



SIA - TP2



Algoritmos Genéticos



Ejercicio 1

Algoritmo genético en imágenes de $N \times N$



Caracterización del problema

- Los individuos serán matrices de $N \times N$ caracteres ASCII
- Cada gen buscará representar una porción de pixels de $M \times M$ de la imagen original
- Los cromosomas consisten de $N \times N$ genotipos
- Por ejemplo, dividimos la imagen en una grilla de 10×10 y cada cuadrado verde debería ser representado por un gen. La imagen resultante de la matriz de 10×10 de caracteres ASCII sería nuestro fenotipo



Fenotipos posibles

Fenotipos de caracteres ASCII:

Representa la imagen original utilizando caracteres ASCII.

Una posibilidad para comparar qué tanto se parece un individuo a la imagen original es comparando cada caracter ASCII a su porción correspondiente, pero conviene comparar la imagen entera para soportar tamaños distintos.



Entrada



Representación de la posible solución

Fenotipo y función de fitness

¡El fenotipo puede afectar a la función de fitness ya que este determina cómo se representa la solución del algoritmo!

Una buena función de fitness sería aquella que mide la similitud entre la solución obtenida y la imagen original.

Función de fitness

Fenotipo de píxeles

- **ECM**: se calcula la diferencia entre la imagen original y se eleva al cuadrado para obtener el ECM. A menor ECM mejor la solución.
- **Similitud estructural**: se mide la similitud estructural entre ambas imágenes a partir de los píxeles. A mayor similitud mejor la solución.

Fenotipo de caracteres

- **Similitud de caracteres**: se mide la similitud de los caracteres utilizados en la solución con respecto a la imagen original. A mayor similitud, mejor es la solución. Se puede, por ejemplo, comparar qué tanto espacio ocupa un carácter contra qué tan oscuro es el pixel correspondiente.
- **Similitud estructural**: idem anterior pero con caracteres.

Ejercicio 2



Estructura del individuo

Genes: proporción de cada color de la paleta inicial.

Alelos: 0 a 100 (float)

Ejemplo:

Si la paleta de colores es Rojo, Verde y Azul. Un individuo podría tener los genes 20% Rojo, 40% Verde y 40% Azul.

Input

- Paleta de colores (RGB)
- Color objetivo (RGB)
- Hiperparámetros

Hiperparámetros:

- Tamaño de la población
 - Método de selección
 - Método de cruza
 - Método de mutación
 - Parámetros de mutación
 - Método de finalización/corte
 - Parámetros de finalización
-

Métodos de selección

- Ruleta
- Elite
- Torneos determinísticos
- Torneos probabilísticos

Métodos de cruza

- Un punto
 - Dos puntos
 - Anular
 - Uniforme
-

Métodos de mutación

- Gen
- Multigen limitado
- Multigen uniforme
- Completo

Parámetros

- Probabilidad de mutar
 - Cambio en porcentaje (sin normalizar)
-

Métodos de finalización

- Por tiempo
- Por generación límite
- Por distancia al objetivo

Parámetros

- Segundos transcurridos
 - Número de generación
 - Distancia al objetivo
-

Output

Proporción de cada color de la paleta a utilizar

Representación gráfica de cada color para mejor entendimiento

Similitud con el color objetivo

Se pueden apreciar todas las generaciones con su output.

Se utilizó la librería *colr* para imprimir todos los colores en RGB.

La similitud se calcula como la distancia euclidiana en LAB space.

Para el fitness se utiliza la máxima distancia posible menos la similitud respectiva

Resultados

Gráficos

Muchos hiperparámetros:
Ceteris paribus

Fundamentación de parámetros

- Paletas de colores iniciales:
 - CMYK
 - RGB
 - Pastel (Brewer)
 - Pantone
 - Colores objetivos:
 - Complementarios
 - Cercanos
-

Gráficos

Cambiamos los
hiperparámetros



Gráficos

Comparación Paletas de colores

Cantidad de repeticiones: 10

Parámetros de corte:

- por tiempo (5s), generacion (12) y distancia (0.4)

Parámetros de mutación:

- Completa
- Probabilidad de mutar: 0.4
- Porcentaje de cambio: 10%

Población Máxima: 20 individuos

Selección: Ruleta

Cruza: Uniforme

Aquí comparamos la versatilidad de cada paleta, intentando medir “Qué tan fácil es formar distintos colores con cada paleta”. La paleta RGB es indudablemente la más capaz, llegando a la menor distancia en promedio para cualquier color probado, y la Pastel es la menos versátil.

Colores objetivo promediados:

- Rojo (255, 0, 0)

- Verde (0, 255, 0)

- Azul (0, 0, 255)

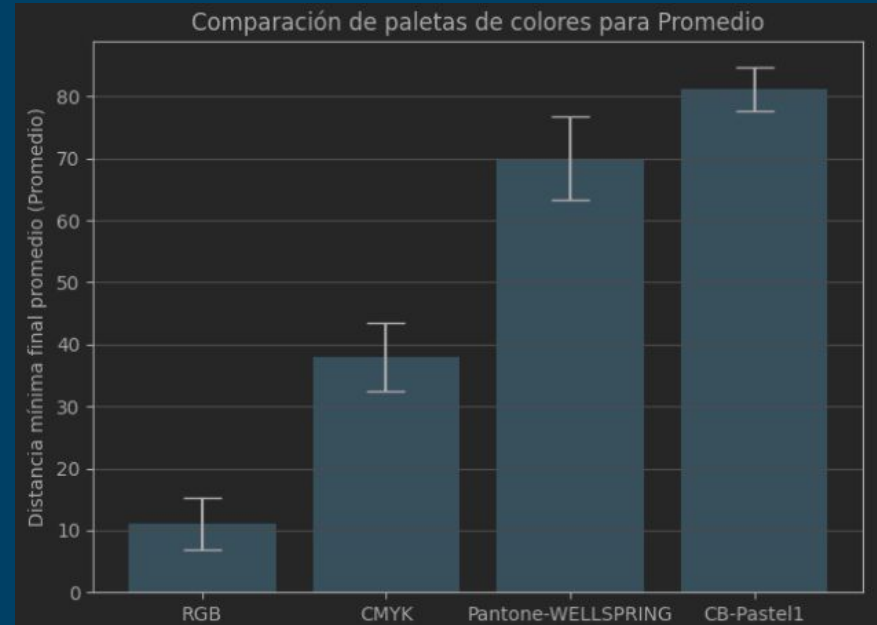
- Cobre (184, 115, 51)

- Naranja Pastel (247, 177, 149)

- Azul Pastel (117, 158, 203)

- Violeta (143, 0, 255)

- Verde Pastel (214, 248, 184)



Gráficos

Comparación Paletas de colores

Cantidad de repeticiones: 10

Colores objetivo: Verde Pastel (214, 248, 184)

Parámetros de corte:

- por tiempo (5s), generacion (12) y distancia (0.4)

Parámetros de mutación:

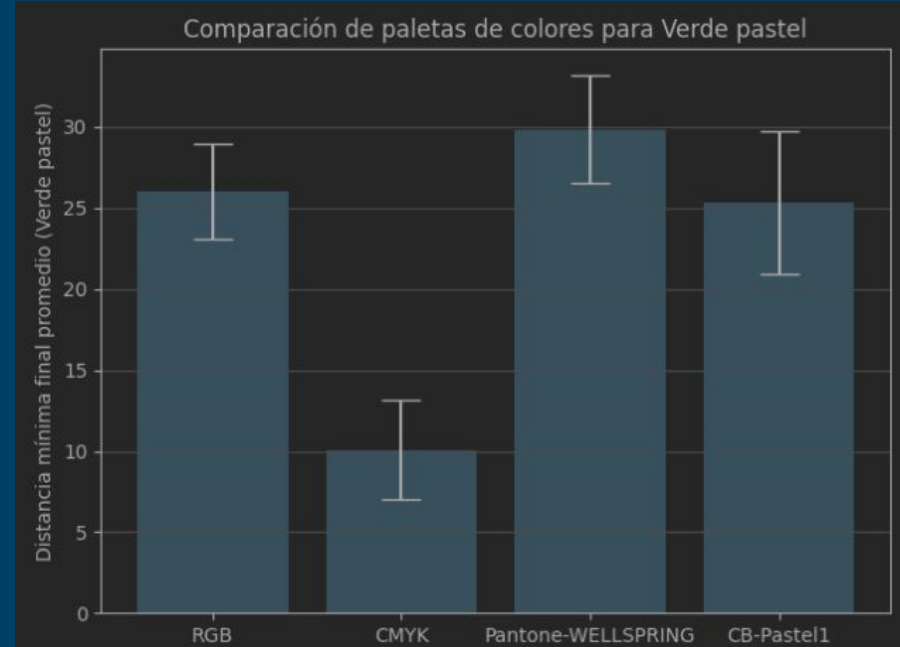
- Completa
- Probabilidad de mutar: 0.4
- Porcentaje de cambio: 10%

Población Máxima: 20 individuos

Selección: Ruleta

Cruza: Uniforme

Si elegimos un color pastel como objetivo, la paleta de colores pastel mejora. También notamos que la paleta RGB, que antes era indiscutiblemente la más versátil, pierde su ventaja.



Gráficos

Comparación Métodos de Selección

Cantidad de repeticiones: 50

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generacion (12) y distancia (0.4)

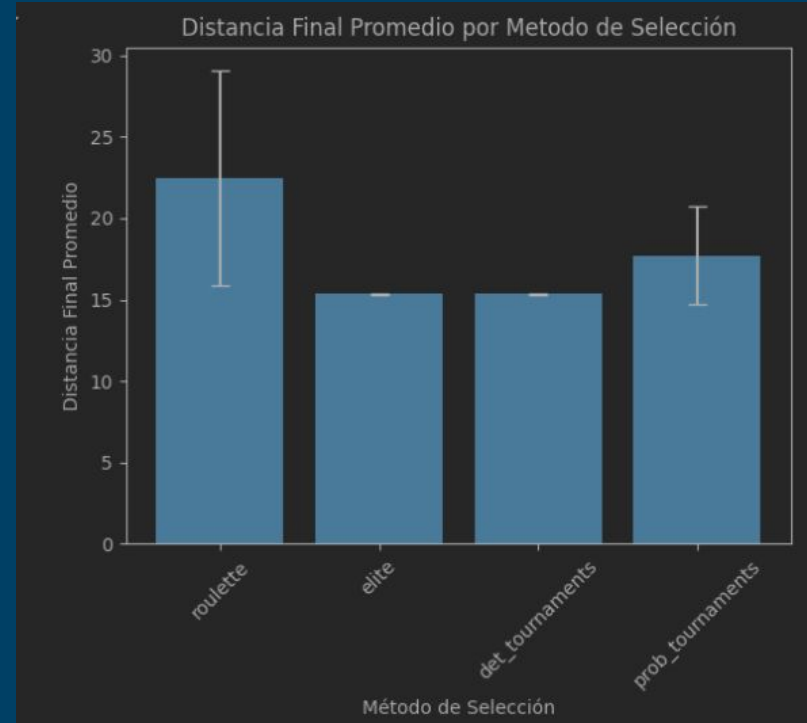
Parámetros de mutación:

- Completa
- Probabilidad de mutar: 0.4
- Porcentaje de cambio: 10%

Población Máxima: 20 individuos

Cruza: Uniforme

El método Elite es claramente superior en este caso particular. Sin embargo, en otras situaciones este podría cortar genes y llevar a poca variación genética, lo que puede llevar a menos soluciones encontradas.



Gráficos

Comparación Métodos de Selección

Cantidad de repeticiones: 10

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generacion (100) y distancia (0.4)

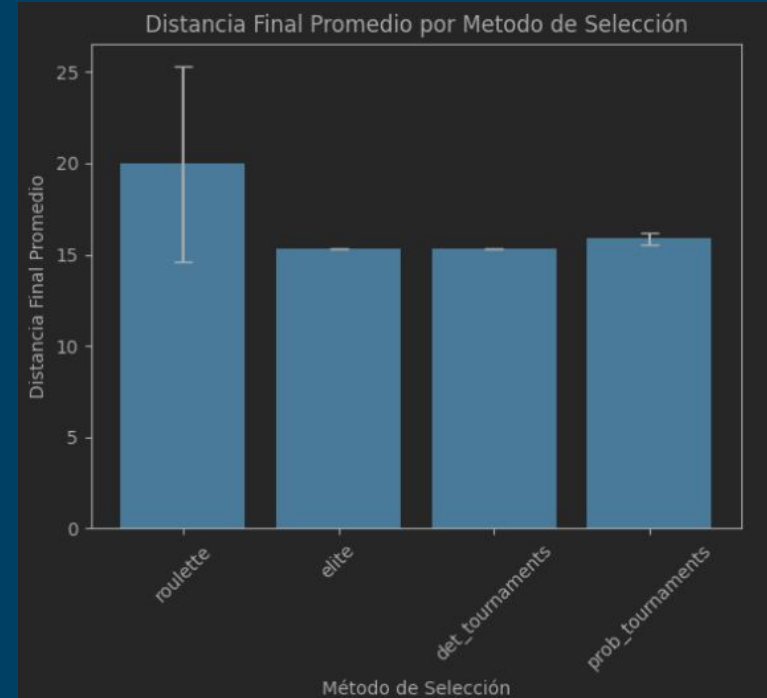
Parámetros de mutación:

- Completa
- Probabilidad de mutar: 0.1
- Porcentaje de cambio: 10%

Población Máxima: 20 individuos

Cruza: Uniforme

Al aumentar la cantidad de generaciones se achica la diferencia de Torneos Probabilísticos con elite y Torneos Determinísticos así cómo también su error.



Gráficos

Comparación Métodos de Selección

Cantidad de repeticiones: 10

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generacion (200) y distancia (0.4)

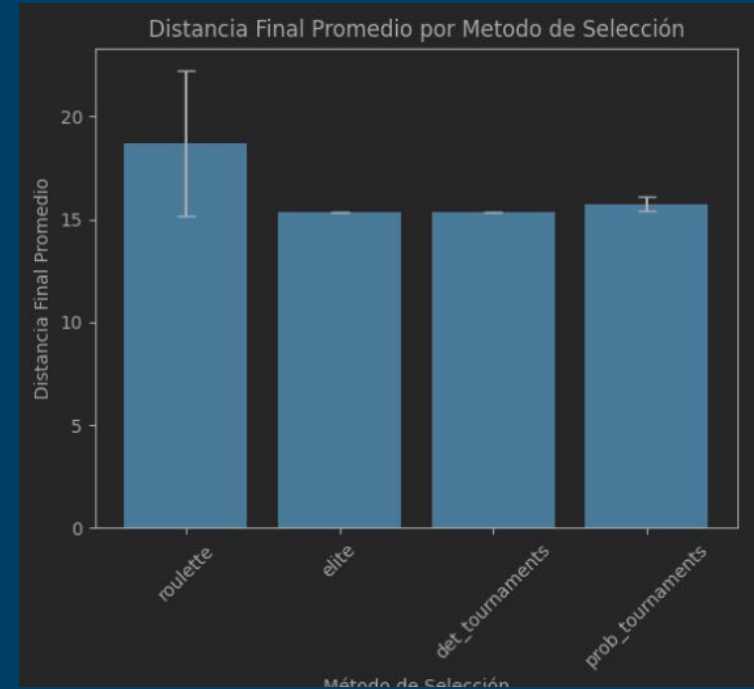
Parámetros de mutación:

- Completa
- Probabilidad de mutar: 0.1
- Porcentaje de cambio: 10%

Población Máxima: 20 individuos

Cruza: Uniforme

Al duplicar las generaciones, predomina la tendencia y mejoran los resultados generales.



Gráficos

Comparación Métodos de Selección

Cantidad de repeticiones: 10

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generacion (500) y distancia (0.4)

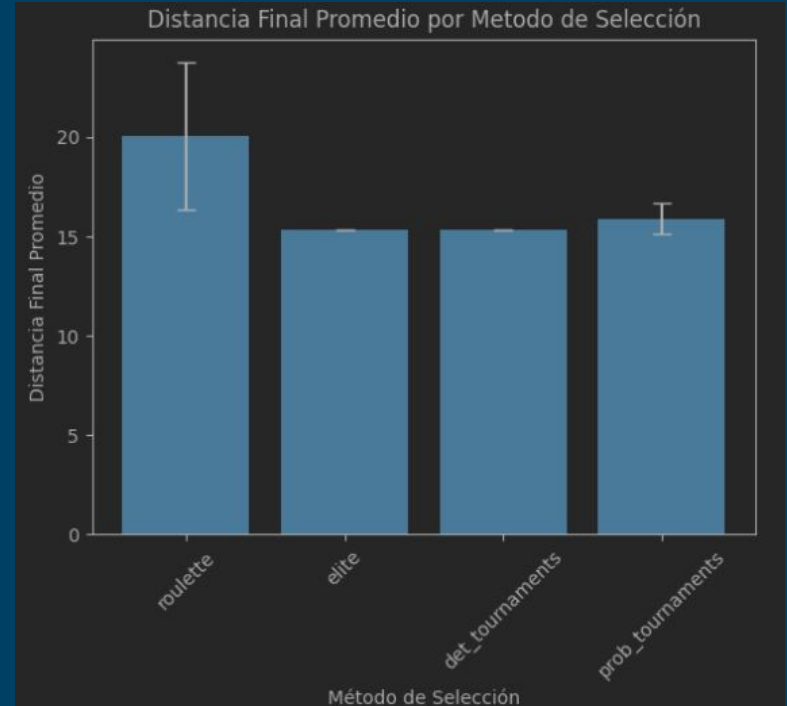
Parámetros de mutación:

- Completa
- Probabilidad de mutar: 0.1
- Porcentaje de cambio: 10%

Población Máxima: 20 individuos

Cruza: Uniforme

Al aumentar la cantidad de generaciones a 500, ya no vemos una mejora, sino que se mantuvo lo que observamos con 200.



Gráficos

Comparación Métodos de Cruza

Cantidad de repeticiones: 50

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generacion (12) y distancia (0.4)

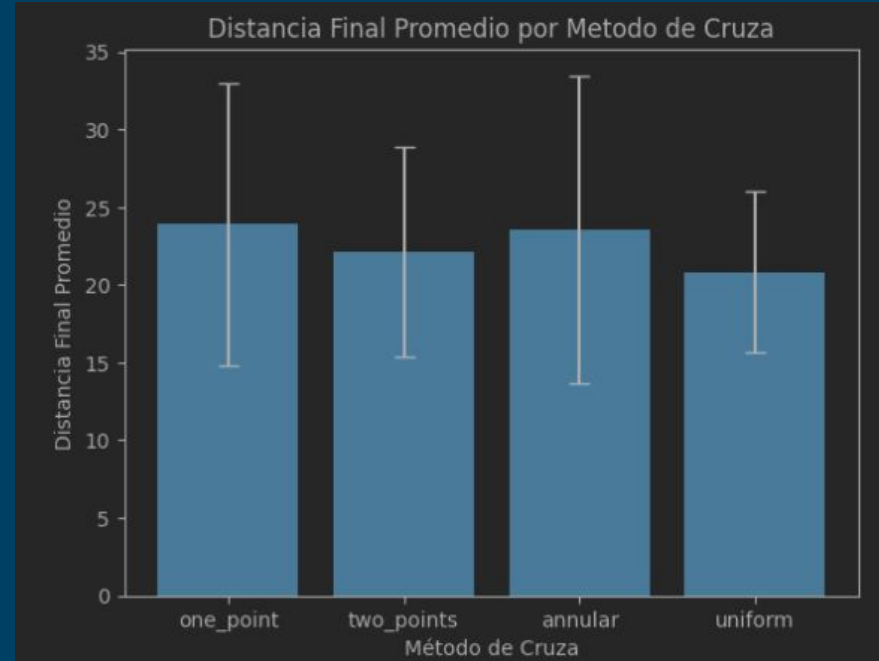
Parámetros de mutación:

- Completa
- Probabilidad de mutar: 0.4
- Porcentaje de cambio: 10%

Población Máxima: 20 individuos

Selección: Ruleta

En este caso particular no se nota tanta diferencia en los promedios en sí tanto como en los desvíos estándar. El Uniform parece producir los resultados más predecibles.



Gráficos

Comparación Métodos de Cruza

Cantidad de repeticiones: 50

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generacion (100) y distancia (0.4)

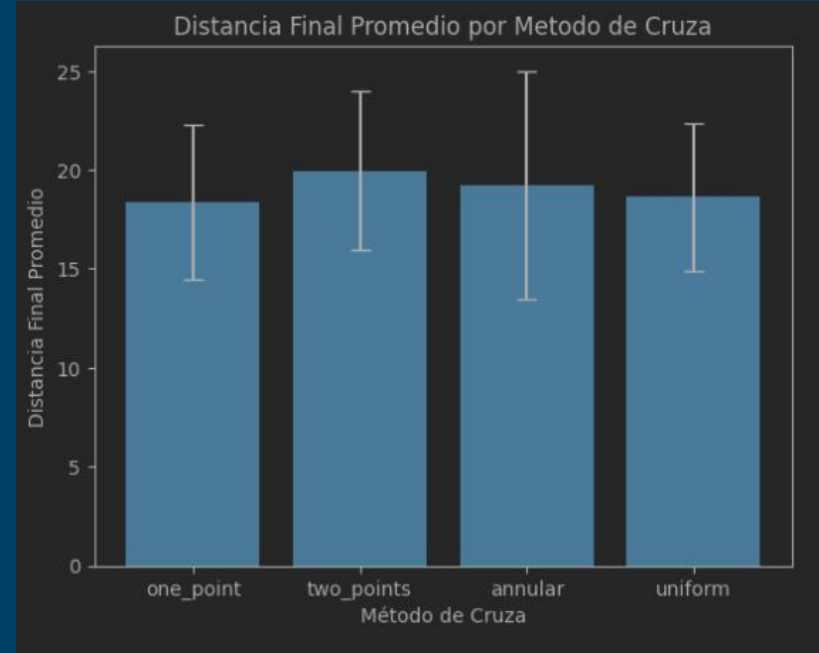
Parámetros de mutación:

- Completa
- Probabilidad de mutar: 0.1
- Porcentaje de cambio: 10%

Población Máxima: 20 individuos

Selección: Ruleta

Mejoran los resultados de distancia final, uniforme parece ser el mejor.



Gráficos

Comparación Métodos de Mutación

Cantidad de repeticiones: 50

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generacion (12) y distancia (0.4)

Parámetros de mutación:

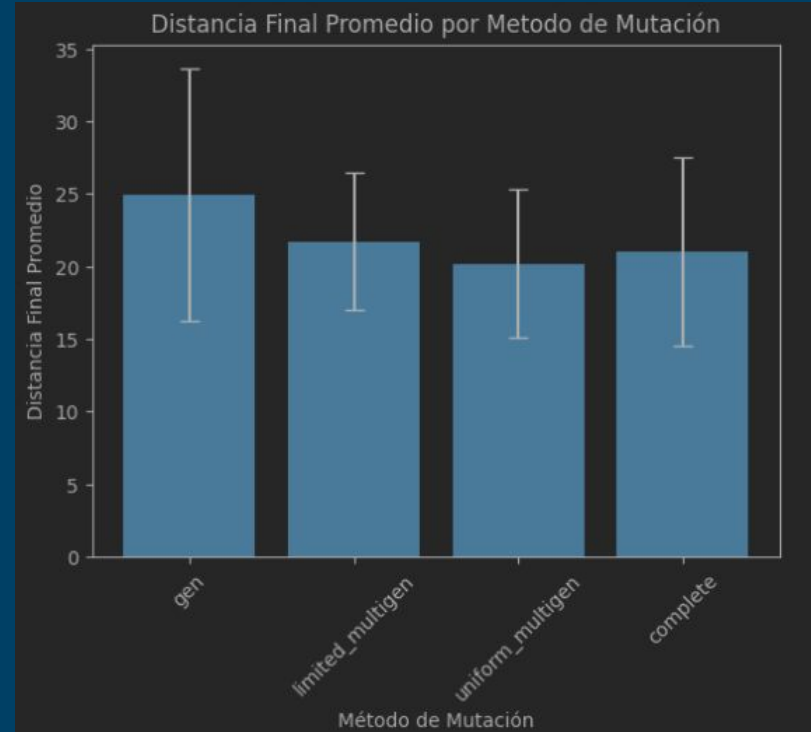
- Probabilidad de mutar: 0.4
- Porcentaje de cambio: 10%

Población Máxima: 20 individuos

Selección: Ruleta

Cruza: Uniforme

El método “gen” muta muy poco comparado a los demás, por eso vemos menos variación de los genes originales, y por ende le cuesta más generaciones poder alejarse de estos y acercarse a la solución.



Gráficos

Comparación Métodos de Mutación

Cantidad de repeticiones: 10

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generación (100) y distancia (0.4)

Parámetros de mutación:

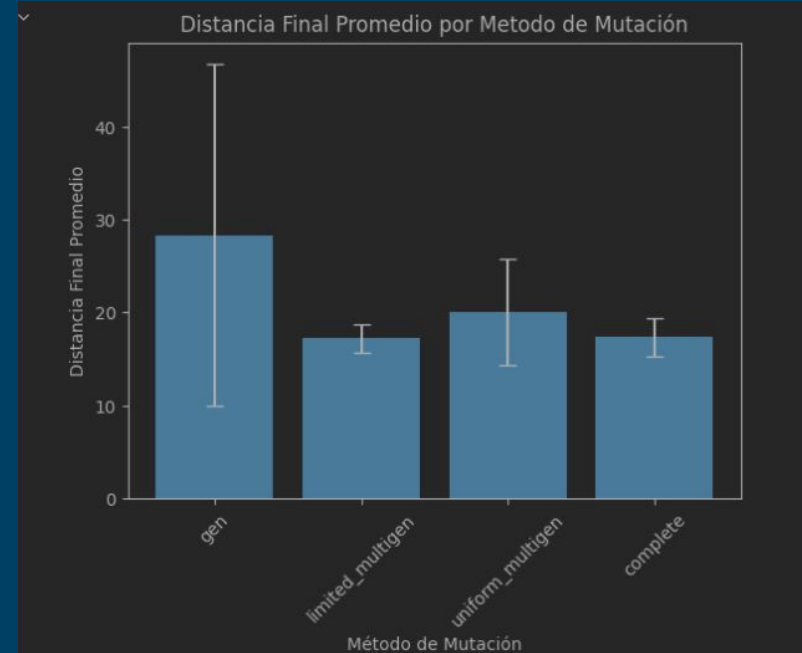
- Probabilidad de mutar: 0.2
- Porcentaje de cambio: 5%

Población Máxima: 20 individuos

Selección: Ruleta

Cruza: Uniforme

Al disminuir las mutaciones se agrava la situación del gráfico anterior para el método “gen”. La distancia final depende mucho más de los genes iniciales que en los otros casos.



Gráficos

Comparación Métodos de Mutación

Cantidad de repeticiones: 10

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generación (**100**) y distancia (0.4)

Parámetros de mutación:

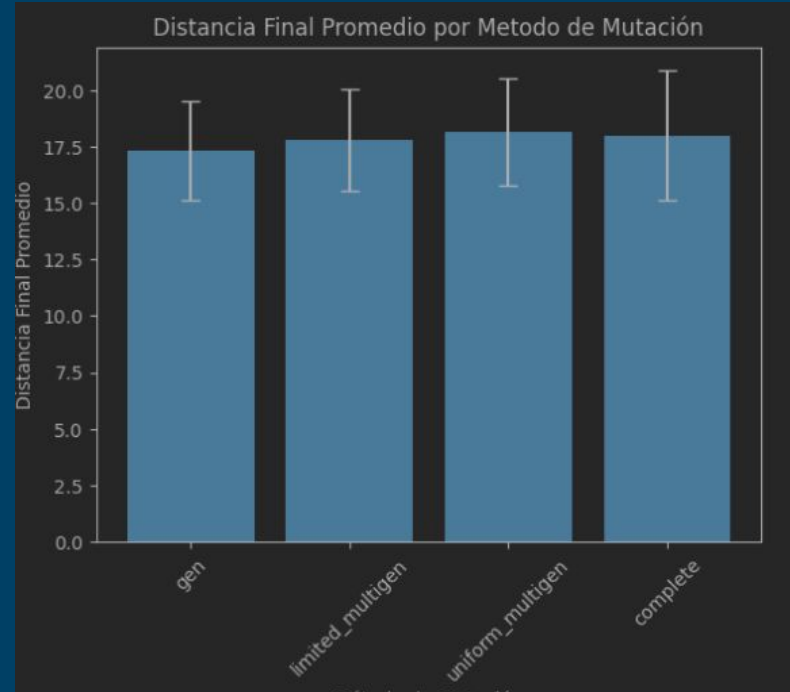
- **Probabilidad de mutar: 0.2**
- **Porcentaje de cambio: 20%**

Población Máxima: 20 individuos

Selección: Ruleta

Cruza: Uniforme

Ahora tomamos la misma probabilidad de mutar, pero aumentamos la cantidad de cambios una vez que ocurre. Vemos que el "gen" se estabiliza relativo a los otros y tenemos menos variación en los resultados entre los distintos métodos



Gráficos

Comparación Métodos de Mutación

Cantidad de repeticiones: 10

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generación (**100**) y distancia (0.4)

Parámetros de mutación:

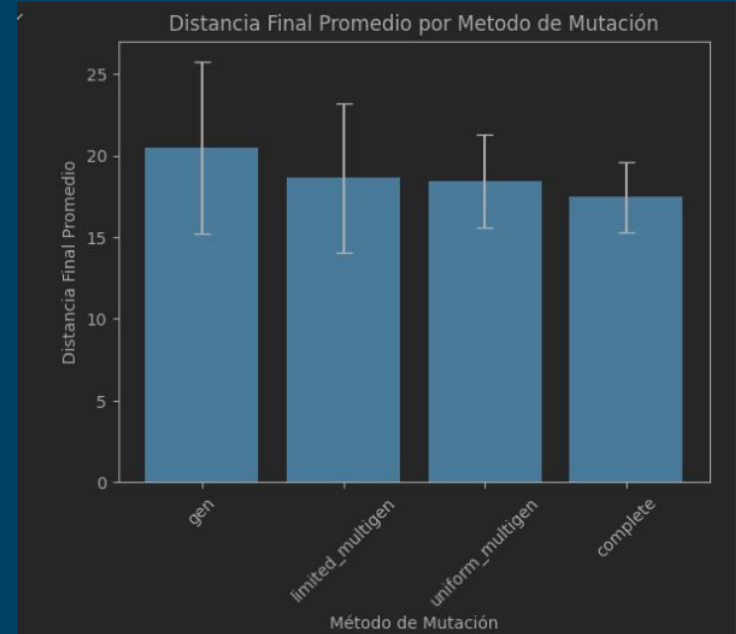
- **Probabilidad de mutar: 0.2**
- **Porcentaje de cambio: 50%**

Población Máxima: 20 individuos

Selección: Ruleta

Cruza: Uniforme

Por otro lado, al incrementar mucho el porcentaje de cambio al algoritmo le es difícil mantenerse cerca de una solución, pues los individuos varían demasiado. Esto se podría evitar reduciendo la cantidad de mutación a medida que avanzan las generaciones.



Gráficos

Comparación Métodos de Mutación

Cantidad de repeticiones: 10

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generación (**100**) y distancia (0.4)

Parámetros de mutación:

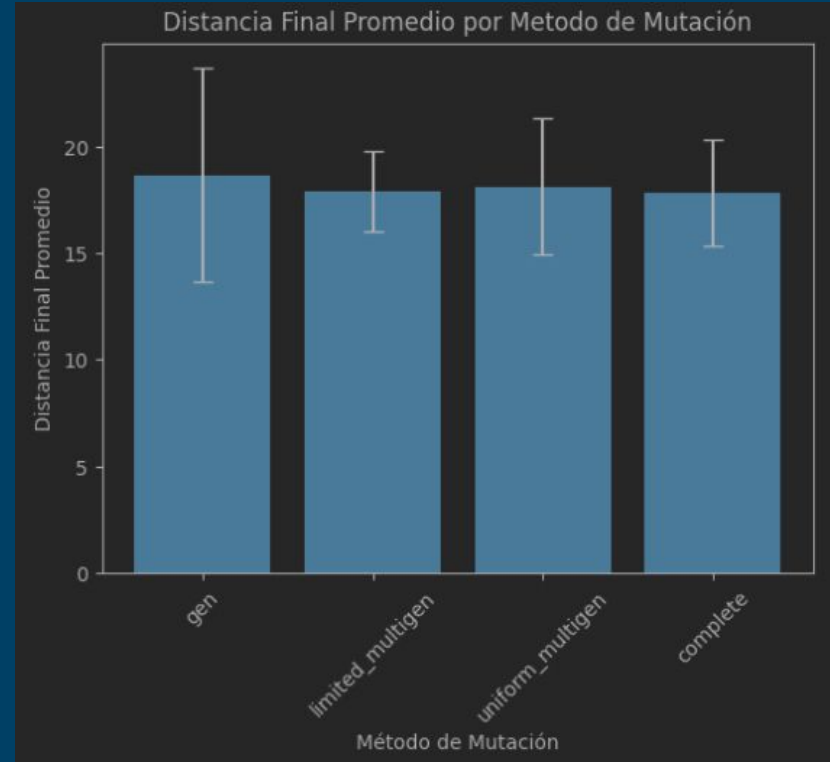
- **Probabilidad de mutar: 1**
- **Porcentaje de cambio: 5%**

Población Máxima: 20 individuos

Selección: Ruleta

Cruza: Uniforme

Por último, probamos mutar levemente pero hacerlo siempre. Vemos una repetición de dos casos atrás, cuando los métodos tampoco se diferenciaban significativamente.



Ejemplo de Solución

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generación (100) y distancia (0.4)

Parámetros de mutación:

- Multigen Limitado
- Probabilidad de mutar: 0.2
- Porcentaje de cambio: 5%

Población Máxima: 20 individuos

Selección: Ruleta

Cruza: Uniforme

Target

Mejor Solución

Mejor distancia: 15.61

Generation 100:									
IND9:		C0:	60.09%	C1:	24.43%	C2:	15.48%	DIST:	15.61
IND3:		C0:	66.71%	C1:	21.94%	C2:	11.35%	DIST:	20.07
IND4:		C0:	54.99%	C1:	32.66%	C2:	12.35%	DIST:	22.52
IND10:		C0:	48.31%	C1:	28.7%	C2:	22.99%	DIST:	24.44
IND1:		C0:	68.72%	C1:	23.52%	C2:	7.77%	DIST:	25.43
IND15:		C0:	64.22%	C1:	17.24%	C2:	18.53%	DIST:	25.52
IND17:		C0:	63.7%	C1:	17.09%	C2:	19.21%	DIST:	26.21
IND6:		C0:	49.59%	C1:	23.76%	C2:	26.65%	DIST:	28.35
IND5:		C0:	74.03%	C1:	16.88%	C2:	9.09%	DIST:	29.8
IND18:		C0:	73.24%	C1:	12.46%	C2:	14.3%	DIST:	33.81
IND11:		C0:	79.37%	C1:	9.52%	C2:	11.11%	DIST:	40.89
IND19:		C0:	82.15%	C1:	9.16%	C2:	8.69%	DIST:	43.58
IND2:		C0:	78.78%	C1:	6.98%	C2:	14.24%	DIST:	45.27
IND12:		C0:	73.71%	C1:	6.53%	C2:	19.75%	DIST:	47.75
IND16:		C0:	71.41%	C1:	28.59%	C2:	0.0%	DIST:	52.37
IND14:		C0:	66.84%	C1:	33.16%	C2:	0.0%	DIST:	53.39
IND8:		C0:	81.43%	C1:	18.57%	C2:	0.0%	DIST:	56.01
IND0:		C0:	68.38%	C1:	3.87%	C2:	27.75%	DIST:	59.33
IND7:		C0:	86.03%	C1:	13.97%	C2:	0.0%	DIST:	60.1
IND13:		C0:	100.0%	C1:	0.0%	C2:	0.0%	DIST:	78.84

Ejemplo de Solución

Colores objetivo: Naranja Pastel (247, 177, 149)

Paleta: RGB

Parámetros de corte:

- por tiempo (5s), generación (100) y distancia (0.4)

Parámetros de mutación:

- Multigen Limitado
- Probabilidad de mutar: 0.2
- Porcentaje de cambio: 5%

Población Máxima: 20 individuos

Selección: **Elite**

Cruza: Uniforme

Perdemos la diversidad en la generación 15, pero llegamos a una levemente mejor solución.

Target

Mejor Solución

Mejor distancia: 15.33

Generation 15:									
IND0:		C0:	59.24%	C1:	25.43%	C2:	15.33%	DIST:	15.35
IND1:		C0:	59.24%	C1:	25.43%	C2:	15.33%	DIST:	15.35
IND2:		C0:	59.28%	C1:	25.88%	C2:	14.84%	DIST:	15.35
IND3:		C0:	59.2%	C1:	25.44%	C2:	15.36%	DIST:	15.35
IND4:		C0:	58.83%	C1:	25.69%	C2:	15.48%	DIST:	15.36
IND5:		C0:	58.83%	C1:	25.69%	C2:	15.48%	DIST:	15.36
IND6:		C0:	59.26%	C1:	25.92%	C2:	14.82%	DIST:	15.36
IND7:		C0:	59.32%	C1:	25.9%	C2:	14.78%	DIST:	15.36
IND8:		C0:	59.5%	C1:	25.77%	C2:	14.74%	DIST:	15.36
IND9:		C0:	59.5%	C1:	25.77%	C2:	14.74%	DIST:	15.36
IND10:		C0:	59.5%	C1:	25.77%	C2:	14.74%	DIST:	15.36
IND11:		C0:	58.65%	C1:	26.0%	C2:	15.35%	DIST:	15.36
IND12:		C0:	59.64%	C1:	25.63%	C2:	14.73%	DIST:	15.36
IND13:		C0:	59.64%	C1:	25.63%	C2:	14.73%	DIST:	15.36
IND14:		C0:	58.52%	C1:	26.08%	C2:	15.4%	DIST:	15.38
IND15:		C0:	58.52%	C1:	26.08%	C2:	15.4%	DIST:	15.38
IND16:		C0:	58.52%	C1:	26.08%	C2:	15.4%	DIST:	15.38
IND17:		C0:	59.82%	C1:	25.57%	C2:	14.61%	DIST:	15.38
IND18:		C0:	59.82%	C1:	25.57%	C2:	14.61%	DIST:	15.38
IND19:		C0:	59.82%	C1:	25.57%	C2:	14.61%	DIST:	15.38

Generation 100:									
IND0:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND1:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND2:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND3:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND4:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND5:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND6:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND7:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND8:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND9:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND10:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND11:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND12:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND13:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND14:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND15:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND16:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND17:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND18:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33
IND19:		C0:	59.12%	C1:	25.7%	C2:	15.18%	DIST:	15.33

Conclusiones

¿Qué es lo que más afecta este problema?

- Se puede destacar que el hiperparámetro más determinante para este problema fue el método de selección
 - La paleta de colores afecta críticamente la capacidad de acercarse y/o alcanzar el color objetivo, independientemente del resto de los parámetros y algoritmos utilizados.
 - El método de selección Elite siempre brindó soluciones muy óptimas, al costo de eliminar por completo la diversidad. Por el contrario, el método de Ruleta también brindó buenas soluciones si se le permite más generaciones, realizando exploración más diversa.
-

Conclusiones

Solución rápida
vs
Solución óptima

El problema de elegir “Elite” contra otro método de selección se parece al problema de si conviene usar un DFS o BFS al querer encontrar una solución.

“Elite”, análogo a DFS, encuentra una solución de forma rápida, sin embargo, a diferencia de BFS, no tiende a ser capaz de encontrar soluciones óptimas por falta de diversidad genética.

Conclusiones

¿Qué tomaríamos para encontrar la mejor solución a este problema?

Método de selección Elite, que combinado con una población inicial variada y una buena cantidad de mutación es capaz evitar caer en máximos/mínimos locales, así encontrando soluciones de forma no sólo rápida, sino también óptima.

Muchas Gracias!

Preguntas?