Algoritmo Mergesort

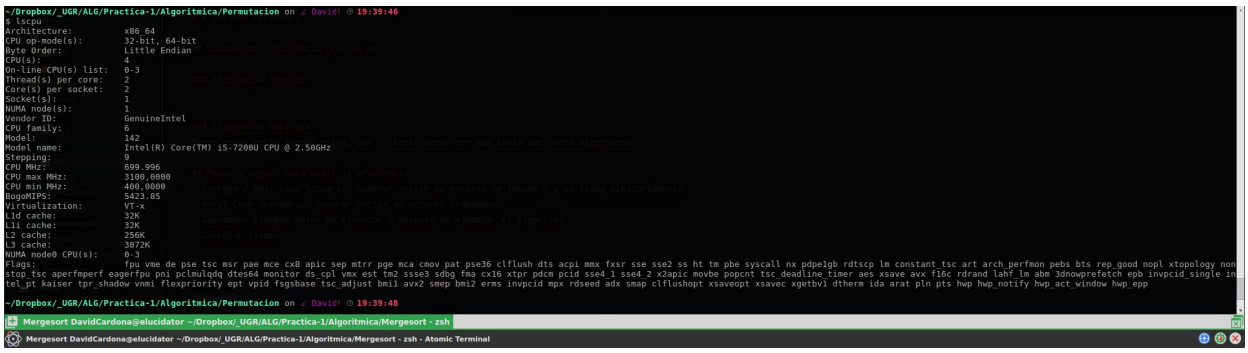
El algoritmo de ordenamiento por mezcla (merge sort en inglés) es un algoritmo de ordenamiento externo estable basado en la técnica divide y vencerás.

Pasos a seguir para medir la eficiencia:

1. Average / Best Case Scenario: Generar vector de enteros de tamano n y se llena aleatoriamente
2. Worst Case Sceneario: Generar Vector de enteros ordenados
3. Guardamos tiempos antes de ejecutar y despues de ejecutar el algortimo
4. Calcular tiempo.

Información

Hardware usado (CPU, velocidad de reloj, memoria RAM, ...)



Compilador utilizado y opciones de compilación

```
gcc -v
Using built-in specs.
COLLECT_GCC=gcc
COLLECT_LTO_WRAPPER=/usr/lib/gcc/x86_64-linux-gnu/5/lto-wrapper
Target: x86_64-linux-gnu
gcc version 5.4.0 20160609 (Ubuntu 5.4.0-6ubuntu1~16.04.9)
```

Compilación

```
g++ -std=c++11 ./src/$PROGRAMA.cpp -o ./bin/$PROGRAMA
```

Usamos el siguiente script:
# Variables:
PROGRAMA=$1
SALIDA=./data/tiempo_worst_case_$1.dat
MENSAJE_INICIO="Se inicia la ejecución del algoritmo $1:"
MENSAJE_FINAL="Fin de la ejecución. Se ha creado un fichero con los resultados."

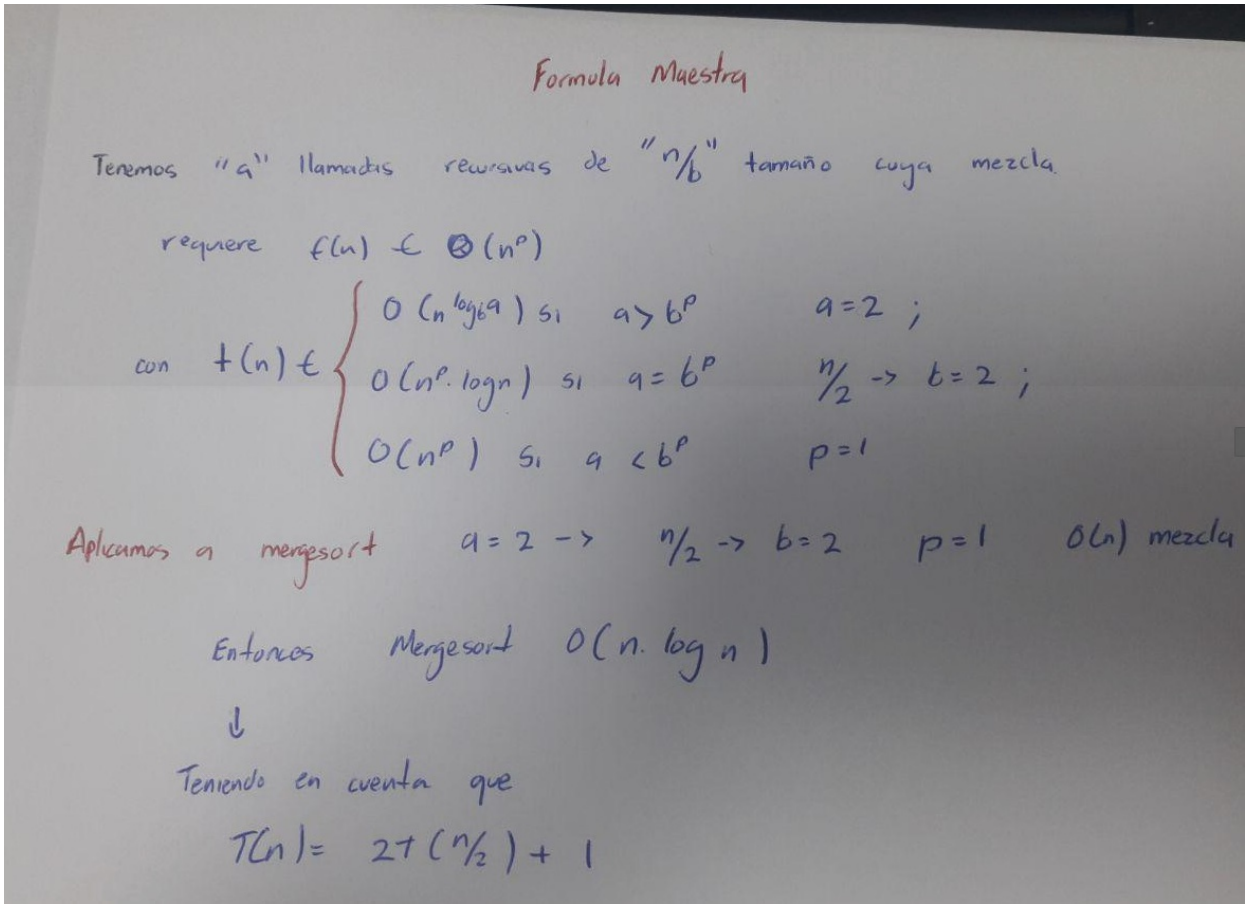
```
# Se genera el ejecutable con el algoritmo de ordenación:
g++ -std=c++11 ./src/$PROGRAMA.cpp -o ./bin/$PROGRAMA
echo "$MENSAJE_INICIO"

# Variables:
INICIO=1000
FIN=50000
INCREMENTO=1000

i=$INICIO
echo > $SALIDA
while [ $i -le $FIN ]
do
    echo Vector size = $i
    echo "`./bin/$PROGRAMA $i `" >> $SALIDA
    i=$((i+$INCREMENTO))
done

rm -fr ./bin/$PROGRAMA

echo "$MENSAJE_FINAL"
```

Sistema operativo

Desarrollo completo del cálculo de la eficiencia teórica y gráfica.

Calculo de eficiencia O(n log n)

El T(n) es igual 2T(n/2) + n (No ""+ 1") Errata.



Eficiencia Empirica

Hemos tomado multiples medidas y sobre esto hemos realizado los ajustes y graficas.

Ejemplo de medidas Worst Case

N Elementos	Tiempo
1000	0.00012729
2000	0.000245837
3000	0.000381419
4000	0.00051691
5000	0.000681268
6000	0.000793639
7000	0.000932129
8000	0.00107566
9000	0.00124717
10000	0.00137646
11000	0.001584
12000	0.001945
13000	0.001859
14000	0.002107
15000	0.002258
16000	0.00231
17000	0.00332
18000	0.003403
19000	0.00378

20000	0.004421
21000	0.004415
.....	.....

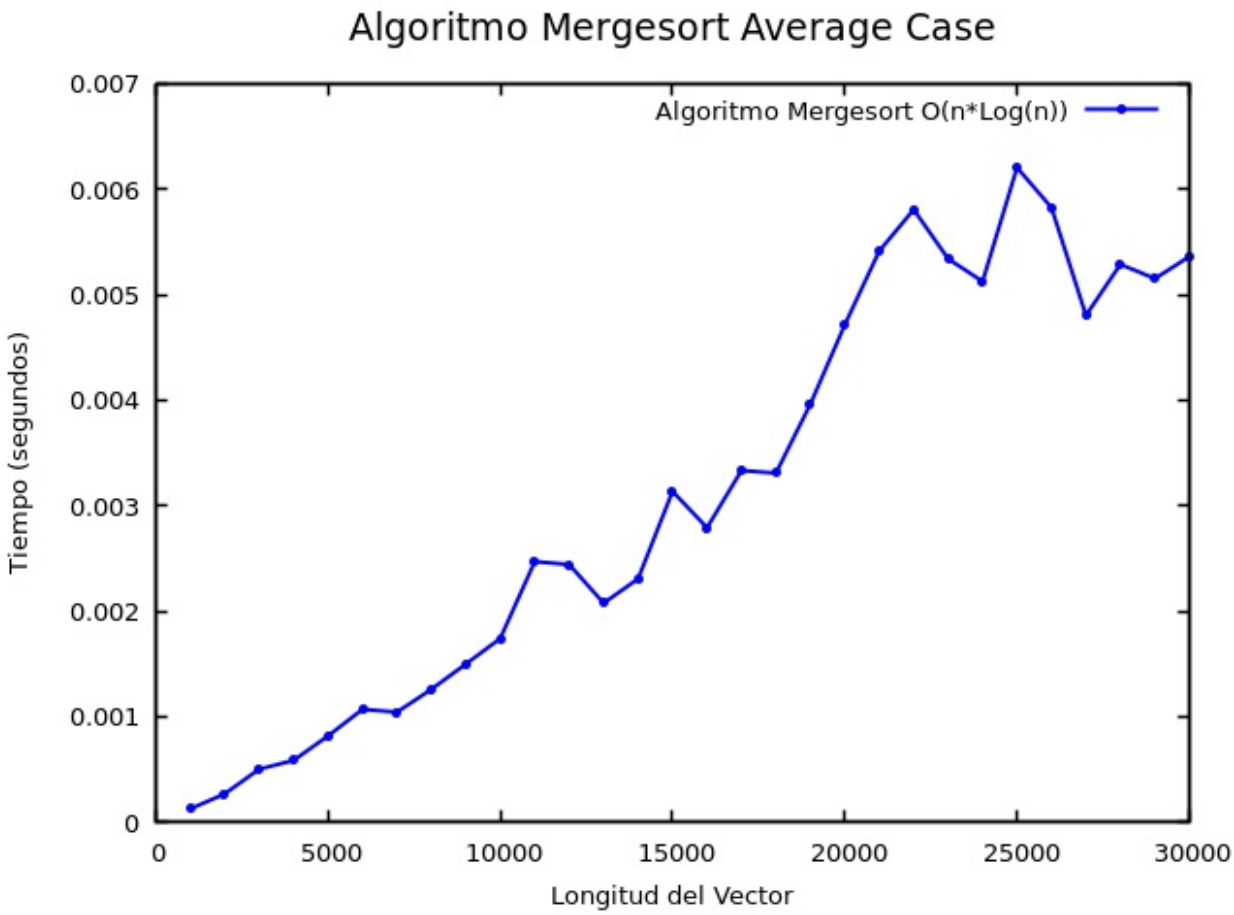
Merge ocupa mucho tiempo extra, por lo tanto suele ser mas lenta que otros algoritmos se ha notado en las graficas realizada, respecto a los otros.

Parámetros usados para el cálculo de la eficiencia empírica y gráfica.

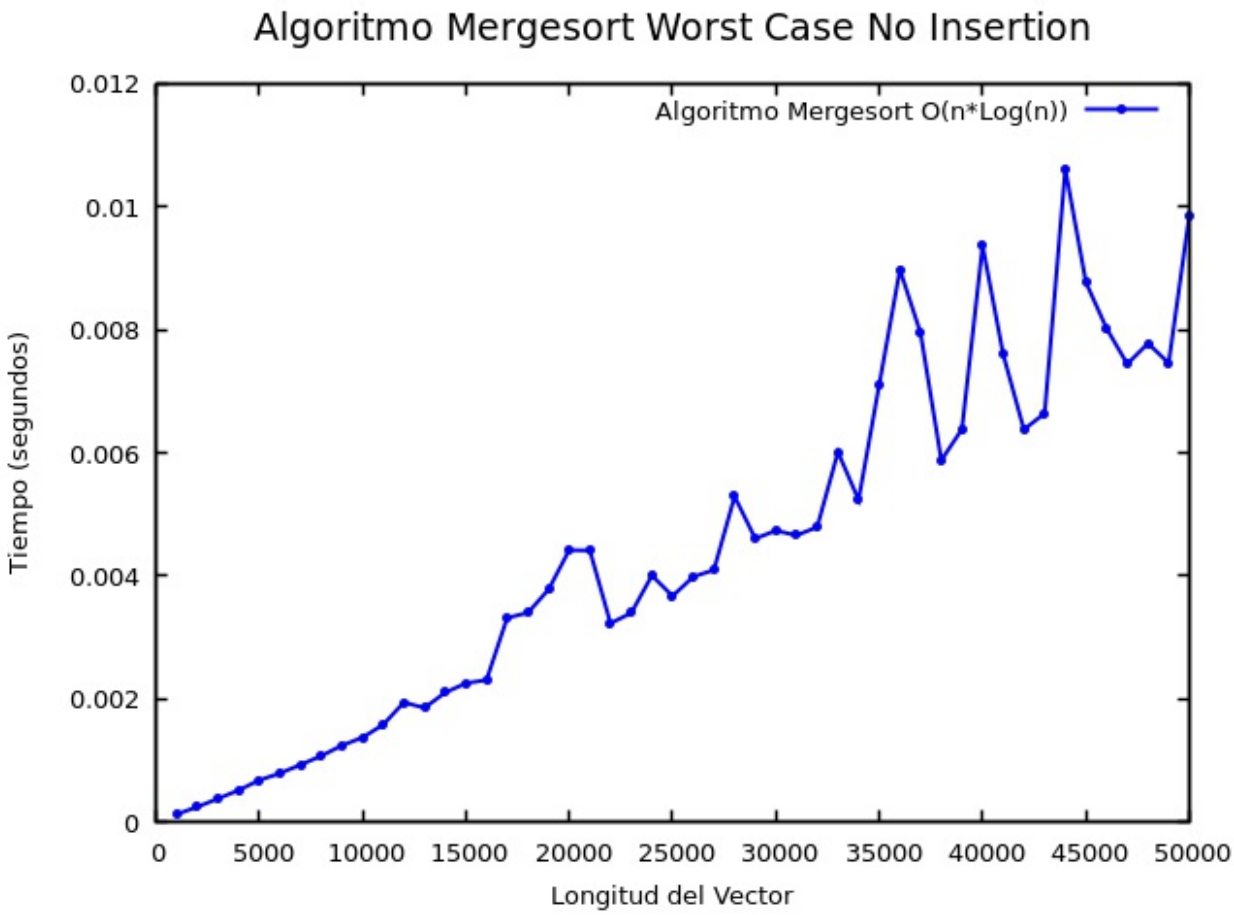
Para el calculo de las gráficas hemos usado el script:

```
#!/bin/bash

#Variables:
OUTPUT=./data/grafica_tiempo_worst_case_mergesort.png
TITULO="Algoritmo Mergesort Worst Case No Insertion"
XLABEL="Longitud del Vector"
YLABEL="Tiempo (segundos)"
LEYENDA="Algoritmo Mergesort O(n*Log(n))"
FICHERO_DATOS="./data/tiempo_worst_case_mergesort.dat"
COLOR=blue

gnuplot<<FIN
# Terminal para png:
set terminal pngcairo enhanced font 'Verdana,10'
set border linewidth 1.5

# Estilo de línea y color:
set style line 1 lc rgb '$COLOR' lt 1 lw 2 pt 7 pi 0 ps 0.5
set pointintervalbox 0

# Nombre de la imagen resultante:
set output '$OUTPUT'

# Título y ejes:
set title "$TITULO" enhanced font 'Verdana,14'
set xlabel "$XLABEL"
set ylabel "$YLABEL"

set autoscale

plot "$FICHERO_DATOS" title '$LEYENDA' with linespoints ls 1
FIN
```

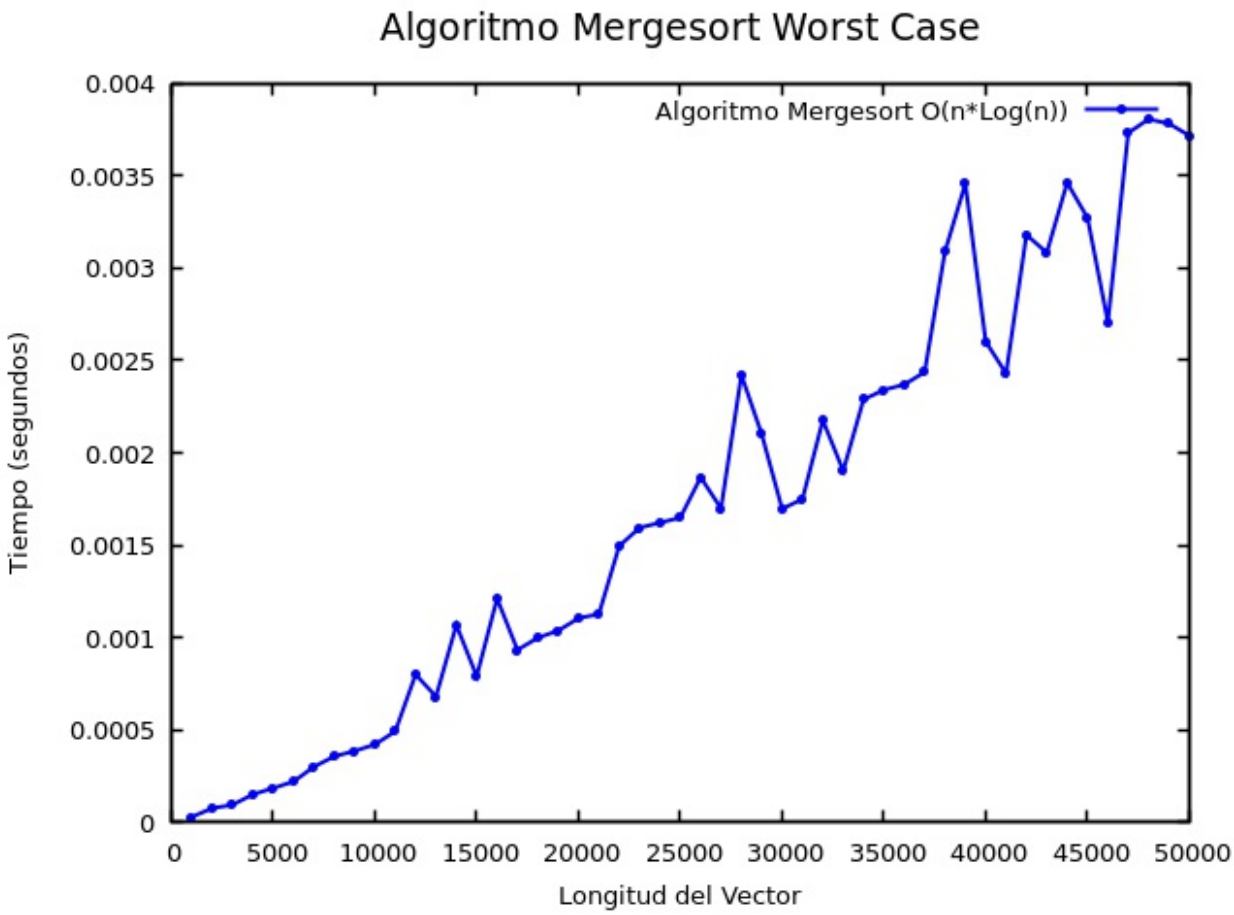
Average


Algoritmo Mergesort Average Case
Algoritmo Mergesort O(n*Log(n))
Tiempo (segundos)
Longitud del Vector
0
0.001
0.002
0.003
0.004
0.005
0.006
0.007
0
5000
10000
15000
20000
25000
30000

No Insertion

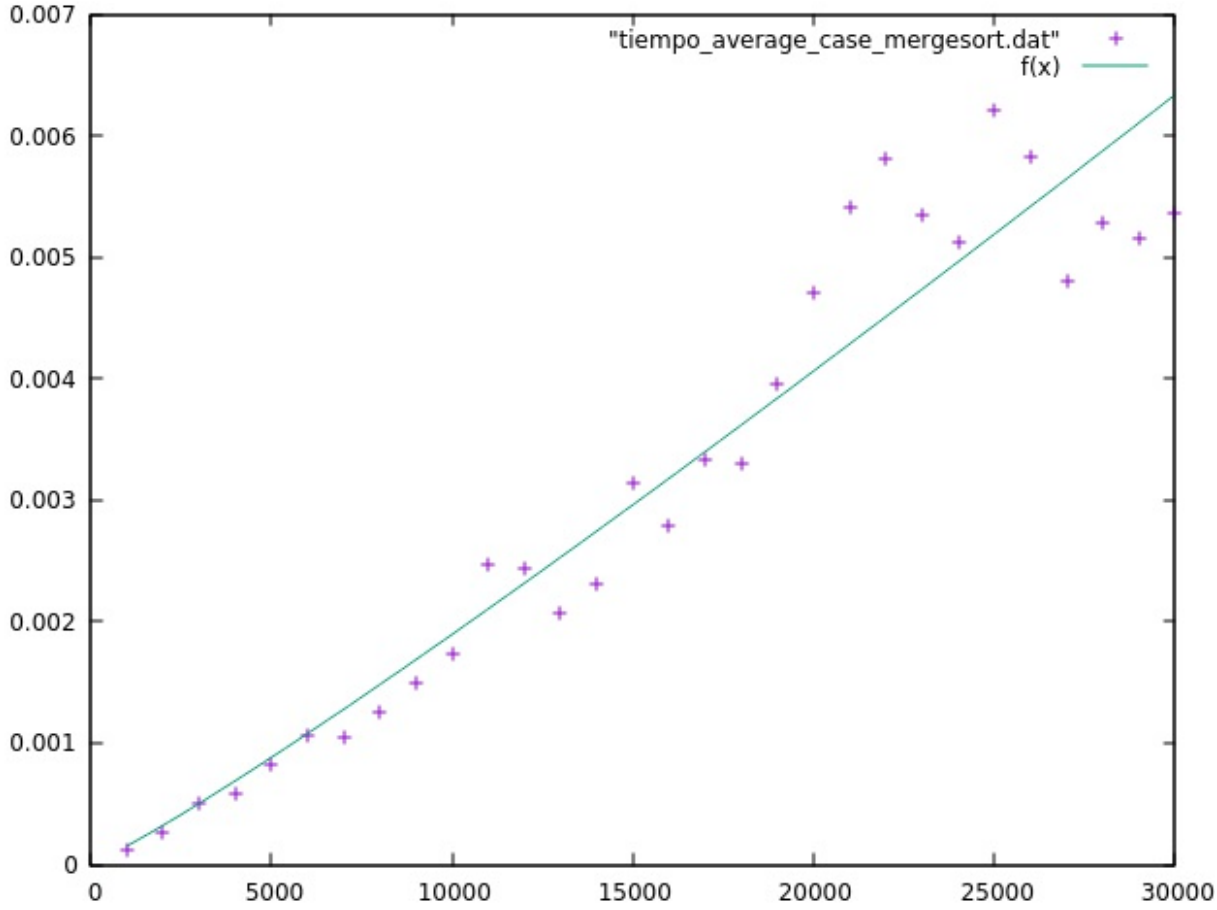


Insertion

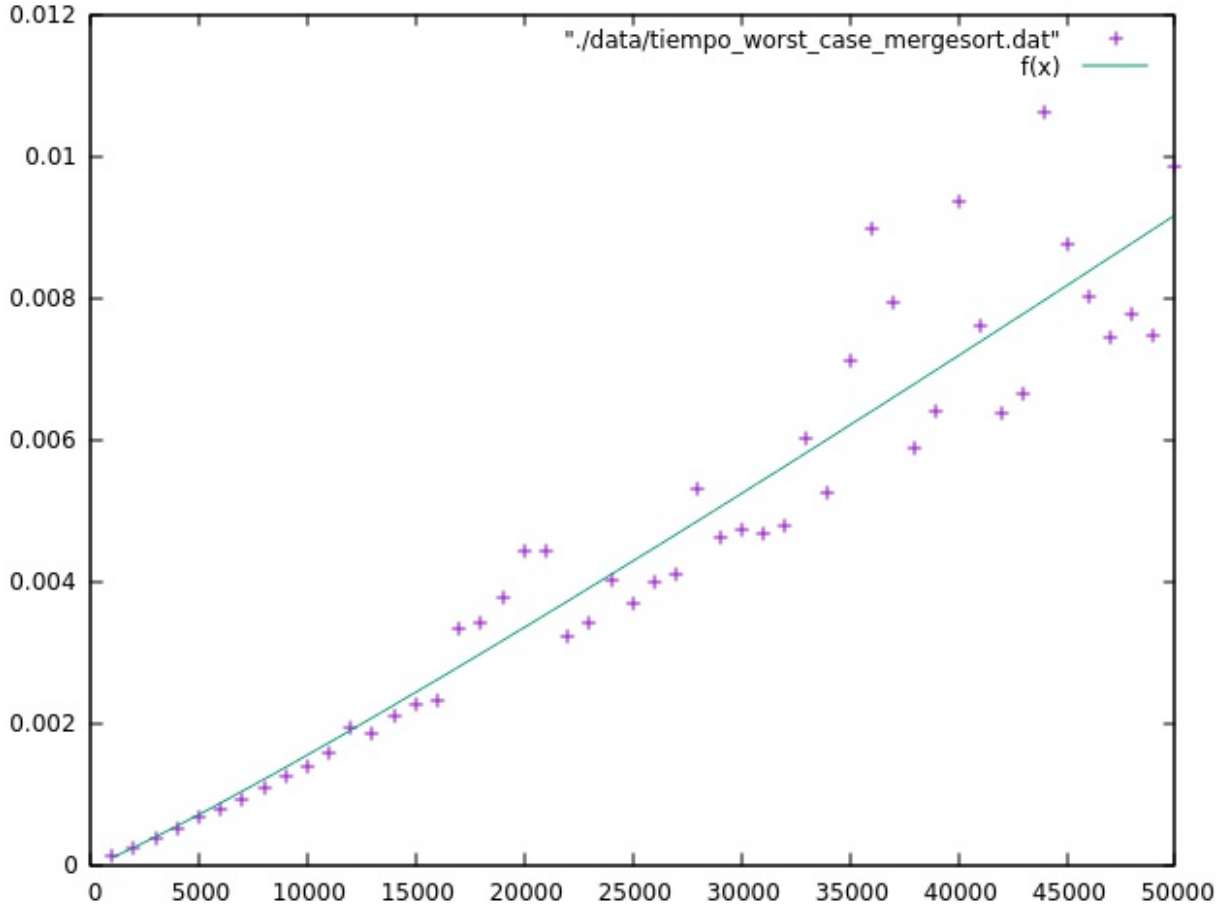

Algoritmo Mergesort Worst Case
Algoritmo Mergesort O(n*Log(n))
Tiempo (segundos)
Longitud del Vector
0
0.0005
0.001
0.0015
0.002
0.0025
0.003
0.0035
0.004
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
50000

Ajuste de la curva teórica a la empírica: mostrar resultados del ajuste y gráfica.

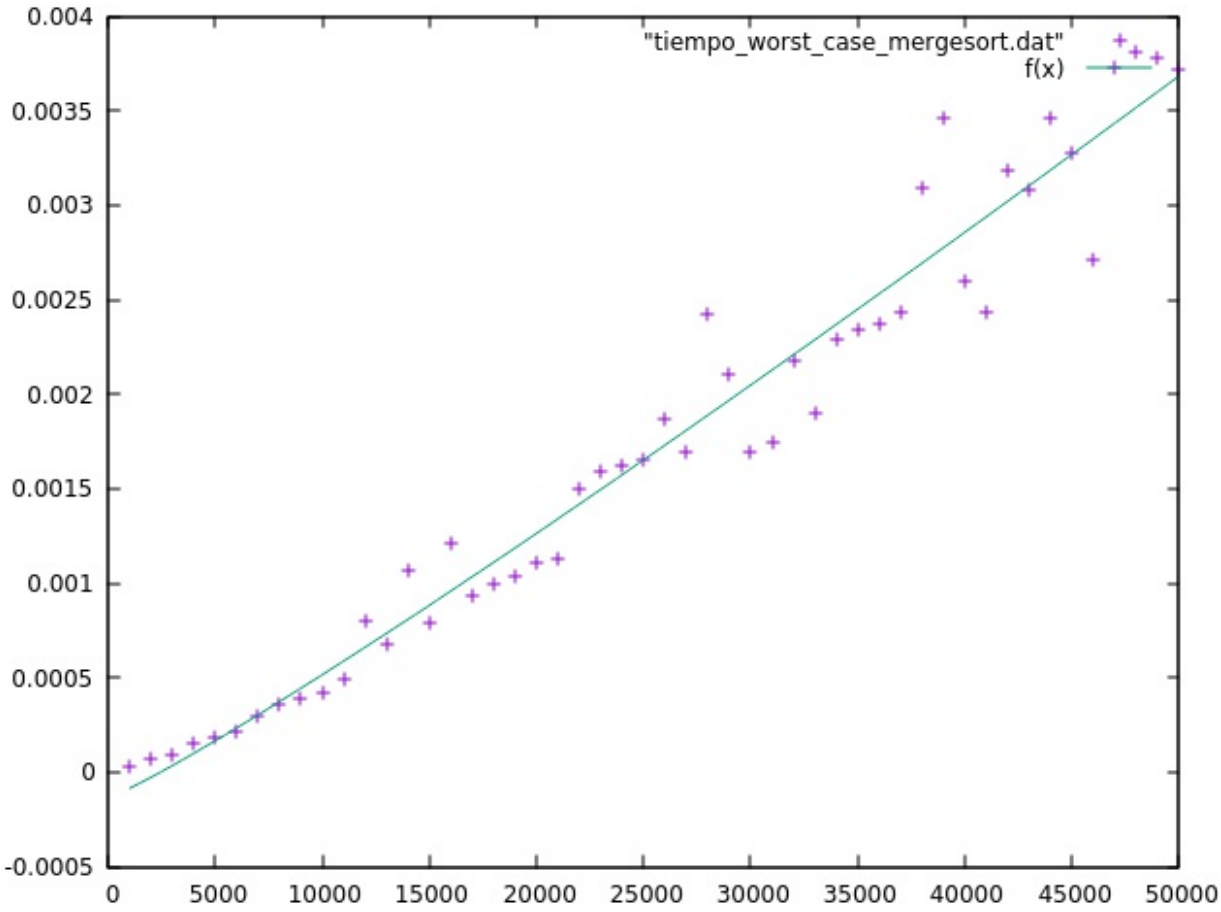
Average



Worst No Insertion



Worst Insertion



Ajuste

```
*******************************************************************************
Sat Mar 17 03:02:37 2018


FIT:    data read from "tiempo_average_case_mergesort.dat"
        format = z
        #datapoints = 30
        residuals are weighted equally (unit weight)

function used for fitting: f(x)
    f(x)=(a*x*log(x)/log(2))+b
fitted parameters initialized with current variable values

iter      chisq       delta/lim  lambda   a             b
   0 1.9669559893e+12   0.00e+00  1.81e+05    1.000000e+00   1.000000e+00
   9 8.9112223740e-06  -2.13e-04  1.81e-04    1.418250e-08   9.523383e-06

After 9 iterations the fit converged.
```

```
final sum of squares of residuals : 8.91122e-06
rel. change during last iteration : -2.12717e-09

degrees of freedom    (FIT_NDF)                       : 28
rms of residuals      (FIT_STDFIT) = sqrt(WSSR/ndf)    : 0.000564144
variance of residuals (reduced chisquare) = WSSR/ndf   : 3.18258e-07

Final set of parameters            Asymptotic Standard Error
=======================            ==========================
a               = 1.41825e-08      +/- 7.835e-10    (5.524%)
b               = 9.52338e-06      +/- 0.0002006    (2107%)

correlation matrix of the fit parameters:
                a      b
a               1.000
b              -0.858  1.000


*******************************************************************************
Sat Mar 17 03:04:05 2018


FIT:    data read from "tiempo_worst_case_mergesort.dat"
        format = z
        #datapoints = 50
        residuals are weighted equally (unit weight)

function used for fitting: f(x)
    f(x)=(a*x*log(x)/log(2))+b
fitted parameters initialized with current variable values

iter      chisq       delta/lim  lambda   a             b
   0 9.0520846224e-04   0.00e+00  4.45e-03    1.418250e-08   9.523383e-06
   5 2.8449514907e-06  -9.73e-01  4.45e-08    4.893749e-09  -1.354714e-04

After 5 iterations the fit converged.
final sum of squares of residuals : 2.84495e-06
rel. change during last iteration : -9.7309e-06

degrees of freedom    (FIT_NDF)                       : 48
rms of residuals      (FIT_STDFIT) = sqrt(WSSR/ndf)    : 0.000243454
variance of residuals (reduced chisquare) = WSSR/ndf   : 5.92698e-08

Final set of parameters            Asymptotic Standard Error
=======================            ==========================
a               = 4.89375e-09      +/- 1.501e-10    (3.066%)
b               = -0.000135471     +/- 6.662e-05    (49.17%)

correlation matrix of the fit parameters:
                a      b
a               1.000
b              -0.856  1.000
```