

Trabajo Práctico Especial 3

Algoritmos Genéticos - Defensor 1

Grupo 8

Colloca, Tomás

López, Noelia Belén

Suárez Bodean, Joaquín

Vera, Juan Sebastián

Indice

Grupo 8	1
Indice	2
Introducción	3
Modelo	4
Métodos de selección	4
Función de la temperatura para Boltzmann	5
Métodos de cruza	6
Métodos de mutación	6
Métodos de reemplazo	6
Métodos de corte	7
Jugador	8
Resultados	9
Análisis de métodos de selección	10
Parámetros de Boltzmann	11
Análisis de métodos de cruza	14
Análisis de mutación	16
Análisis de métodos de reemplazo	17
Análisis de los parámetros N y k	19
Análisis de otras configuraciones	21
Individuo más apto	24
Conclusiones	25

Introducción

Se implementó un programa que, utilizando algoritmos genéticos, busca una solución aproximada e idealmente óptima para problemas que tienen varias variables y cada una de ellas con distintos valores.

En particular, se utilizó el programa realizado para buscar la configuración óptima de un jugador de un videojuego el cual tiene una altura que debe pertenecer a un rango dado, y distintos posibles equipamientos para distintas partes del cuerpo, cada una de ellas afectando indirectamente a los dos parámetros principales del jugador: ataque y fuerza. De esta forma, y a partir del tipo de jugador, se buscó la mejor configuración de altura y equipamientos para el mismo.

Modelo

El programa realizado en Java consiste de una clase principal `GeneticAlgorithms` que se ocupa de correr el algoritmo. La misma recibe un generador aleatorio de individuos, que debe implementar la interfaz `IndividualFactory`, con el cual se crea una población inicial de tamaño N . Los individuos a su vez deben extender de `Individual`. Cada individuo tiene un listado de genes, representados por la clase `Gene`. A su vez, cada gen tiene una referencia al conjunto de sus posibles alelos, modelado con la interfaz `AlleleSet`.

De esta forma, para poder aplicar el programa a un problema, sólo se debe implementar un generador de individuos aleatorio, la función de fitness de los individuos, y los conjuntos de alelos.

Respecto al algoritmo, existen clases que contienen las distintas variantes de métodos de selección, cruce, mutación, reemplazo y condición de corte. Algunos de estos métodos reciben el listado de individuos, mientras que otros la población (`Population`) que contiene además una referencia a la generación actual.

Métodos de selección

Los métodos de selección implementados son los siguientes:

- **Elite:** Se toman los k individuos con fitness más alto.
- **Random:** Se toman k individuos al azar.
- **Ruleta:** Se toman k individuos probabilísticamente a partir de la frecuencia relativa de fitness de cada individuo, utilizando k números aleatorios.
- **Universal:** Se toman k individuos probabilísticamente a partir de la frecuencia relativa del *fitness* de cada individuo a partir de un único número aleatorio.
- **Boltzmann:** Se toman k individuos utilizando una distribución de Boltzmann donde la energía está representada por el fitness del individuo, y la temperatura es una función exponencial decreciente en función de la generación t , dada por la ecuación:

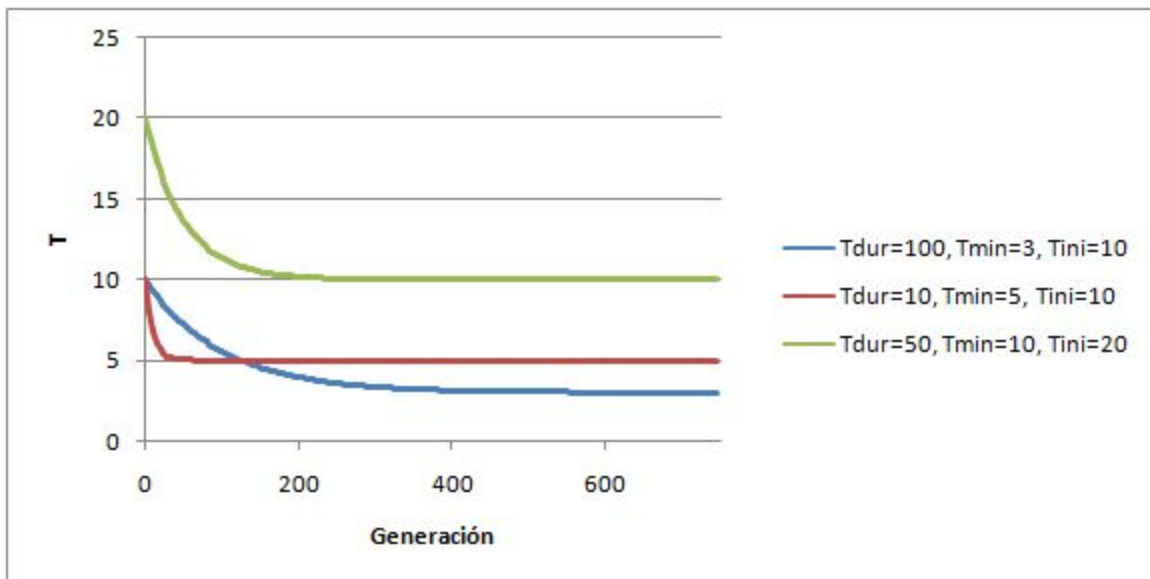
$$T(t) = e^{-1/T_{duration} * t + \ln(T_{initial} - T_{min})} + T_{min} .$$

Notar que $T(0) = T_{initial}$ y que $\lim_{t \rightarrow +\infty} T(t) = T_{min}$, y por otro lado cuanto mayor es $T_{duration}$ más tarda en converger a T_{min} .

- **Torneo determinístico:** Se toman k veces subconjuntos de m individuos y se elige siempre al de mejor fitness.
- **Torneo probabilístico:** Se toman k veces subconjuntos de m individuos y se elige con una probabilidad de 0.75 al de mejor fitness.
- **Ranking:** Se toman k individuos donde cada uno tiene una probabilidad $p = \frac{N-i}{S}$ donde N es la cantidad de individuos, i es la posición del individuo según su fitness (1 es el valor de i para el individuo con mayor fitness, y $N - 1$ para el de menor fitness), y $S = \sum_{k=1}^N k$.

Función de la temperatura para Boltzmann

En la siguiente Figura se puede ver la evolución de la temperatura según los distintos parámetros mencionados anteriormente:



Métodos de cruce

Los métodos de cruce implementados son los siguientes:

- **Un punto:** Se realiza la cruce a partir de un gen elegido aleatoriamente.
- **Dos puntos:** Se realiza la cruce entre dos genes elegidos aleatoriamente.
- **Anular:** Se realiza la cruce a partir de un gen y con una longitud l , ambos elegidos aleatoriamente. Si la longitud es mayor a la longitud del cromosoma L , se aplica modulo L .
- **Uniforme:** Se realiza la cruce gen por gen con una probabilidad $p_{uniforme}$.

La cruce además se realiza con una probabilidad p_{cruce} .

Métodos de mutación

Los métodos de mutación implementados son los siguientes:

- **Un único gen:** Se muta un único gen elegido aleatoriamente con una probabilidad $p_{mutación}$.
- **Multigen:** Cada gen tiene una probabilidad $p_{mutación}$ de ser mutado.

Métodos de reemplazo

Los métodos de reemplazo implementados son los siguientes:

- **Método 1:** Se toman $\frac{N}{2}$ parejas siendo N el total de individuos, y los 2 hijos de cada pareja pasan a la siguiente generación.
- **Método 2:** Se utiliza brecha generacional con $G = \frac{k}{N}$.
- **Método 3:** $N - k$ individuos pasan a la próxima generación sin modificación. Se obtienen k individuos, se los empareja de dos en dos y se obtienen k hijos. De la población en la generación actual (N individuos) junto con los k hijos (en total $N + k$), se toman k individuos que pasan a la siguiente generación.

- **Brecha generacional:** A partir del parámetro G , se obtiene $k = G * N$. Luego se obtienen k individuos, se los empareja de dos en dos y se obtienen k hijos que pasan a la siguiente generación. De los N individuos de la generación actual, se toman $N - k$ que pasan a la siguiente generación.

Cabe destacar que a la hora de seleccionar los individuos inicialmente y luego al realizar el reemplazo, se pueden combinar hasta dos métodos de selección con porcentajes $100 * p_{selección}$ y $100 * (1 - p_{selección})$; y $100 * p_{reemplazo}$ y $100 * (1 - p_{reemplazo})$, respectivamente.

Métodos de corte

Los métodos de corte implementados son los siguientes:

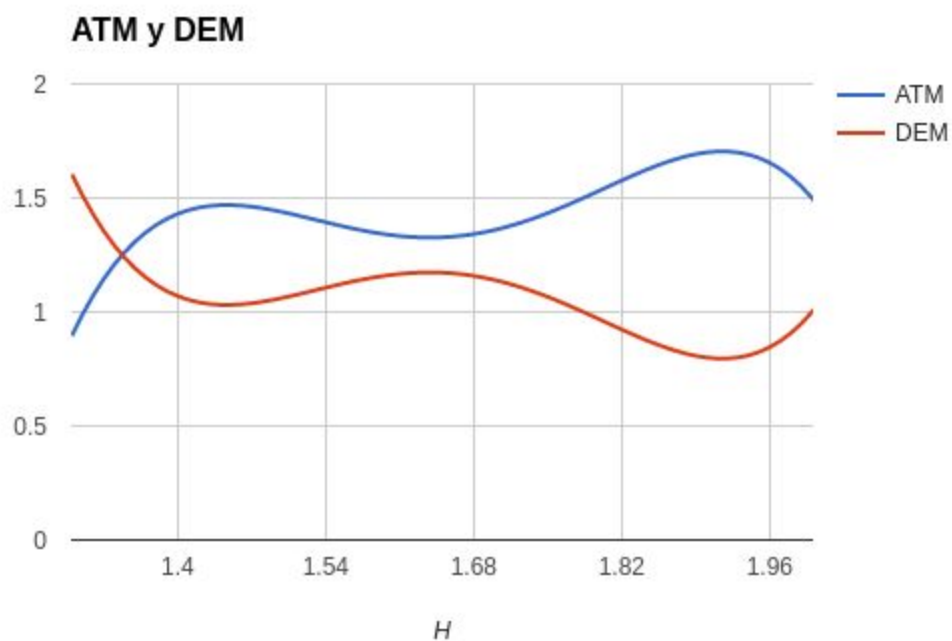
- **Cantidad máxima de generaciones:** Límite en la cantidad de generaciones a correr.
- **Fitness mínimo:** Fitness mínimo al que se debe alcanzar antes de cortar el algoritmo.
- **Estructura:** Porcentaje de individuos con el mismo genotipo que se encuentran en la generación $t - 1$ y t .
- **Contenido:** Cantidad máxima de generaciones donde el máximo fitness es el mismo.

Cabe destacar que se pueden elegir uno o más de dichos métodos, y en particular, se puede elegir aplicarlos todos.

Jugador

Tocó aplicar el algoritmo para el Defensor 1. El mismo tiene multiplicadores de defensa y ataque 0.9 y 0.1 respectivamente. Por esta razón, se espera que el jugador priorice defensa sobre ataque.

Este es el gráfico de los multiplicadores ATM y DEM en función de la altura. A partir de lo dicho anteriormente, se podría esperar que la altura del jugador que arroje el algoritmo sea cercana a 1.3m, pues es donde la función DEM alcanza su máximo.



Resultados

Los siguientes resultados se realizaron con la siguiente configuración default:

$N = 100$, $k = 50$

Métodos de selección: 50% Elite y 50% Roulette.

Parámetros para Boltzmann: $T_{duration} = 10$, $T_{min} = 3$, $T_{initial} = 10$.

Parámetro de torneo: $m = 2$.

Método de cruce ($p_{cruza} = 0.775$): uniforme con $p_{uniforme} = 0.5$.

Método de mutación: multigen con $p_{mutación} = 0.005$.

Método de reemplazo: Método 2 con 50% Elite y 50% Roulette a la hora de seleccionar.

Parámetro G de brecha generacional: 0.7.

Métodos de corte: máximo de 5000 generaciones, fitness mínimo de 28.08 y contenido = 1000 generaciones.

Además, se corrió el algoritmo para los archivos de prueba con 10 equipamientos distintos.

Análisis de métodos de selección

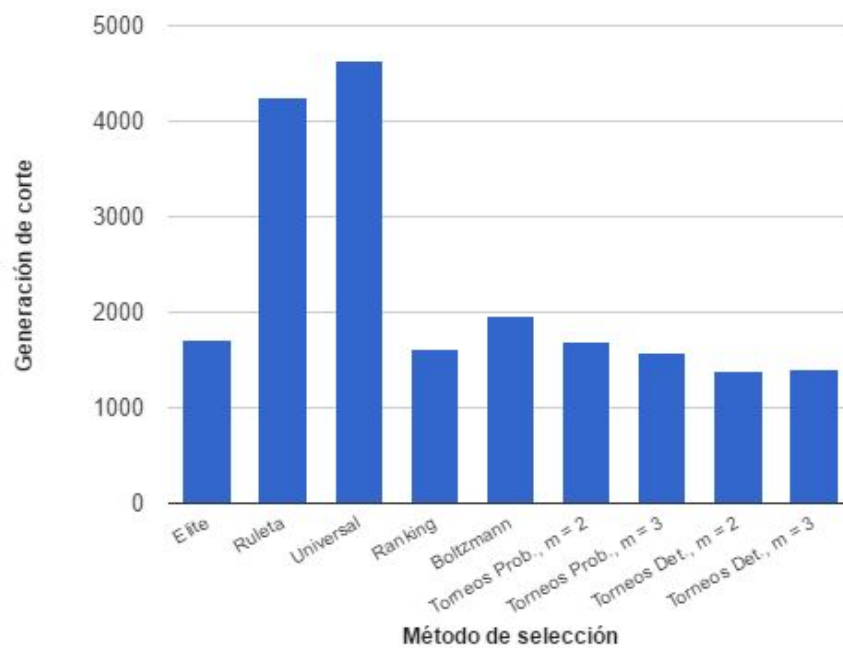


Figura 1: Gráfico del promedio de la generación de corte sobre 10 poblaciones aleatorias para distintos métodos de selección.

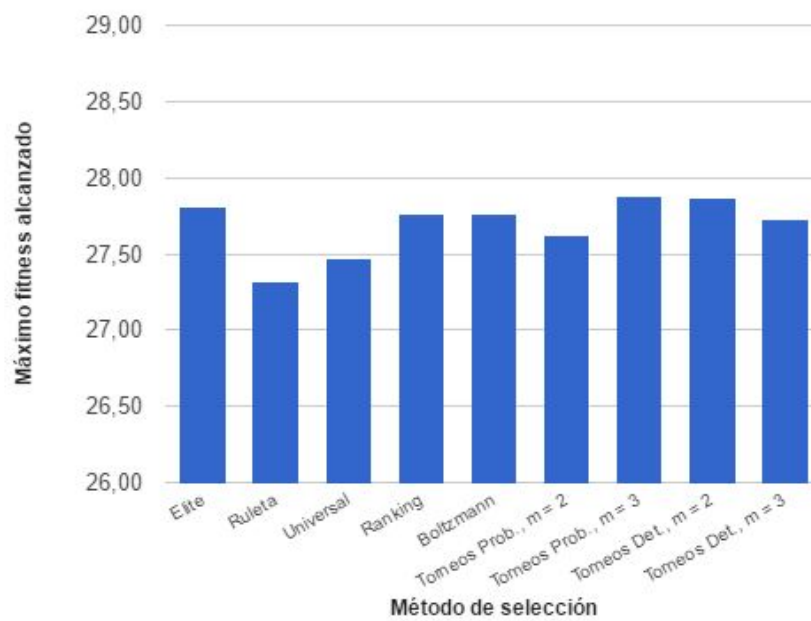


Figura 2: Gráfico del promedio del máximo fitness alcanzado sobre 10 poblaciones aleatorias para distintos métodos de selección.

Parámetros de Boltzmann

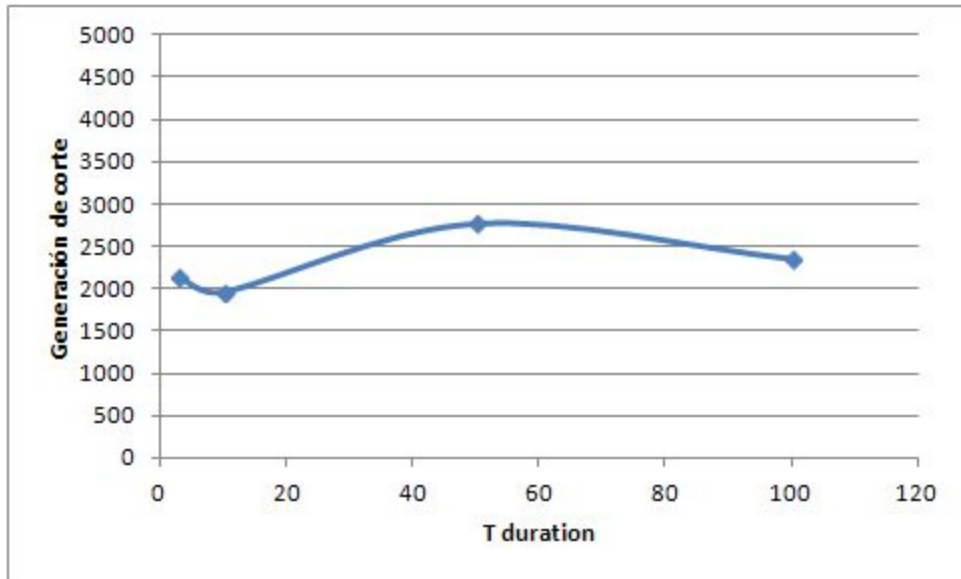


Figura 3: Gráfico del promedio de la generación de corte sobre 10 poblaciones aleatorias utilizando el método de selección de Boltzmann en función de $T_{duration}$.

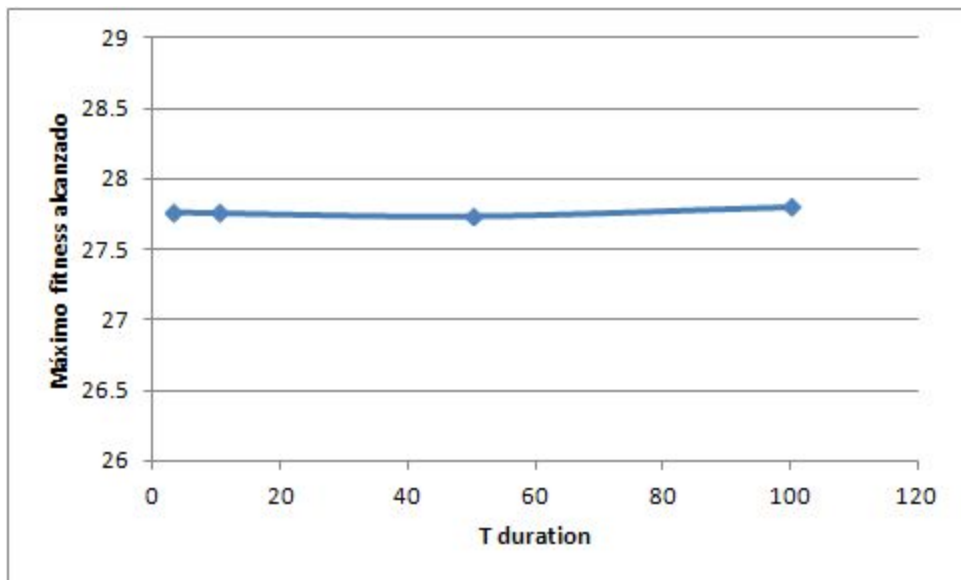


Figura 4: Gráfico del promedio del máximo fitness alcanzado sobre 10 poblaciones aleatorias utilizando el método de selección de Boltzmann en función de $T_{duration}$.

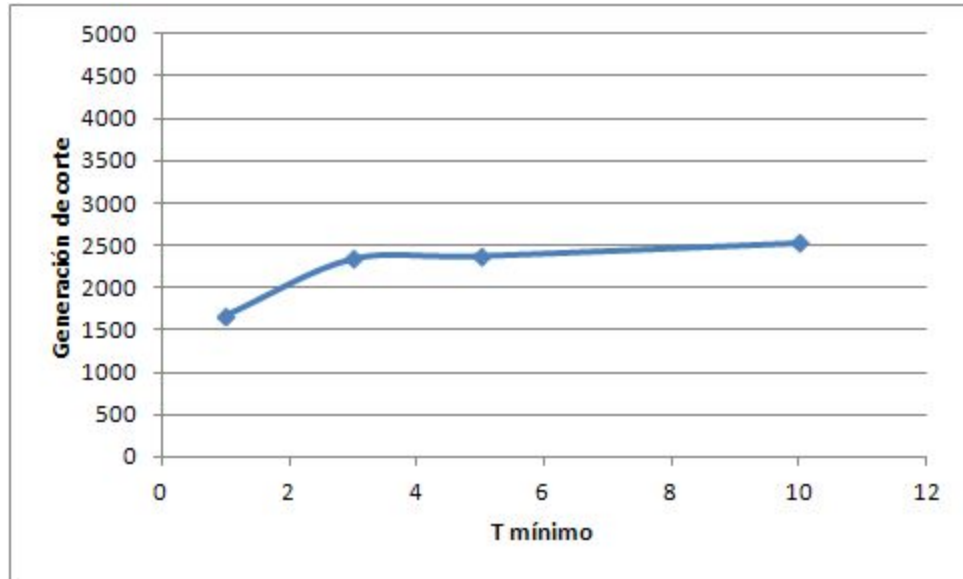


Figura 5: Gráfico del promedio de la generación de corte sobre 10 poblaciones aleatorias utilizando el método de selección de Boltzmann en función de la temperatura mínima a la que converge la función de la temperatura.

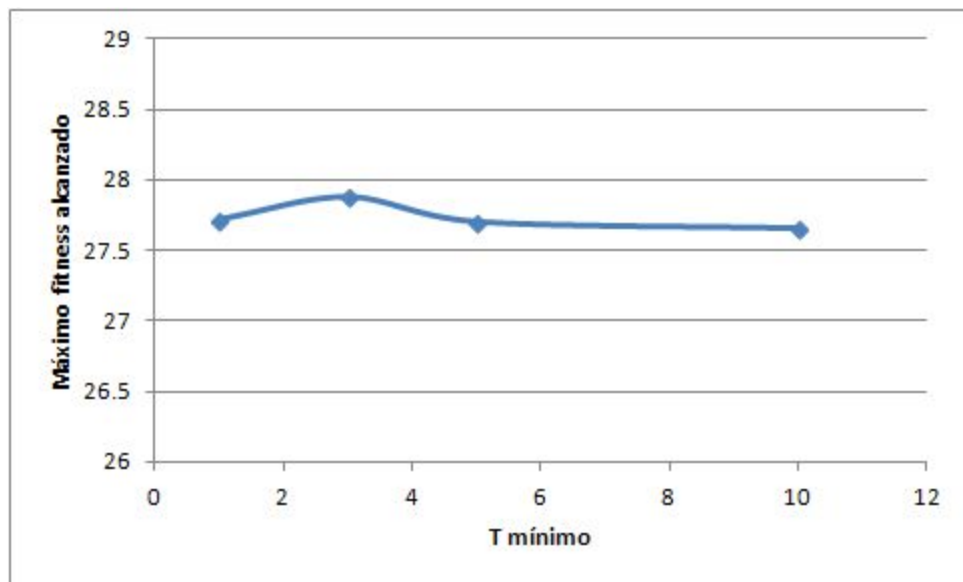


Figura 6: Gráfico del promedio del máximo fitness alcanzado sobre 10 poblaciones aleatorias utilizando el método de selección de Boltzmann en función de la temperatura mínima a la que converge la función de la temperatura.

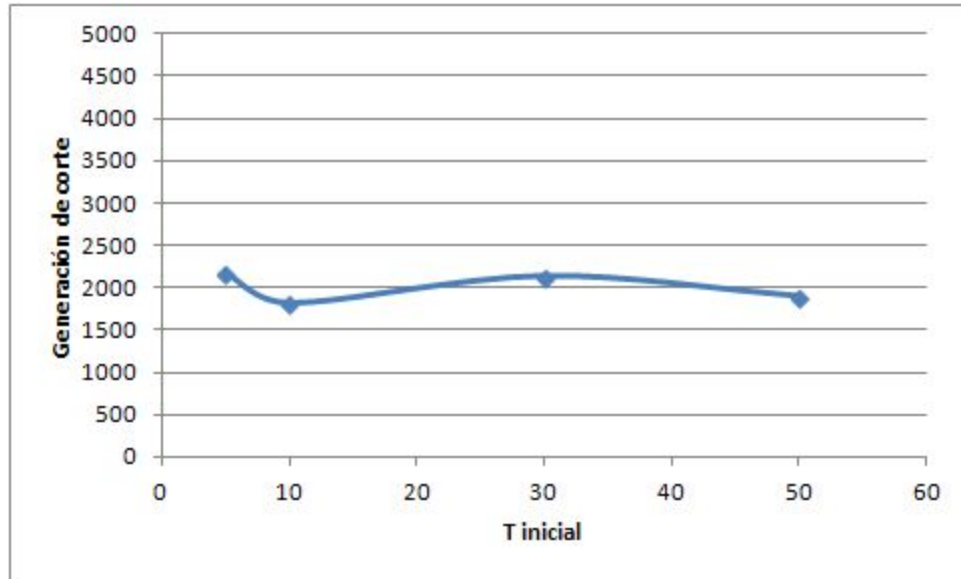


Figura 7: Gráfico del promedio de la generación de corte sobre 10 poblaciones aleatorias utilizando el método de selección de Boltzmann en función de la temperatura inicial.

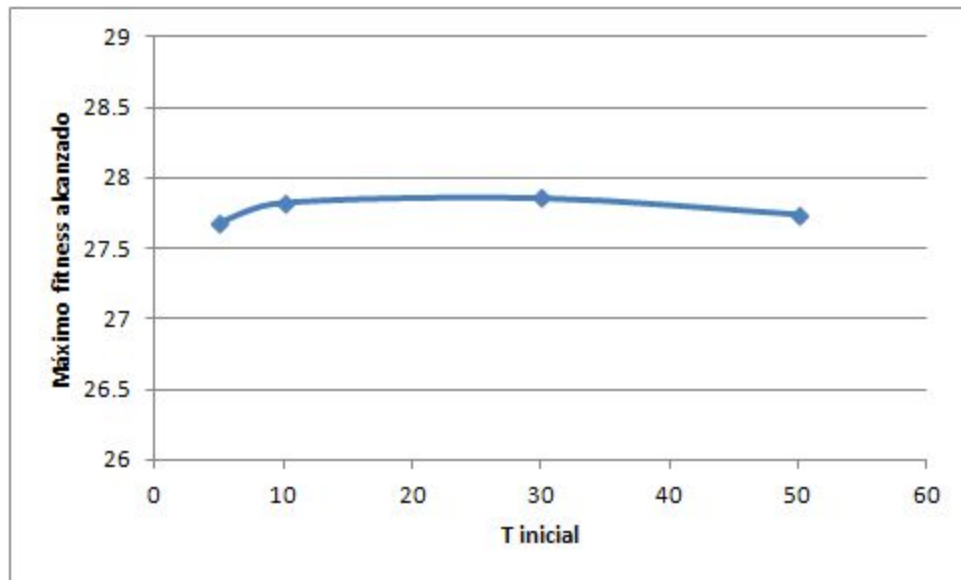


Figura 8: Gráfico del promedio del máximo fitness alcanzado sobre 10 poblaciones aleatorias utilizando el método de selección de Boltzmann en función de la temperatura inicial.

Análisis de métodos de cruce

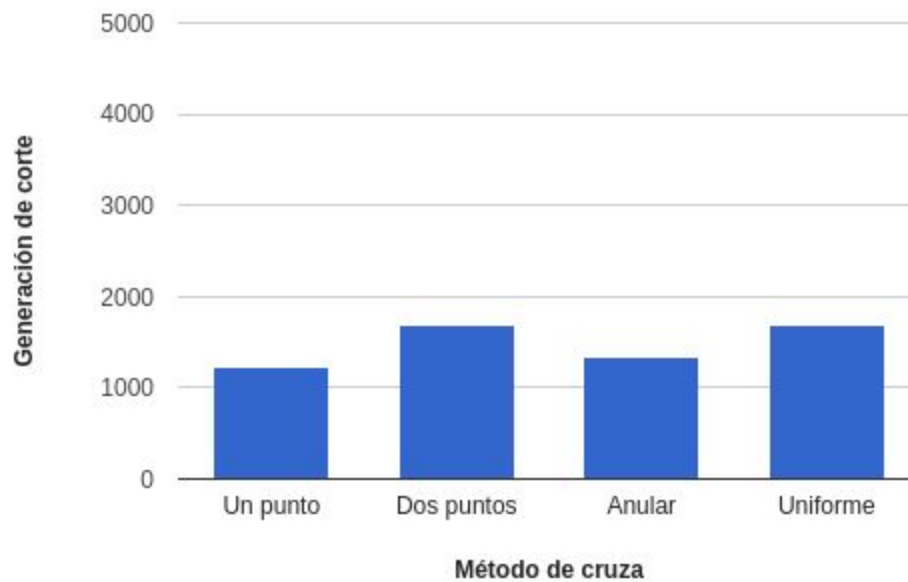


Figura 9: Gráfico del promedio de la generación de corte sobre 10 poblaciones aleatorias para distintos métodos de cruce.

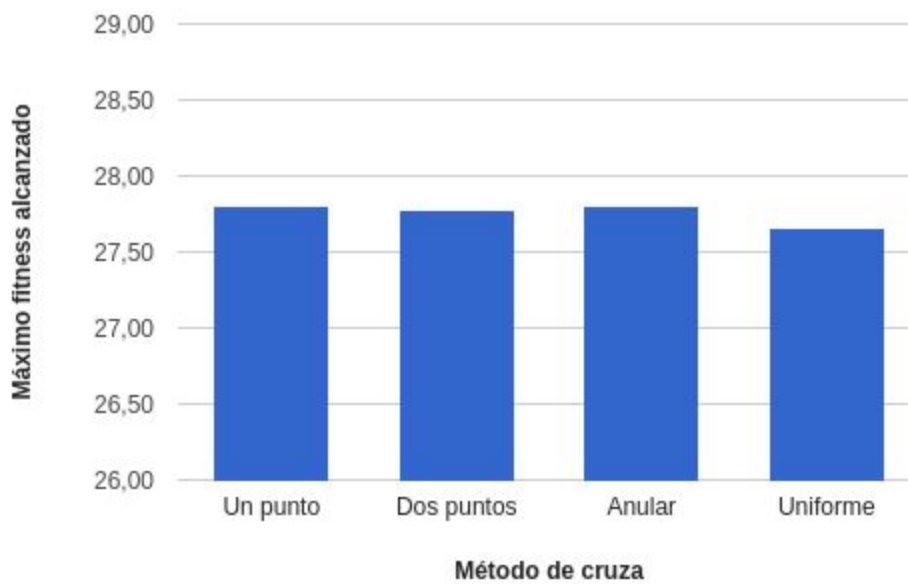


Figura 10: Gráfico del promedio del máximo fitness alcanzado sobre 10 poblaciones aleatorias para distintos métodos de cruce.

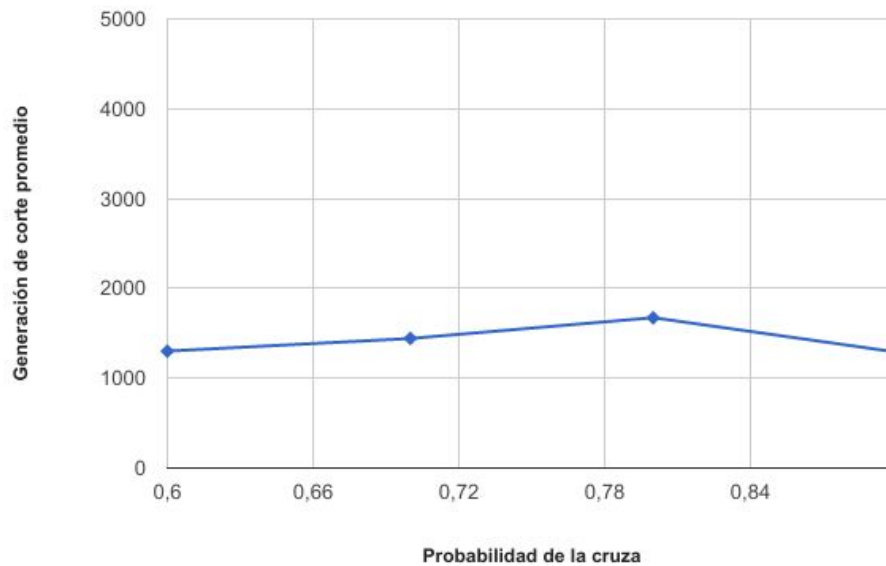


Figura 11: Gráfico del promedio de la generación de corte sobre 10 poblaciones aleatorias en función de la probabilidad de cruza.

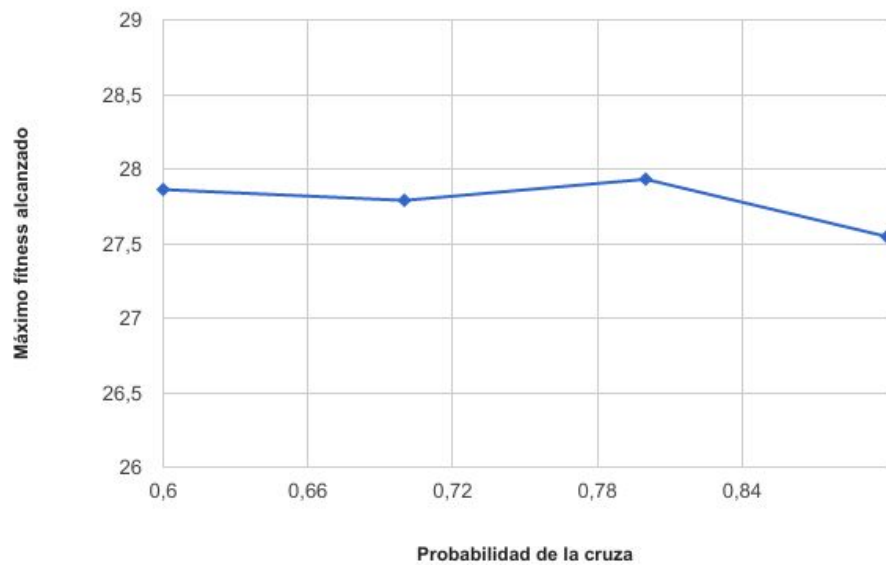


Figura 12: Gráfico del promedio del máximo fitness alcanzado sobre 10 poblaciones aleatorias en función de la probabilidad de cruza.

Análisis de mutación

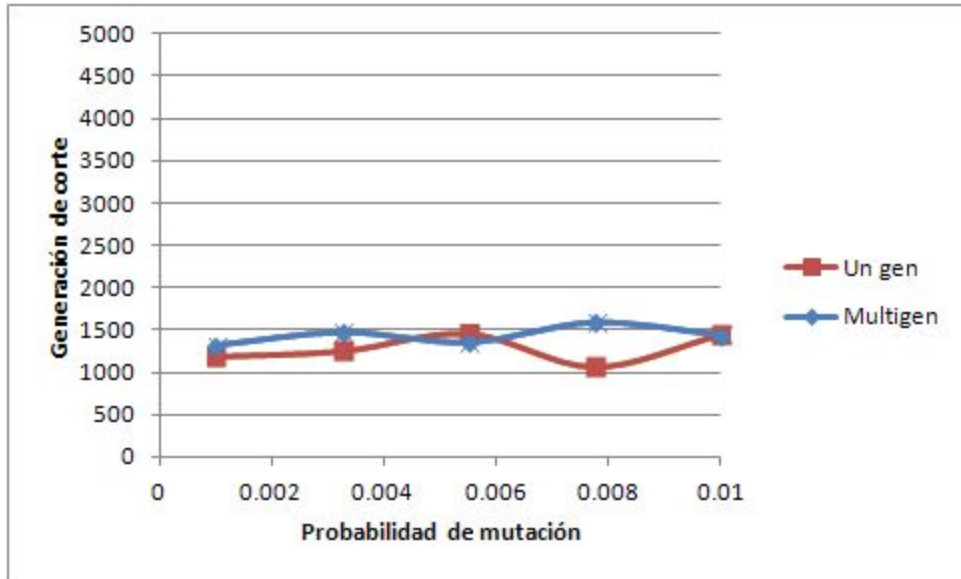


Figura 13: Gráfico del promedio de la generación de corte sobre 10 poblaciones aleatorias en función de la probabilidad de mutación y el método de mutación.

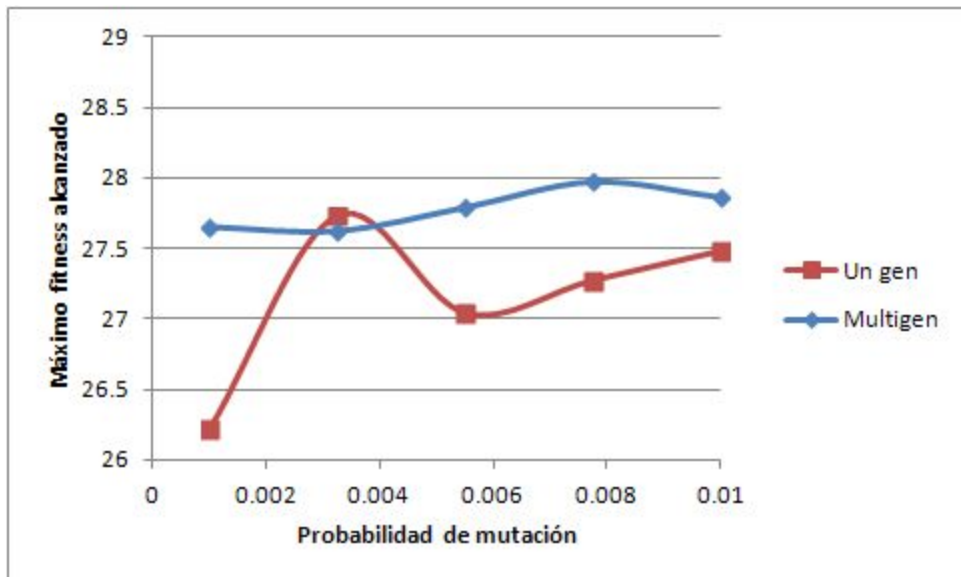


Figura 14: Gráfico del promedio del máximo fitness alcanzado sobre 10 poblaciones aleatorias en función de la probabilidad de mutación y el método de mutación.

Análisis de métodos de reemplazo

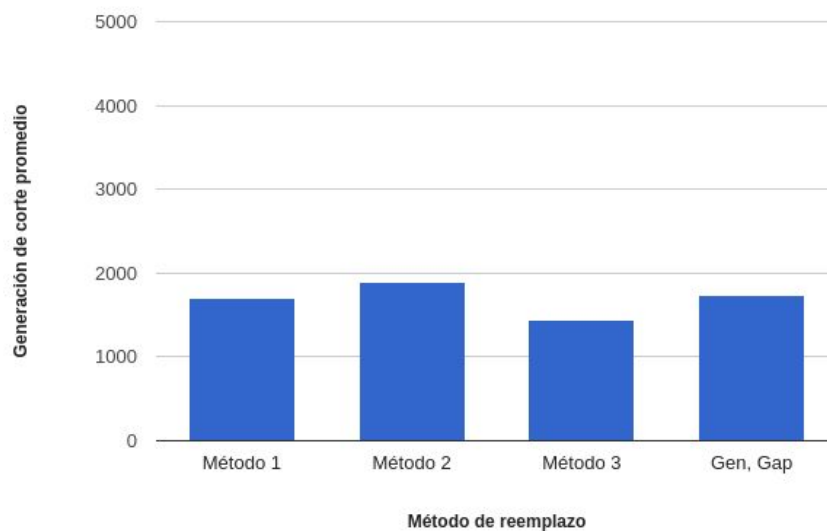


Figura 15: Gráfico del promedio de la generación de corte sobre 10 poblaciones aleatorias para distintos métodos de reemplazo.

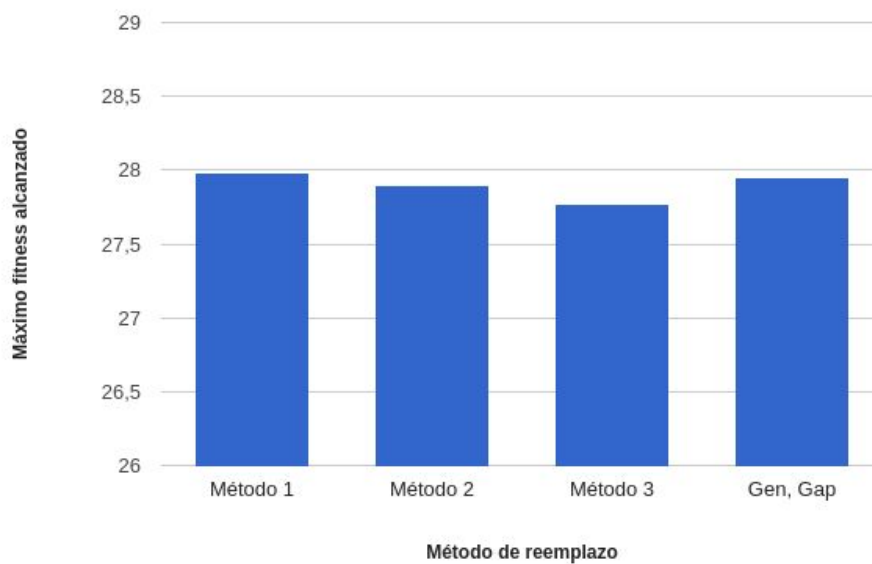


Figura 16: Gráfico del promedio del máximo fitness alcanzado sobre 10 poblaciones aleatorias para distintos métodos de reemplazo.

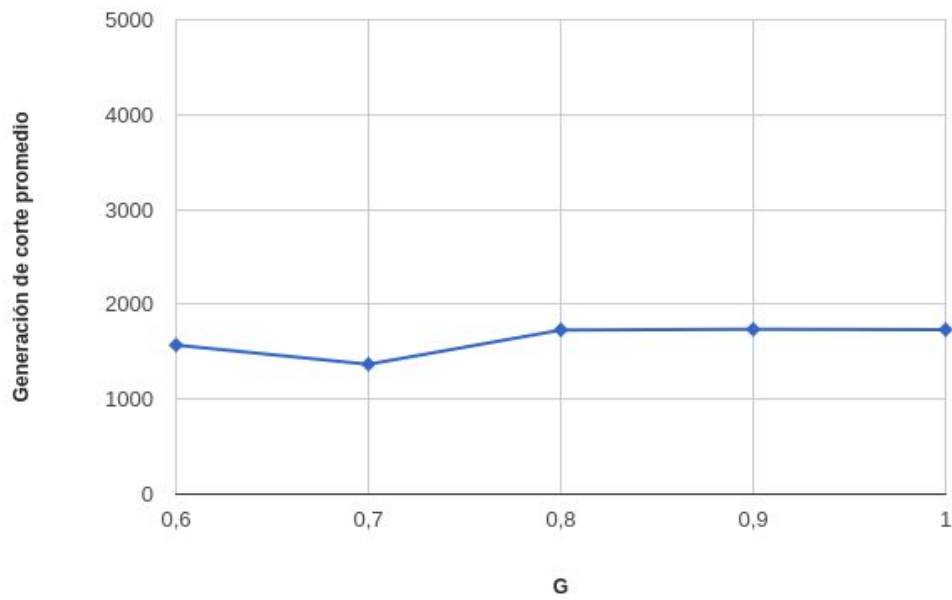


Figura 17: Gráfico del promedio de la generación de corte en función de G sobre 10 poblaciones aleatorias utilizando brecha generacional en función de G .

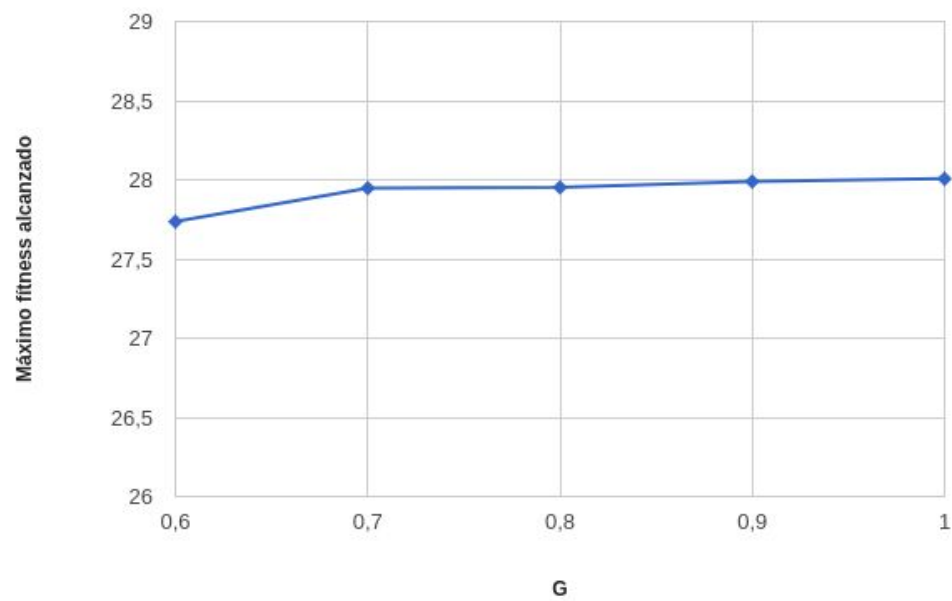


Figura 18: Gráfico del promedio del máximo fitness alcanzado sobre 10 poblaciones aleatorias utilizando brecha generacional en función de G .

Análisis de los parámetros N y k

Los siguientes resultados se corrieron con los archivos que pedía la cátedra con la siguiente configuración (variando solo N y k) la cual se decidió a partir de los resultados anteriores:

Métodos de selección: 50% Elite y 50% Ranking.

Método de cruce ($p_{\text{cruza}} = 0.8$): Anular.

Método de mutación: multigen con $p_{\text{mutación}} = 0.01$.

Método de reemplazo: Método 3 con 70% Elite y 30% Ranking a la hora de seleccionar.

Métodos de corte: máximo de 10000 generaciones, y contenido = 1000 generaciones.

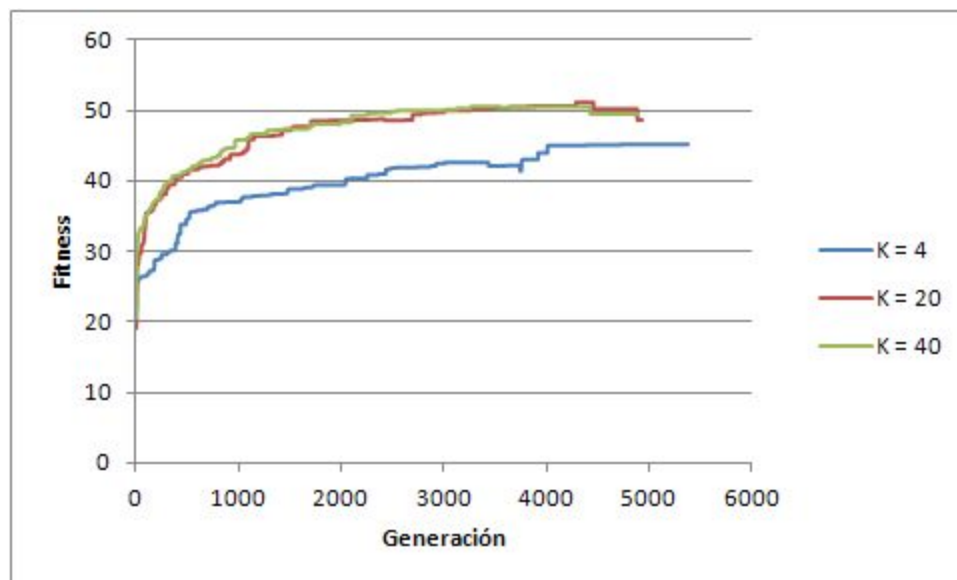


Figura 19: Gráfico del promedio del fitness sobre 5 poblaciones aleatorias en función del paso de las generaciones para $N = 50$.

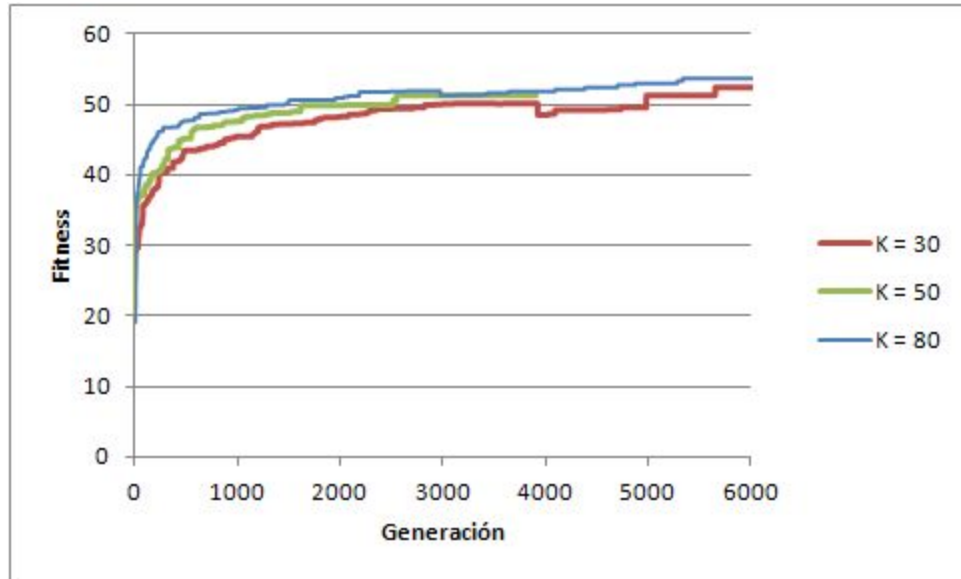


Figura 20: Gráfico del promedio del fitness sobre 5 poblaciones aleatorias en función del paso de las generaciones para $N = 100$.

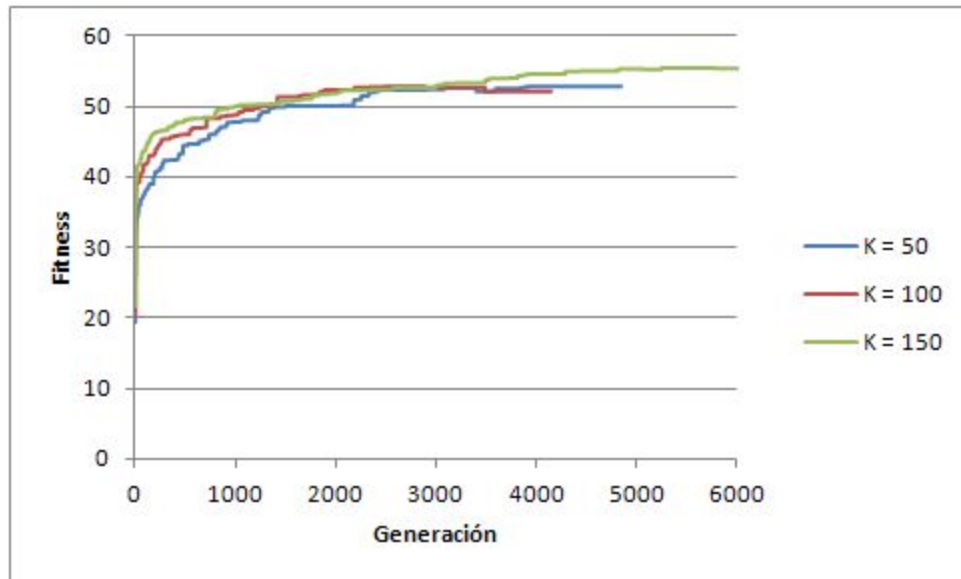


Figura 21: Gráfico del promedio del fitness sobre 5 poblaciones aleatorias en función del paso de las generaciones para $N = 200$.

Análisis de otras configuraciones

Los resultados a continuación referencian a las siguientes configuraciones:

Configuración 1:

$N = 100$, $k = 50$

Métodos de selección: 50% Boltzmann y 50% Ruleta.

Parámetros para Boltzmann: $T_{duration} = 10$, $T_{min} = 3$, $T_{initial} = 10$.

Método de cruce ($p_{cruza} = 0.8$): uniforme con $p_{uniforme} = 0.5$.

Método de mutación: Multigen con $p_{mutación} = 0.01$.

Método de reemplazo: Brecha generacional con $G = 0.7$ y con 70% Boltzmann y 30% Ruleta a la hora de seleccionar.

Métodos de corte: máximo de 10000 generaciones y contenido = 1000 generaciones.

Configuración 2:

$N = 100$, $k = 50$

Métodos de selección: 30% Elite y 70% Torneo probabilístico.

Parámetro de torneo: $m = 3$.

Método de cruce ($p_{cruza} = 0.6$): Dos puntos.

Método de mutación: Un único gen con $p_{mutación} = 0.01$.

Método de reemplazo: Método 1 con 80% Elite y 20% Torneo probabilístico a la hora de seleccionar.

Métodos de corte: máximo de 10000 generaciones y contenido = 1000 generaciones.

Configuración 3:

$N = 100$, $k = 50$

Métodos de selección: 70% Universal y 30% Torneo determinístico.

Parámetro de torneo: $m = 2$.

Método de cruce ($p_{\text{cruza}} = 0.75$): Anular.

Método de mutación: Multigen con $p_{\text{mutación}} = 0.001$.

Método de reemplazo: Método 2 con 20% Universal y 80% Torneo determinístico a la hora de seleccionar.

Métodos de corte: máximo de 10000 generaciones y contenido = 1000 generaciones.

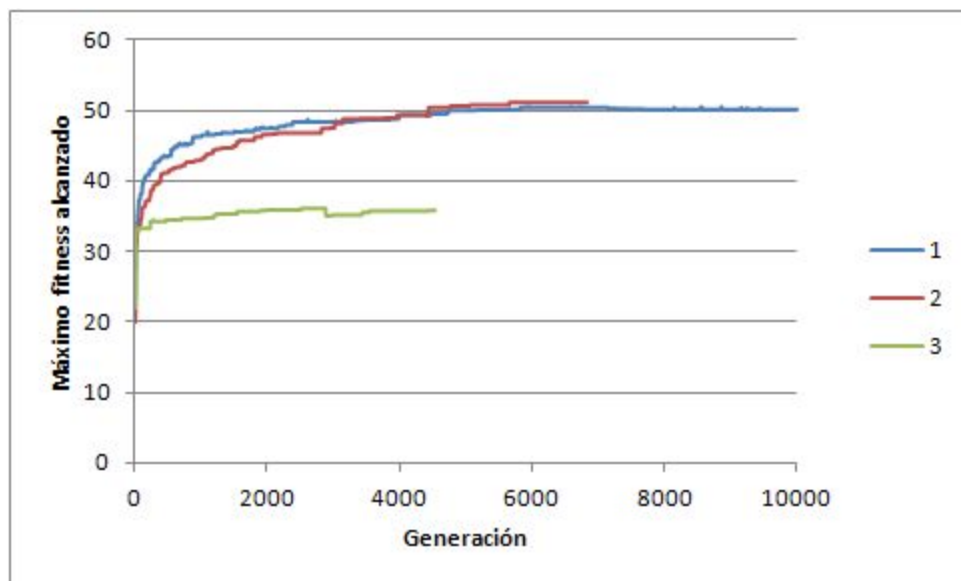


Figura 22: Gráfico del promedio del fitness sobre 5 poblaciones aleatorias en función del paso de las generaciones para $N = 100$ y $K = 50$ y las configuraciones 1, 2 y 3.

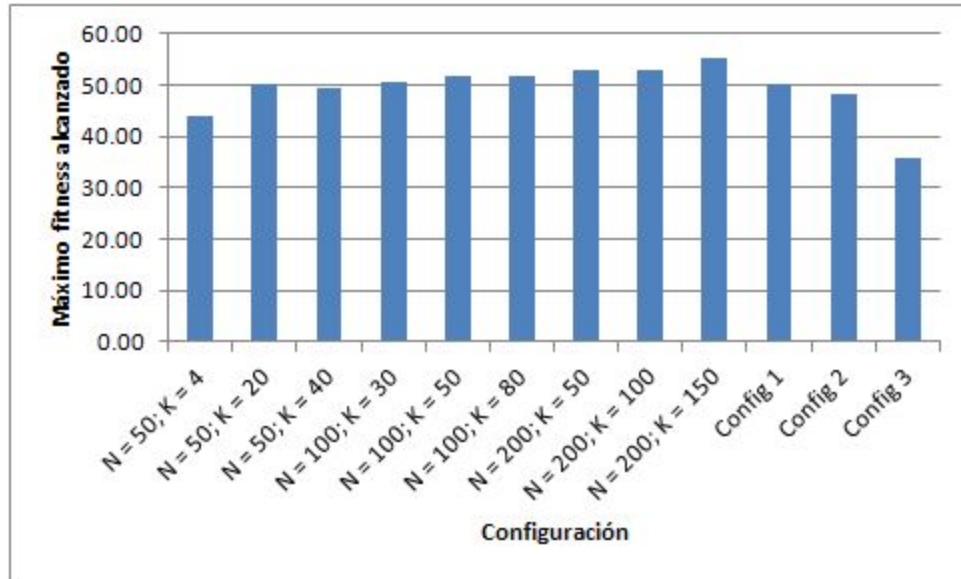


Figura 23: Gráfico del promedio del máximo fitness alcanzado sobre 5 poblaciones aleatorias para distintos valores de N y K , y las configuraciones 1, 2 y 3.

Individuo más apto

El individuo más apto que se logró obtener es el siguiente:

Altura: 1.3000056m,

Pechera nº315806

Guantes nº542040

Casco nº918849

Armas nº426547

Botas nº975383

El fitness obtenido fue 55.72.

Conclusiones

Como se puede ver en las figuras 1 y 2, los algoritmos de selección que parecen alcanzar mejores valores de fitness son elite, ranking, Boltzmann y torneos. Hay que tener en cuenta que tanto con elite como con boltzmann una vez que la temperatura bajo suficientemente que puede que haya convergencia prematura, es por esta razón, que no conviene seleccionar a todos los individuos con dichos métodos.

Por otro lado, respecto al algoritmo de selección Boltzmann, se puede ver en la figura 6 que el valor ideal de T_{min} es 3, ya que de esta forma se permite elegir individuos que no son óptimos, pero con baja probabilidad comparado con temperaturas más altas.

Respecto a la mutación, se puede ver en la figura 14 que es conveniente utilizar la variante multigen con una probabilidad de mutación cercana a 0.01.

También se pudo ver como se muestra en la figuras 18, 19, 20 y 21 que normalmente conviene que la brecha generacional sea mayor, es decir, que sea alto el porcentaje de individuos de la población a ser reemplazados por los hijos obtenidos en la cruce.

En adición, en la figura 23 se puede ver que la configuración que se decidió tomar para buscar el individuo más apto la cual se decidió a partir de los primeros resultados obtenidos, es mejor que las otras tres configuraciones que se decidieron probar.

Finalmente, en la figura 23 se puede ver que utilizar una población más grande aumenta significativamente el máximo fitness alcanzado. Esto se debe a que desde un inicio hay una mayor cantidad de alelos presentes en la población inicial, y por lo tanto es más fácil obtener aquellos que mejoran el fitness.