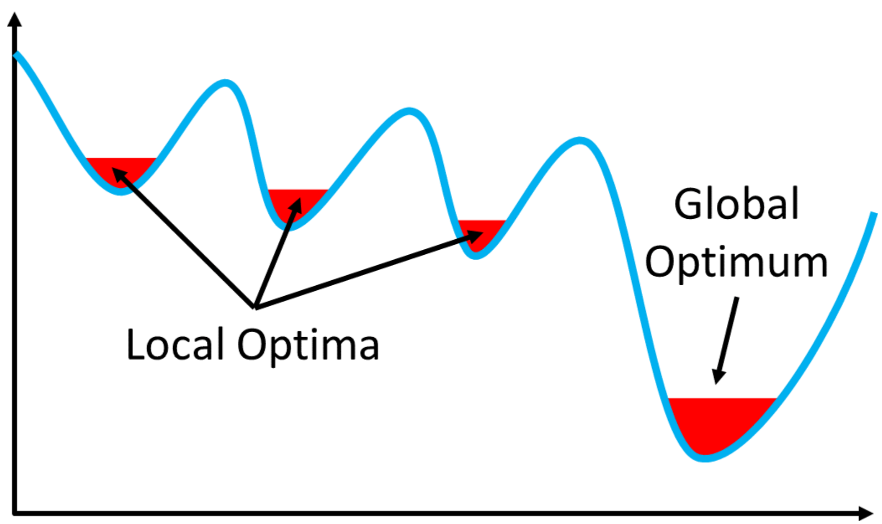
12:00 – 13:30

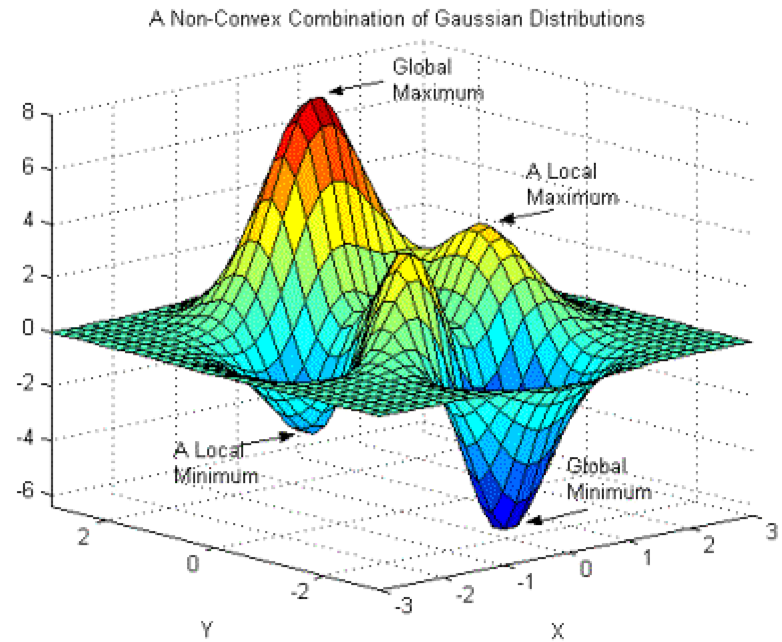
Contenido Unidad III

- ¿Que es la Optimización?
- Tipos de Problemas de Optimización
- Taxonomía de Modelos de Optimización
- Conceptos básicos Metaheurísticas
- Técnicas Metaheurísticas

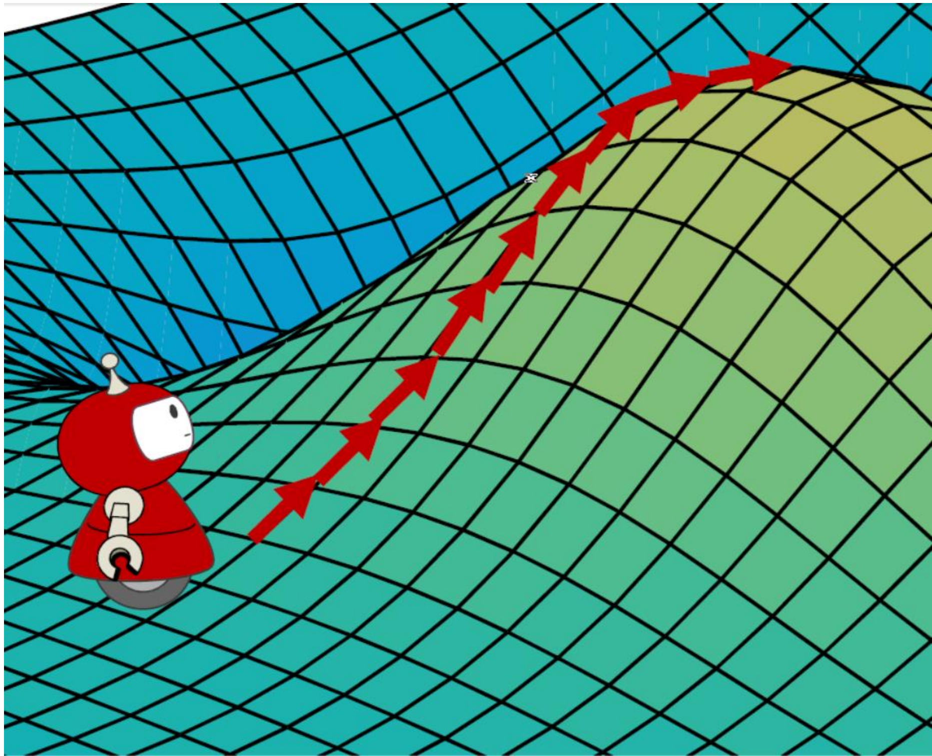
- Los problemas de optimización se encuentran en diversos dominios como: ciencias, ingeniería, administración, negocios, etc.
- Un problema de optimización puede ser definido por la dupla (S, f), donde S representa el conjunto de soluciones factibles, y $f : S \rightarrow \mathbb{R}$ es la función objetivo a optimizar. La función f permite definir una relación de orden total entre cualquier par de soluciones en el espacio de búsqueda.

- **Optimo Global:** Una solución $s^* \in S$ es un optimo global si tiene un mejor valor de acuerdo a la función objetivo que todas las soluciones del espacio de búsqueda, esto es, $\forall s \in S, f(s^*) \leq f(s)$

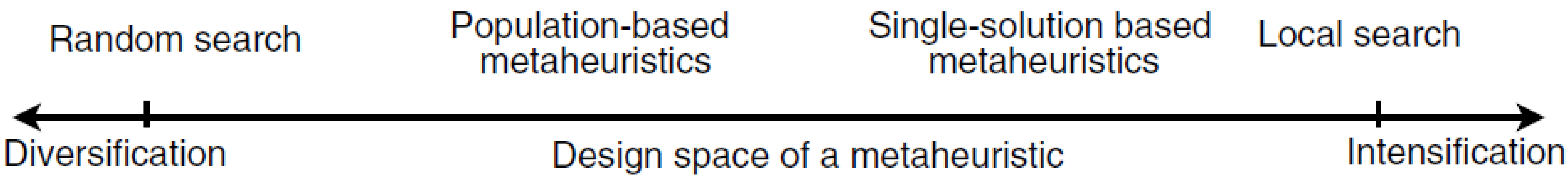




- Los tipos de optimización se puede clasificar de acuerdo a la naturaleza de las variables de decisión:
  - Optimización continua
  - Optimización discreta
    - Binaria
    - Combinatoria
  - Optimización Mixta

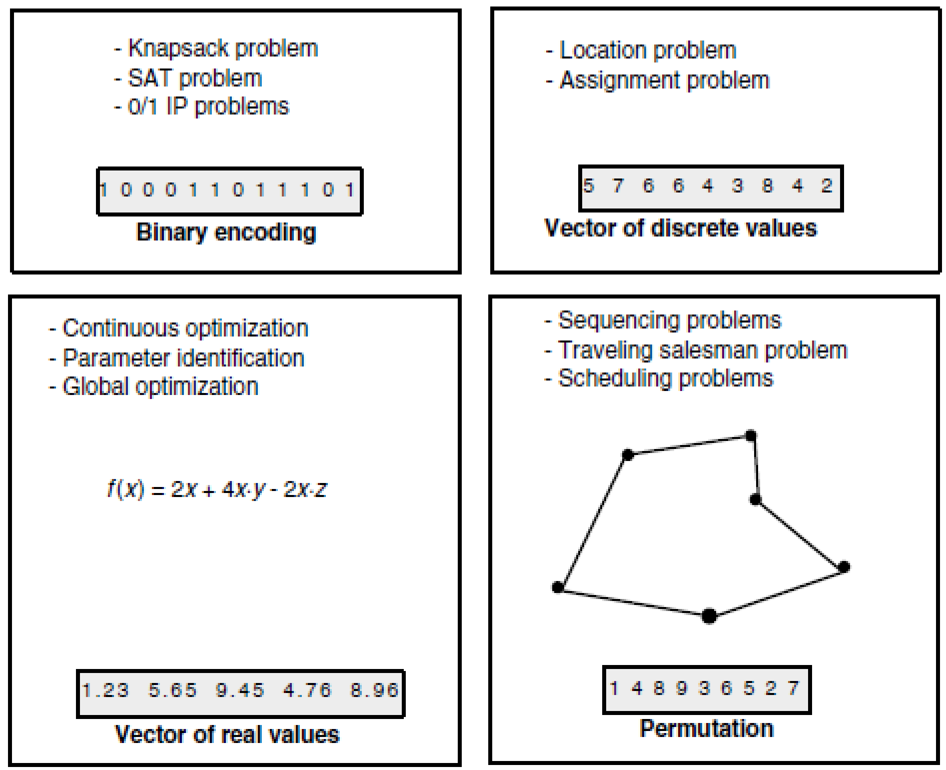


- Métodos Exactos
  - Rama y X
  - Programación de restricciones
  - Programación dinámica
  - A*, IDA*
- Métodos de Aproximación
  - Algoritmos de aproximación
  - Algoritmos Heurísticos
    - Heurísticas específicas al problema
    - Metaheurísticas
      - Basadas en una sola solución
      - Basadas en una población



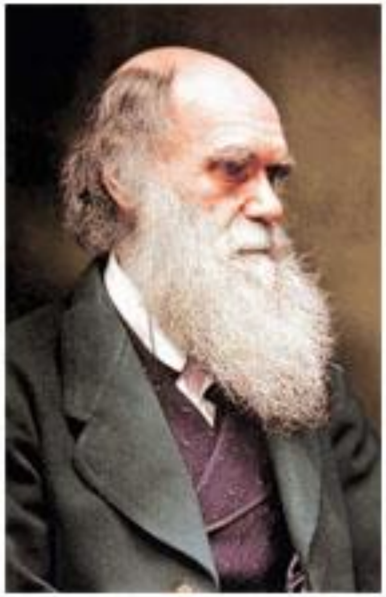
- Metaheurísticas:
  - Evolutivos
  - Bioinspirados
  - Basados en fenómenos físicos

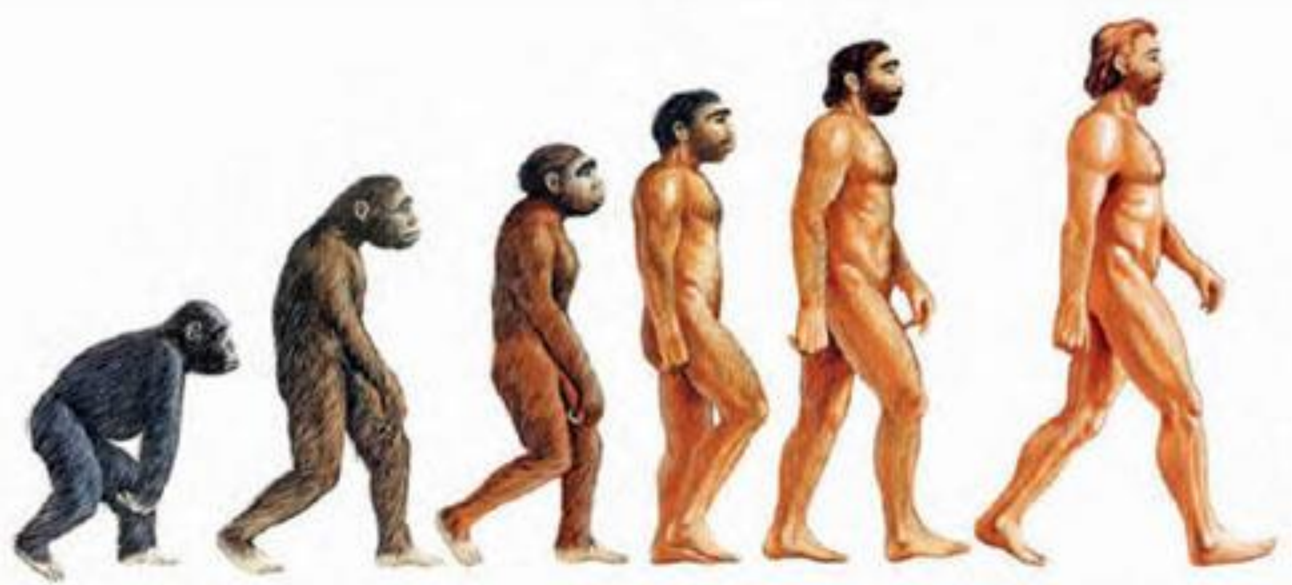
- Representación, una solución debe de poseer:
  - Completitud
  - Conexión
  - Eficiencia



- Es un algoritmo evolutivo, propuesto por John Holland en 1975.
- Es una simulación de la selección natural que puede resolver problemas de optimización.
- Características que simula un AG acerca de la selección natural:
  - Un sistema biológico consta de una población de individuos, muchos de los cuales tienen la habilidad de reproducirse.
  - Los individuos tienen una vida útil finita
  - Hay variación en la población
  - La habilidad para sobrevivir está correlacionada positivamente con la habilidad de reproducirse.

- Características:
  - Representación: cadenas de bits (1s y 0s)
  - Selección: Ruleta (Proporcional)
  - Variaciones: Cruza.





- Inicio
  - $t = 0$
  - **Inicializar** $P(t)$
  - **Evaluar** estructuras en $P(t)$
  - Mientras no se llegue a la condición de paro, hacer
    - $t = t + 1$
    - **Seleccionar** $R(t)$ de $P(t-1)$
    - **Cruzar** $R(t)$ para obtener $\mathrm{C}(t)$
    - **Mutar** $\mathrm{C}(t)$ para obtener $\mathrm{C}'(t)$
    - **Evaluar** estructuras en $C'(t)$
    - **Reemplazar** $P(t)$ con $C'(t)$ y $P(t-1)$

- Optimizar Función One Max $(X \in \mathfrak{B}^d)$
  - $\max f(x) = \sum_{i=1}^{d} x_i$
  - $\min f(x) = -\sum_{i=1}^{d} x_i$
- $d = 4$, $\text{pop\_size} = 10$
  - Inicializar y evaluar $P(t)$

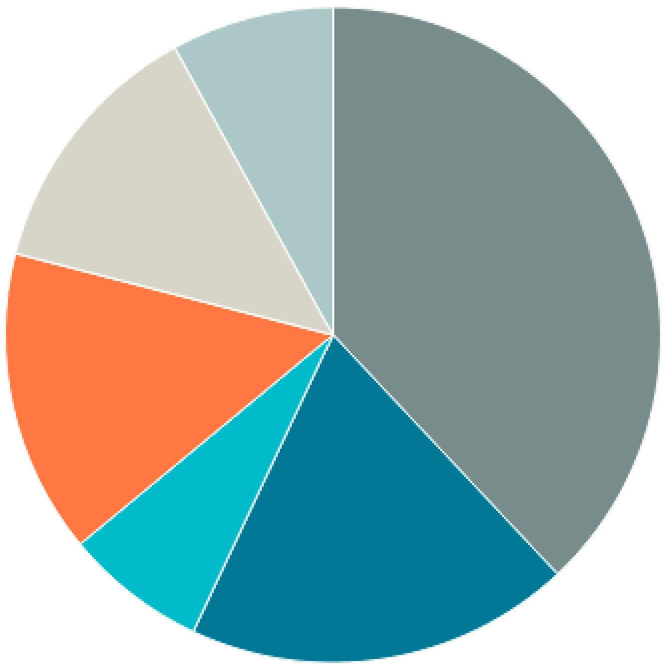
$P(t)$

x_1	x_2	x_3	x_4	$f(X)$
0	1	1	0	2
1	0	1	1	3
1	1	0	1	3
1	0	1	0	2
0	0	0	0	0
0	1	0	1	2
1	0	1	1	3
0	1	0	0	1
0	0	0	1	1
1	1	1	0	3

– **Seleccionar** (Ruleta) $R(t)$ de $P(t-1)$

$P(t)$

#	x_1	x_2	x_3	x_4	$f(X)$	$f_{norm}(X$
1	0	1	1	0	2	0.1
2	1	0	1	1	3	0.15
3	1	1	0	1	3	0.15
4	1	0	1	0	2	0.1
5	0	0	0	0	0	0
6	0	1	0	1	2	0.1
7	1	0	1	1	3	0.15
8	0	1	0	0	1	0.05
9	0	0	0	1	1	0.05
10	1	1	1	0	3	0.15
					20	1



– **Cruzar** (Cruza en un Punto) $R(t)$ para obtener C(t)

$R(t)$

Pareja	x_1	x_2	x_3	x_4
1	1	0	1	1
1	1	1	0	1
2	0	1	0	1
2	1	1	0	1
3	0	1	0	0
3	0	1	1	0
4	1	0	1	0
4	1	1	1	0
5	1	1	1	0
5	1	0	1	1

$C(t)$

Pareja	Punto	x_1	x_2	x_3	x_4
1	2	1	0	0	1
1	2	1	1	1	1
2	3	0	1	0	1
2	3	1	1	0	1
3	3	0	1	0	0
3	3	0	1	1	0
4	1	1	1	1	0
4	1	1	0	1	0
5	3	1	1	1	1
5	3	1	0	1	0

– **Mutar** (Mutación de 1 bit) C(t) para obtener C$'(t)$

$C(t)$

x_1	x_2	x_3	x_4
1	0	0	1
1	1	1	1
0	1	0	1
1	1	0	1
0	1	0	0
0	1	1	0
1	1	1	0
1	0	1	0
1	1	1	1
1	0	1	0

$C'(t)$

x_1	x_2	x_3	x_4
1	1	0	1
1	1	0	1
0	1	0	0
1	1	1	1
0	0	0	0
1	1	1	0
1	0	1	0
1	0	1	1
0	1	1	1
1	1	1	0

– **Evaluar** estructuras en $C'(t)$

$C'(t)$

x_1	x_2	x_3	x_4	$f(X)$
1	1	0	1	3
1	1	0	1	3
0	1	0	0	1
1	1	1	1	4
0	0	0	0	0
1	1	1	0	3
1	0	1	0	2
1	0	1	1	3
0	1	1	1	3
1	1	1	0	3

– **Reemplazar** $P(t)$ con $C'(t)$

$P(t)$

x_1	x_2	x_3	x_4	$f(X)$
1	1	0	1	3
1	1	0	1	3
0	1	0	0	1
1	1	1	1	4
0	0	0	0	0
1	1	1	0	3
1	0	1	0	2
1	0	1	1	3
0	1	1	1	3
1	1	1	0	3

- El principal objetivo de la selección es procurar las mejores soluciones en una población.
- Se utiliza en:
  - Reproducción - Elección de buenas soluciones para crear (idealmente) mejores soluciones.
  - Reemplazo – Copias de una o mas buenas soluciones son colocadas en la población siguiente.
- Operadores:
  - Selección Aleatoria
  - Selección Proporcionada (Ruleta)
  - Selección Torneo n-ario
- ¿Presión de selección?
- ¿Brecha generacional?

- Es un mecanismo el cual intercambia subcadenas entre las cadenas de símbolos de los individuos.
- Operadores:
  - Cruza en Un Punto
  - Cruza en n Putos (n=2, sugerido)

Cruza en 2 Puntos

$R(t)$

Pareja	x_1	x_2	x_3	x_4
1	1	0	1	1
1	1	1	0	1
2	0	1	0	1
2	1	1	0	1
3	0	1	0	0
3	0	1	1	0
4	1	0	1	0
4	1	1	1	0
5	1	1	1	0
5	1	0	1	1

$C(t)$

Pareja	Punto1	Punto2	x_1	x_2	x_3	x_4
1	1	3	1	1	0	1
1	1	3	1	0	1	1
2	2	3	0	1	0	1
2	2	3	1	1	0	1
3	2	3	0	1	1	0
3	2	3	0	1	0	0
4	1	3	1	1	1	0
4	1	3	1	0	1	0
5	2	3	1	1	1	1
5	2	3	1	0	1	1

- Es un mecanismo para proveer diversidad a las soluciones generadas en la cruza
- Operadores:
  - Muta Negador de 1 bit
  - Muta Negador de N bits
  - Muta con base a una Probabilidad

- El problema SAT es el problema de saber si, dada una expresión booleana con variables y sin cuantificadores, hay alguna asignación de valores para sus variables que hace la expresión verdadera.
- Por ejemplo, dadas x_1, x_2, x_3, x_4, determine los valores para que la siguiente expresión sea verdadera:
- $(x_1 \vee x_3) \wedge (\neg x_2 \vee x_3 \vee \neg x_4)$

- La representación canónica de los algoritmos genéticos consiste de vectores binarios de longitud fija ℓ; es decir, el espacio de búsqueda es $\mathrm{I} = \{0,1\}^{\ell}$ y los individuos $a = (a_1, \dots, a_{\ell}) \in \{0,1\}^{\ell}$. Esta representación es adecuada para problemas de optimización de la forma $f\colon \{0,1\}^{\ell} \rightarrow \mathbb{R}$.

- Los algoritmos genéticos también pueden ser usados en problemas de optimización de la forma $f: S \to \mathbb{R}$, donde S difiere de un vector binario del espacio $\{0,1\}^{\ell}$. Un ejemplo de esto, es la aplicación del algoritmo genético para problemas de optimización con parámetros continuos $f: \mathbb{R}^n \to \mathbb{R}$.

- Los mecanismos de codificación y decodificación entre dos espacios diferentes $\{0,1\}^{\ell}$ y $\mathbb{R}^n$ requieren restringir el espacio continuo a intervalos finitos $[u_i, v_i]$ para cada variable $\mathrm{x} \in \mathbb{R}$.
- El vector binario se divide en n segmentos de (en la mayoría de los casos) longitud igual a ℓ_x, tal que $\ell = n\ell_x$.
- El subsegmento $\left(a_{(i-1)\ell_x+1}, \dots, a_{i\ell_x}\right)$ $(i = 1, \dots, n)$ es la codificación binaria de la variable x_i.

- La decodificación $\Gamma^i\colon \{0,1\}^{\ell} \rightarrow [\mathrm{u_i}, \mathrm{v_i}]$, se realiza de la siguiente manera

$$\Gamma^i(a_1, \ldots, a_\ell) = u_i + \frac{v_i - u_i}{2^{\ell_x} - 1}\left(\sum_{j=0}^{\ell_x - 1} a_{i\ell_x - j} 2^j\right)$$

- La ecuación anterior puede ser adaptada a usar la codificación Gray, la cual asegura que los valores enteros adyacentes son representados por vectores binarios con distancia de Hamming igual a 1.

- Considere la función esfera ($n = 2$):

$$f(x) = \sum_{i=1}^{2} (x_i)^2 \, ; x \in \mathbb{R}^n$$

  – Cada componente x_i será acotada al intervalo $u_i = -5.21$ y $v_i = 5.21$.

  – Cada componente x_i será codificada por un subsegmento binario $\ell_x = 3$; por lo tanto, la longitud total del vector binario será $\ell = n\ell_x = 6$.

a_1	a_2	a_3	a_4	a_5	a_6
1	1	1	1	0	1
x_1			x_2		

Representación: Ejemplo (Decodificación)

a_1	a_2	a_3	a_4	a_5	a_6
1	1	1	1	0	1
x_1			x_2		

- Decodificación de x_1

$$\Gamma^1 = x_1 = -5.21 + \frac{5.21 - (-5.21)}{2^3 - 1}\left(\sum_{j=0}^{3-1} a_{1(3)-j} 2^j\right)$$

$$\Gamma^1 = x_1 = -5.21 + \frac{10.42}{7}(1 + 2 + 4) = 5.21$$

- Decodificación de x_2

$$\Gamma^2 = x_2 = -5.21 + \frac{5.21 - (-5.21)}{2^3 - 1}\left(\sum_{j=0}^{3-1} a_{2(3)-j} 2^j\right)$$

$$\Gamma^2 = x_2 = -5.21 + \frac{10.42}{7}(1 + 0 + 4) = 2.23$$

- El vector decodificador $x = [5.21, 2.23]$, es evaluado en la función objetivo
$$f(x) = 5.21^2 + 2.23^2 = 32.117$$

UNIVERSIDAD DE
GUANAJUATO