"Conjunto de instrucciones MIPS"

Federico Colangelo, Padrón Nro. 89.869 federico.colangelo@semperti.com

Facundo Rossi, Padrón Nro. 86.707 frossi85@gmail.com

Federico Martín Rossi, Padrón Nro. 92.086 federicomrossi@gmail.com

1
er. Cuatrimestre 2013 66.20 Organización de Computadoras Facultad de Ingeniería, Universidad de Buenos Aires



Índice

1.	Introducción	1				
2.	Compilación	1				
3.	Utilización 3.1. Implementación en C	1 1 1				
4.	Implementaciones 4.1. Implementación en C					
5.	Debugging	6				
6.	Pruebas	7				
7.	Conclusiones	8				
\mathbf{A}_{l}	ppendices	9				
Α.	A.1. $tp0.c$. Implementación del main del programa A.2. $bubblesort.h$. Declaración de la librería Bubblesort A.3. $bubblesort.c$. Definición de la librería Bubblesort A.4. $shellsort.h$. Declaración de la librería Shellsort A.5. $shellsort.c$. Definición de la librería Shellsort A.6. $fileloader.h$. Declaración de la librería File Loader A.7. $fileloader.c$. Definición de la librería File Loader	14 15 15 16				
В.	Implementación completa de Shellsort en assembly MIPS B.1. tp0.c. Implementación del main del programa	20 21				

1. Introducción

En este trabajo presentamos la comparación entre dos algoritmos de ordenamiento: el Bubblesort y el Shell-sort.

Para realizar la comparación entre ambos se realizó la implementación en lenguaje C de cada uno. A su vez, para comparar la performance entre código de alto nivel y código nativo, se hizo la implementación de shellsort en assembler MIPS, con el fin de poder comparar los tiempos de ejecución de ambos programas.

Todo el trabajo se realizo en una plataforma NetBSD/MIPS-32 mediante el GXEmul [1]

Todos los archivos y códigos fuente aquí mencionados, así como también el presente informe, pueden ser descargados de la sección Downloads del repositorio del grupo¹.

2. Compilación

La herramienta para compilar tanto el código asembly como C será el GCC [2]. Para tratar de equiparar al máximo el S generado por ambas implementaciones, se utilizará el flag de gcc "-O0" para que no realice optimizaciones sobre el código en lenguaje C.

Para automatizar las tareas de compilación se hace uso de la herramienta $GNU\ Make$. Los Makefiles utilizados para la compilación se incluyen junto al resto de los archivos fuentes del presente trabajo 2 .

3. Utilización

En los siguientes apartados se especifica la forma en la que deben ser ejecutados los programas implementados tanto en C como en assembly.

3.1. Implementación en C

El resultado de compilación con "make" será un programa ejecutable, de nombre $tp\theta$, que podrá ser invocado con los siguientes parámetros:

- -h: Imprime ayuda para la utilización del programa;
- -V: Imprimer la versión actual del programa;
- -b [ARGS]: El programa recibe nombres de archivos de texto o strings ingresados por stdin, ordenandolos utilizando el algoritmo bubblersort. Para utilizar stdin deberá omitirse [ARGS] y luego introducir las palabras;
- -s [ARGS]: El programa recibe nombres de archivos de texto o strings ingresados por stdin, ordenandolos utilizando el algoritmo shellsort. Para utilizar stdin deberá omitirse [ARGS] y luego introducir las palabras.

3.2. Implementación en Assembly

El resultado de compilación con "make" será un programa ejecutable, de nombre $tp\theta$, el cual aceptará un archivo de texto como argumento y lo ordenará con el algoritmo Shellsort.

4. Implementaciones

En lo que sigue de la sección, se presentarán los códigos fuente de las implementaciones de los algoritmos. Aquellos lectores interesados en la implementación completa de los dos programas, pueden dirigirse a los apéndices ubicados al final del presente informe. Recordamos que se han separado las implementaciones en dos de manera de poder mantener un orden entre ambos.

¹URI del Repositorio: https://code.google.com/p/tp-orga-computadoras/

²Los archivos se encuentran separados según la implemetación a la que pertenecen, por lo que habrán dos Makefiles distintos, uno para la implementación en lenguaje C y otro para la implementación en assembly

4.1. Implementación en C

La implementación del programa fue divida en los siguientes módulos:

- tp0: Programa principal responsable de interpretar los comandos pasados por la terminal de modo que realice las tareas solicitadas por el usuario. Su principal función es encadenar el funcionamiento de los otros módulos y mostrar por pantalla el resultado obtenido;
- fileloader: Módulo encargado de levantar un archivo de texto desde el filesystem y convertirlo en un array donde cada posición es una palabra. Permite además dimensionar el tamaño en memoria necesario para cargar todo el archivo como un arreglo de palabras;
- bubblesort: Módulo encargado de levantar un archivo de texto desde el filesystem y convertirlo en un array donde cada posición es una palabra. Permite además dimensionar el tamaño en memoria necesario para cargar todo el archivo como un arreglo de palabras;
- shellsort: Módulo encargado de implementar el algoritmo de shellsort. Recibe como paramentros un arreglo de palabras desordenado y el tamaño del mismo. Como resultado devuelve dicho arreglo ordenado.

4.1.1. Algoritmo Bubblesort

En el C'odigo~1 se muestra el header de la librería, donde se declara la función Bubblesort, mientras que en el C'odigo~2 se muestra la definición de la librería.

```
Código 1: "bubblesort.h"
   * LIBRERIA BUBBLESORT
11 #ifndef BUBBLESORT_H
12 #define BUBBLESORT_H
13
14
15
16 // Funcion que aplica el algoritmo de ordenamiento Bubblesort para ordenar un
17 // arreglo de palabras.
18 // PRE: 'words' es un puntero a un arreglo de punteros a caracter; 'arraysize'
19 // es el tamanio de dicho arreglo.
20 // POST: el arreglo 'words' queda ordenado.
void bubblesort(char* words[], int arraysize);
23
24
  #endif
```

```
16 // Funcion que aplica el algoritmo de ordenamiento Bubblesort para ordenar un
17 // arreglo de palabras.
18 // PRE: 'words' es un puntero a un arreglo de punteros a caracter; 'arraysize'
19 // es el tamanio de dicho arreglo.
20 // POST: el arreglo 'words' queda ordenado.
void bubblesort(char* words[], int arraysize)
22 {
    // Variables de procesamiento
    bool huboIntercambio;
24
25
    int i;
    int n = arraysize;
26
27
    char* sAux;
    // Recorremos el arreglo haciendo intercambios hasta que ya no se registre
29
    // ningun cambio realizado.
30
    do
31
    {
32
      huboIntercambio = false;
33
34
      for(i = 1; i < n; i++)
35
36
        // Si el de indice menor es mayor que el de indice superior, los
37
         // intercambiamos
38
39
        if(strcasecmp(words[i-1], words[i]) > 0)
40
41
           sAux = words[i-1];
           words[i-1] = words[i];
42
           words[i] = sAux;
43
44
           // Cambiamos el flag para registrar que hubo un cambio
45
           huboIntercambio = true;
46
        }
47
48
49
      // Como el elemento del indice superior se encuentra ya ordenado una
50
      // vez finalizada la pasada, se reduce en uno la cantidad de indices
51
52
      // a iterar en la proxima pasada.
      n -= 1;
53
54
55
    } while(huboIntercambio);
56 }
```

4.1.2. Algoritmo Shellsort

En el Código 3 se muestra el header de la librería, donde se declara la función Shellsort, mientras que en el Código 4 se muestra la definición de la librería.

```
22
23
24
25 #endif
```

```
Código 4: "shellsort.c"
* LIBRERIA SHELLSORT
  #include "shellsort.h"
11 #include <string.h>
12
15 // Funcion que aplica el algoritmo de ordenamiento Shellsort para ordenar un
16 // arreglo de palabras.
17 // PRE: 'words' es un puntero a un arreglo de punteros a caracter; 'arraysize'
18 // es el tamanio de dicho arreglo.
19 // POST: el arreglo 'words' queda ordenado.
void shellsort(char* words[], int arraysize)
21 {
22
    int intervalo, k, j, i;
23
    char* sAux;
24
25
    // Tomamos como intervalo el modulo de la mitad del arreglo
    intervalo = arraysize / 2;
26
    // Iteramos hasta que el intervalo sea nulo
28
    while(intervalo > 0)
29
30
     for(i = intervalo-1; i < arraysize; i++)</pre>
31
32
       j = i - intervalo;
33
34
       // Procesamos hasta caer en un indice fuera de rango del arreglo
35
       while(j >= 0)
36
37
       {
38
         k = j + intervalo;
39
40
         // Si el de indice menor es mayor que el de indice superior, los
41
         // intercambiamos
         if(strcasecmp(words[j], words[k]) > 0)
42
43
         {
44
           sAux = words[j];
           words[j] = words[k];
45
           words[k] = sAux;
47
         }
48
         else
           j = 0;
50
         j -= intervalo;
51
52
53
54
      // Tomamos como nuevo intervalo el modulo de la mitad del intervalo
55
      // anterior
56
57
     intervalo = intervalo / 2;
   }
58
59 }
```

4.2. Implementación en Assembly

La implementación del programa fue divida en los siguientes módulos:

- **tp0**: Solamente recibe un texto como argumento por linea de comandos y lo imprime ordenandolo mediante shellsort. Esta implementado en C;
- fileloader: Se utiliza el mismo de la otra implementación;
- swap: Función implementada en assembler para intercambiar el valor de dos registros recibidos como parametros;
- compare: Función similar a stremp de C. Recibe dos argumentos de tipo texto y decide cual es el que precede alfabeticamente.
- shellsort: Implementación en assembler del algoritmo de ordenamiento.

4.2.1. Algoritmo Shellsort

En el Código 5 se muestra la implementación en assembly del algoritmo Shellsort.

Código 5: "shellsort.S" #ifndef USE_MIPS_ASSEMBLY 2 #define USE_MIPS_ASSEMBLY 3 #include <mips/regdef.h> 4 #include <sys/syscall.h> 6 .text 7 .align 2 8 .globl shellsort 9 .extern swap 10 .extern compare 11 13 shellsort: subu sp,sp,24 # creo stack frame, es de 24 porque necesito guardar ra sw ra,16(sp) 14 sw \$fp,12(sp) 15 16 **sw** gp,8(sp) move \$fp,sp 17 18 #Guardo los parametros recibidos en fp 19 sw a0,24(\$fp) # words = char ** 20 sw a1,28(\$fp) # size 22 #Guardo los parametros cargados en fp a variables temporales 23 24 lw s6, 24(\$fp) #en 24 esta el words lw s7, 28(\$fp) #en 28 esta el size 25 26 #Iniciaizo variables locales 27 srl s0, s7, 1 # intervalo = size/2 28 # k **li** s1, 0 **li** s2, 0 # j 30 31 **li** s3, 0 # i **li** s4, 0 # sAux 32 33 34 #add v0, s0, zero 35 #j fin 36 37 #while(intervalo > 0) 38 39 loop1: bgt s0, zero, initloop2 #branch grater than, salta si s0 es mayor a 0 41 #for(i = intervalo-1; i < arraysize; i++)</pre> 43 initloop2: addi s3, s0, -1 # i = interlvalo - 1 loop2: bge s3, s7, endloop2 # si i >= size fin de for, bge=branch greater than or equal 44 **sub** s2, s3, s0 #j = i - intervalo; 46 47 #while(j >= 0) 49 loop3: bge s2, zero, loop3task addi s3, s3, 1 j loop2

```
53
54
55
56 loop3task: add s1, s2, s0
                                 \#k = j + intervalo;
       #Para comparar cadenas, la forma mas sencilla es iterar sobre cada caracter de cada cadena, y restarlos. \hookleftarrow
58
            Si el resultado es 0, son
_{59} #iguales. Si no, a continuacion, si el resultado es> 0, entonces la primera cadena es mayor de que la otra \leftrightarrow
        cadena, de lo contrario la segunda cadena es inferior y #habria que intercambiarlas. Si te quedas sin \leftrightarrow
        alguna de las cadenas antes que el otro, y son iguales hasta llegar a ese punto, la cadena mas corta es \leftrightarrow
        menor que el que #mas.
       #CARGO LAS CADENAS A COMPARAR
60
       #words[j]
61
       ##Guardo en variables auxiliares las direcciones de memoria a intercambiar
62
       #words = s6
63
       #Guardo i y j por q las voy a modificar y no las quiero perder
       sw s1, 0($fp)
65
       sw s2, 4($fp)
66
67
68
69
70
       sll s1, s1, 2
       sll s2, s2, 2
71
       addu s1, s1, s6
       addu s2, s2, s6
74
       #Cargo los strings a comparar en ax que luego compare utiliza para trabajar
75
       lw a0. 0(s1)
                        #OPTIMIZAR CODIGO PARA ISAR MENOS ti
76
77
       lw a1, 0(s2)
78
       jal compare
                      #llamo a compare
79
       lw s1, 0($fp)
                        #Cargo el valor de k en stack
       lw s2, 4($fp)
                        #Cargo el valor de j en stack
81
82
83
       bltz v0, swapping #swap si compare > 0
       add s2, zero, zero \#j=0
84
85
       i endif
86
87
       # poner el valor a utilizar en los registros correctos y llamar a swap
   swapping: add a0, s6,0
       add a1, s1, 0
89
90
       add a2, s2, 0
91
       jal swap
92
   endif:
93
              sub s2, s2, s0
                               #j = j - intervalo;
       j loop3
94
95
   endloop2: srl s0, s0, 1 #intervalo = intervalo / 2;
97
98
       j loop1
99
100
101
       ##Comienzo retorno
102
         move sp, $fp
103 fin:
       lw gp, 8(sp)
       lw $fp, 12(sp)
105
106
       lw ra, 16(sp)
       addiu sp, sp, 24
107
                  #ra = return address es la direccion a la q quiero volver luego de llamar la funcion
       ir ra
108
109
110 #endif
```

5. Debugging

Para analizar el correcto funcionamiento de los programas se crearon programas adhoc que fueran corroborando el correcto funcionamiento de cada uno de los modulos de manera individual. La idea era simular el concepto de test unitario de lenguajes de alto nivel.

Una vez que todos los módulos funcionaban de la manera esperada, se utilizó el programa principal como test de integración. Finalmente se utilizó la herramienta *Valgrind* para realizar un análisis minucioso del uso de memoria, corrigiendo así las pérdidas de memoria que se presentaban en cada módulo.

6. Pruebas

Para realizar las pruebas de cada algoritmo se procesaron cuatro textos provistos por la cátedra. Estos son "Alicia en el País de las Maravillas" (173kB), "Beowulf" (220kB), una "Enciclopedia" (643kB) y "Don Quijote" (2147kB).

A cada ejecución se le midió el tiempo de procesamiento mediante el comando *GNU* "time" [5]. A continuación un ejemplo de su utilización:

\$ time ./tp0 -b alice.txt

Esto imprimirá por pantalla el tiempo de ejecución que fue tabulado para cada caso. En el $Cuadro\ 1$ se muestran los valores de tiempo obtenidos.

Archivo de texto	bubblesort.c [s]	shellsort.c [s]	shellsort.S [s]
alice.txt	571,715	2,207	1,035
beowulf.txt	958,527	3,418	1,500
cyclopedia.txt	>3600	12,184	6,914
elquijote.txt	>3600	48,957	33,098

Cuadro 1: Tiempos en segundos obtenidos en la ejecución de los distintos algoritmos.

En los graficos que siguen se muestran distintas comparaciones hechas con estos datos obtenidos, los cuales nos llevarán a la conclusión final.

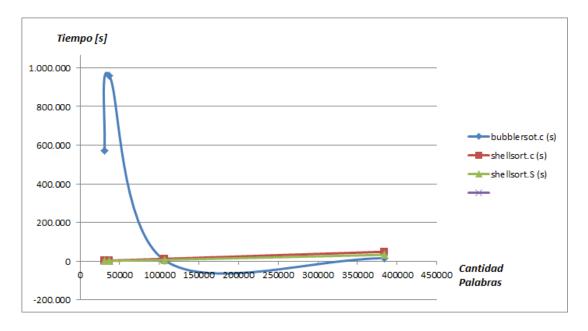


Figura 1: Gráfico de tiempo insumido (en segundos) contra el tamaño de la muestra (cantidad de palabras)

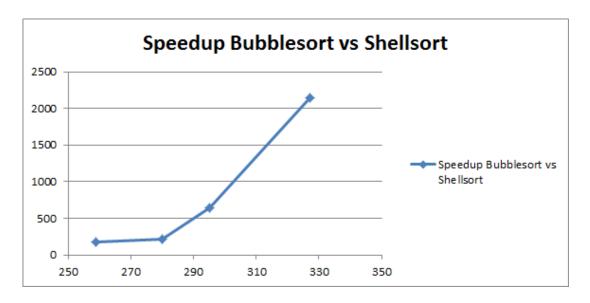


Figura 2: Gráfico del speedup del Bubblesort contra Shellsort para los diversos tamaños de archivo.

7. Conclusiones

Al realizar la comparación entre los algoritmos bubblesort y shellsort en C, vemos que la diferencia entre ambos es notable. Estamos hablando de varios ordenes de maginitud.

Esto se debe a que por su codigo bubblesort tiene un coste de $O(n^2)$ mientras que si bien no podemos determinarlo fácilmente para shellsort, ya que depende mucho del gap elegido, estamos hablando de algo cercano a O(nlog(n)). Esto hace que para textos grandes no sea una opción viable la utilización de bubblesort, mientras que shellsort prueba ser muy eficiente y el esfuerzo de codificación no es demsiado grande.

En cuanto a la comparación de assembler y C, notamos una ligera ventaja del primero. Creemos que esto se debe a que muchas de las cosas que el algoritmo en C utiliza como variables en memoria principal, assembler utiliza registros los cuales presentan un tiempo de lecto/escritura notablemente menor.

Referencias

- [1] The NetBSD project, http://www.netbsd.org/
- [2] GCC, the GNU Compiler Collection, http://gcc.gnu.org/
- [3] Bubblesort, http://en.wikipedia.org/wiki/Bubble_sort
- [4] Shellsort, http://en.wikipedia.org/wiki/Shell_sort
- [5] time man page, http://unixhelp.ed.ac.uk/CGI/man-cgi?time
- [6] MIPS ABI, http://www.sco.com/developers/devspecs/mipsabi.pdf
- [7] J. L. Hennessy and D. A. Patterson, "Computer Architecture. A Quantitative Approach," 4th Edition, Morgan Kaufmann Publishers, 2000.

Apéndices

A. Implementación completa en lenguaje C

A.1. *tp0.c.* Implementación del main del programa

```
Código 6: "tp\theta.c"
  * PROGRAMA DE ORDENAMIENTO DE PALABRAS
  * Facultad de Ingenieria - UBA
  * 66.20 Organizacion de computadoras
   * Trabajo Practico: Conjunto de instrucciones MIPS
   * ALUMNOS:
10
  * Facundo Rossi (86707) - frossi85@gmail.com
   * Federico Colangelo (89869) - federico.colangelo@semperti.com
  * Federico Martin Rossi (92086) - federicomrossi@gmail.com
15
  * Programa que permite aplicar metodos de ordenamiento a un texto o textos
  * especificados por el usuario. Estos textos pueden ser ingresados a traves de
  * la entrada estandar o bien a traves de la especificacion de los nombres de
* archivo que se desean procesar. Cabe resaltar que al especificar varios
* nombres de archivos, los metodos seran aplicados al conjunto total de los
  * textos contenidos en estos. Es decir, el resultado del ordenamiento sera
  * unico ya que contendra la mezcla de palabras de todos los archivos.
  * Los metodos de ordenamiento soportados por la presente version del programa
25
  * son:
26
       - Bubblesort
27
28
        - Shellsort
29
31
   * FORMA DE USO
34
   * El programa debera ser ejecutado por consola mediante el siguiente comando:
        # ./tp0 [OPCION] [archivos...]
37
38
   * donde, la opcion puede ser:
39
           Muestra la ayuda
           Muestra la informacion de la version
42
  * -b,
           Indica la utilizacion del algoritmo Bubblesort
            Indica la utilizacion del algoritmo Shellsort
  * y en archivos se debe especificar los nombres de archivo (incluyendo su
   * extension) separados uno del otro por un espacio.
  * En caso de desear ingresar el texto a traves de la entrada estandar, debe
  * utilizarse la siguiente variacion del comando anterior:
50
        # ./tp0 [OPCION]
51
53
54
55
58 #include <ctype.h>
59 #include <stdio.h>
60 #include <stdlib.h>
61 #include <string.h>
62 #include <unistd.h>
```

```
63 #include "bubblesort.h"
64 #include "shellsort.h"
65 #include "fileloader.h"
67
69
* CONSTANTES
74
75
76 // Maximo de caracteres permitidos por texto de entrada estandar
77 #define MAX_CHARS 1000
80
* FUNCIONES AUXILIARES
83
86 // Funcion que imprime por pantalla el mensaje de ayuda del programa.
  void ayudaImprimirSalidaEstandar()
88 {
     printf("tp0 [OPTIONS] [files...]\n");
    printf("-h, \t\t display this help and exit.\n");
90
    printf("-V, \t\t display version information and exit.\n");
91
    printf("-b, \t \t use the bubblesort algorithm.\n");
     printf("-s, \t\t use the shellsort algorithm.\n\n");
93
    printf("$echo -n echo 'El tractorcito rojo que silbo y bufo' > entrada.txt\n");
94
    printf("$tp0 -b entrada.txt\n");
     printf("bufo El que rojo silbo tractorcito y\n\n");
96
97
     printf("cat letters.txt\n");
    printf("aAbBcCdDeEfFgGhHiIjJkKlLmMnNoOpPqQrRsStTuUvVwWxXyYzZ$\n\n");
98
     printf("$tp0 letters.txt\n");
99
     printf("aAbBcCdDeEfFgGhHiIjJkKlLmMnNoOpPqQrRsStTuUvVwWxXyYzZ\n$\n\n");
    printf("$tp0 letters.txt entrada.txt\n");
101
     printf("aAbBcCdDeEfFgGhHiIjJkKlLmMnNoOpPqQrRsStTuUvVwWxXyYzZ bufo El que rojo silbo tractorcito y\n$\n");
102
103 }
104
105
106 // Funcion que imprime por pantalla la version del programa.
void versionImprimirSalidaEstandar()
    printf("Programa de ordenamiento de palabras\nVersion: v0.1\n");
109
110 }
113 // Funcion que imprime los resultados del ordenamiento a traves de la salida
115 // PRE: 'palabras' es un arreglo de palabras; 'arraysize' es el tamanio del
116 // arreglo
void ordenamientoImprimirSalidaEstandar(char* palabras[], int arraysize)
118 {
119
    for(i = 0; i < arraysize; printf("%s ", palabras[i++]));</pre>
120
    printf("\n");
122 }
125 // Funcion que se encarga leer archivos de texto devolviendo una cadena con
126 // el contenido de estos.
127 // PRE: 'iniArchivos' es el indice inicial desde donde 'argv' posee nombres
128 // de archivo; 'argc' y 'argv' son los parametros de entrada de la funcion
130 // POST: se devuelve una unica cadena con el contenido completo de archivos.
char* cargarTextosDeArchivos(int iniArchivos, int argc, char **argv)
132
133
     char * texto = "";
134
     // Sensamos cada archivo ingresado y concatenamos su contenido
136
    for(i = iniArchivos; i < argc; i++)</pre>
137
```

```
// Cargamos contenido de archivo
       char * texto_tmp = file_loader(argv[i]);
140
141
       // Recalculamos el tamanio del arreglo
142
       int len = strlen(texto) + strlen(texto_tmp) + 2;
143
144
       // Solicitamos espacio de memoria nuevo
145
       char * texto_alloc = (char *) malloc(len * sizeof(char));
146
       // Copiamos contenido antigui y concatenamos el nuevo
148
149
       strcpy(texto_alloc, texto);
       if(i > iniArchivos) strcat(texto_alloc, " ");
150
       strcat(texto_alloc, texto_tmp);
151
152
       free(texto_tmp);
153
       // Liberamos espacio de memoria antiguio
154
       if(i > iniArchivos) free(texto);
156
       // Asociamos al nuevo arreglo
157
       texto = texto_alloc;
158
     }
159
160
161
     return texto;
162 }
164
165 // Funcion que se encarga de leer el texto ingresado por entrada estandar.
166 // POST: se devuelve una unica cadena con el texto ingresado.
char* cargarTextosDeEntradaEstandar()
168
169
     char * texto = (char *) malloc(MAX_CHARS * sizeof(char) + 1);
     if(fgets(texto, MAX_CHARS + 1, stdin) == NULL) return "";
170
     return texto;
173 }
174
176
   void testFileLoader(const char * url)
177 {
178
     char * text;
     char ** words = NULL;
179
     int wordsSize = 0;
180
181
182
     text = file_loader(url);
183
184
     if (text != NULL)
185
     {
       //fputs(text, stdout);
186
188
     wordsSize = to_words(text, &words);
189
190
     for(int i = 0; i < wordsSize; i++)</pre>
191
192
       fputs("\n", stdout);
193
       fputs(words[i], stdout);
194
195
       fputs("\n", stdout);
196
197 }
198
199
200
201
202 /* ************
   * PROGRAMA PRINCIPAL
    204
205
206
int main(int argc, char **argv)
208
209
210
211
     //File loader variables
     char * text;
212
     char ** words = NULL;
213
     int wordsSize = 0;
```

```
opterr = 0;
216
217
     while ((c = getopt(argc, argv, "hVb:s:")) != -1)
218
219
220
        switch (c)
        {
221
          // Ayuda
222
          case 'h':
223
            ayudaImprimirSalidaEstandar();
224
225
            break;
226
          // Version
227
          case 'V':
            versionImprimirSalidaEstandar();
229
230
            break:
          // Ejecucion de bubblesort
232
          case 'b':
233
234
            // Cargamos contenidos de archivos
235
236
            text = cargarTextosDeArchivos(2, argc, argv);
237
            // Si el/los archivos esta/n vacio/s, salimos
238
            if (text == NULL) break;
240
241
            wordsSize = to_words(text, &words);
242
            if (wordsSize == -1) break;
243
244
            // Ejecutamos bubblesort
245
            bubblesort(words, wordsSize);
246
            // Enviamos a salida estandar
248
249
            ordenamientoImprimirSalidaEstandar(words, wordsSize);
250
            free(words);
251
252
            free(text);
            break;
253
254
          // Ejecucion de shellsort
          case 's':
256
257
            // Cargamos contenidos de archivos
258
            text = cargarTextosDeArchivos(2, argc, argv);
259
260
            // Si el/los archivos esta/n vacio/s, salimos
261
            if (text == NULL) break;
262
            wordsSize = to_words(text, &words);
264
265
            if (wordsSize == -1) break;
266
267
            // Ejecutamos shellsort
268
            shellsort(words, wordsSize);
269
270
            // Enviamos a salida estandar
            ordenamientoImprimirSalidaEstandar(words, wordsSize);
272
273
            free(words);
274
            free(text);
275
276
            break;
277
          // No se especifica nombre de archivo
278
279
          case '?':
280
            // Ejecutar bubblesort con texto desde entrada estandar
281
            if(optopt == 'b')
282
            {
283
284
              // Cargamos texto
              text = cargarTextosDeEntradaEstandar();
285
286
              // Si no se ingreso texto, salimos
              if (text == NULL) break;
288
              wordsSize = to_words(text, &words);
289
```

```
if (wordsSize == -1) break;
292
              // Ejecutamos bubblesort
293
              bubblesort(words, wordsSize);
294
295
              // Enviamos a salida estandar
              ordenamientoImprimirSalidaEstandar(words, wordsSize);
297
298
              free(words);
              free(text);
300
301
              break;
            }
302
303
            // Ejecutar shellsort con texto desde entrada
            // estandar
305
            else if(optopt == 's')
306
            {
              // Cargamos texto
308
              text = cargarTextosDeEntradaEstandar();
309
310
              // Si no se ingreso texto, salimos
311
312
              if (text == NULL) break;
              wordsSize = to_words(text, &words);
313
314
              if (wordsSize == -1) break;
316
317
              // Ejecutamos shellsort
              shellsort(words, wordsSize);
318
319
              // Enviamos a salida estandar
              ordenamientoImprimirSalidaEstandar(words, wordsSize);
321
322
              free(words);
              free(text);
324
325
              break;
326
327
328
          default:
            opterr = 1;
329
330
            break;
331
332
333
334
      // Manejo de casos en que no se ingresan argumentos validos
     if(opterr || argc <= 1)</pre>
335
336
        fprintf (stderr, "Los argumentos no son validos.\n");
337
        printf("Tipee 'tp0 -h' para ver el modo de uso.\n");
338
339
340
341
     return 0;
342
343 }
```

A.2. bubblesort.h. Declaración de la librería Bubblesort

```
14
15
16 // Funcion que aplica el algoritmo de ordenamiento Bubblesort para ordenar un
17 // arreglo de palabras.
18 // PRE: 'words' es un puntero a un arreglo de punteros a caracter; 'arraysize'
19 // es el tamanio de dicho arreglo.
20 // POST: el arreglo 'words' queda ordenado.
21 void bubblesort(char* words[], int arraysize);
22
23
24
25 #endif
```

A.3. bubblesort.c. Definición de la librería Bubblesort

```
Código 8: "bubblesort.c"
  * LIBRERIA BUBBLESORT
#include "bubblesort.h"
#include <stdbool.h>
12 #include <string.h>
14
15
16 // Funcion que aplica el algoritmo de ordenamiento Bubblesort para ordenar un
17 // arreglo de palabras.
18 // PRE: 'words' es un puntero a un arreglo de punteros a caracter; 'arraysize'
19 // es el tamanio de dicho arreglo.
20 // POST: el arreglo 'words' queda ordenado.
void bubblesort(char* words[], int arraysize)
22 {
    // Variables de procesamiento
23
24
    bool huboIntercambio;
    int i;
25
26
    int n = arraysize;
    char* sAux;
28
    // Recorremos el arreglo haciendo intercambios hasta que ya no se registre
    // ningun cambio realizado.
30
31
    do
32
33
     huboIntercambio = false;
34
      for(i = 1; i < n; i++)</pre>
35
36
       // Si el de indice menor es mayor que el de indice superior, los
37
       // intercambiamos
38
       if(strcasecmp(words[i-1], words[i]) > 0)
39
40
         sAux = words[i-1];
41
42
         words[i-1] = words[i];
         words[i] = sAux;
43
44
         // Cambiamos el flag para registrar que hubo un cambio
45
         huboIntercambio = true;
46
       }
47
48
49
     // Como el elemento del indice superior se encuentra ya ordenado una
50
     // vez finalizada la pasada, se reduce en uno la cantidad de indices
51
     // a iterar en la proxima pasada.
52
53
     n -= 1;
```

```
55  } while(huboIntercambio);
56 }
```

A.4. shellsort.h. Declaración de la librería Shellsort

A.5. shellsort.c. Definición de la librería Shellsort

```
Código 10: "shellsort.c"
* LIBRERIA SHELLSORT
  #include "shellsort.h"
#include <string.h>
15 // Funcion que aplica el algoritmo de ordenamiento Shellsort para ordenar un
16 // arreglo de palabras.
17 // PRE: 'words' es un puntero a un arreglo de punteros a caracter; 'arraysize'
18 // es el tamanio de dicho arreglo.
19 // POST: el arreglo 'words' queda ordenado.
void shellsort(char* words[], int arraysize)
21 {
   int intervalo, k, j, i;
22
23
   char* sAux;
24
   // Tomamos como intervalo el modulo de la mitad del arreglo
25
   intervalo = arraysize / 2;
27
```

```
// Iteramos hasta que el intervalo sea nulo
    while(intervalo > 0)
29
30
      for(i = intervalo-1; i < arraysize; i++)</pre>
31
32
33
         j = i - intervalo;
34
         // Procesamos hasta caer en un indice fuera de rango del arreglo
35
         while(j >= 0)
36
37
           k = j + intervalo;
38
39
           // Si el de indice menor es mayor que el de indice superior, los
40
           // intercambiamos
           if(strcasecmp(words[j], words[k]) > 0)
42
43
           {
             sAux = words[j];
             words[j] = words[k];
45
46
             words[k] = sAux;
47
           else
48
49
             j = 0;
50
           j -= intervalo;
51
52
53
54
       // Tomamos como nuevo intervalo el modulo de la mitad del intervalo
55
       // anterior
56
57
      intervalo = intervalo / 2;
58
    }
59 }
```

A.6. fileloader.h. Declaración de la librería File Loader

```
Código 11: "fileloader.h"
* LIBRERIA DE CARGA DE ARCHIVO DE PALABRAS
11 #ifndef FILELOADER_H
12 #define FILELOADER_H
14 #define INITIAL_BUFFER_SIZE 100
16 #include <stdbool.h>
17
19 // Funcion que carga un archivo de texto a memoria
20 // PRE: 'fileUrl' es la url completa al archivo a ordenar.
21 // POST: devuelve la lista de palabras cargadas en memoria
22 char* file_loader(const char * fileUrl);
24 // Funcion que separa las palabras de un texto en unidades.
25 // PRE: 'text' es una cadena que contiene el texto; 'result' es un puntero en donde se
26 // insertan las palabras del texto.
27 // POST: en 'result' se almacenaron las palabras.
int to_words(char * text, char *** result);
_{\mbox{\scriptsize 30}} // Funcion que verifica si un caracter es un fin de palabra o un separador.
31 // PRE: 'letter' es el caracter a procesar.
32 // POST: devuelve true si el caracter es un fin de palabra o separador, o
33 // false en su defecto.
```

```
34 bool isEndOfWord(char letter);
35
36 #endif
```

A.7. fileloader.c. Definición de la librería File Loader

```
Código 12: "fileloader.c"
   * LIBRERIA DE CARGA DE ARCHIVO DE PALABRAS
10 #include <stdio.h>
11 #include <stdlib.h>
12 #include <string.h>
13 #include "fileloader.h"
15
16
18 // Funcion que carga un archivo de texto a memoria
19 // PRE: 'fileUrl' es la url completa al archivo a ordenar.
20 // POST: devuelve la lista de palabras cargadas en memoria
char * file_loader(const char * fileUrl)
22 {
     FILE * file;
23
     long fileSize;
24
25
     char * buffer;
     size_t result;
26
     file = fopen (fileUrl, "r");
29
     if (file==NULL)
30
31
      //Fail to load file
32
      fputs ("The file can not be open or not exist \n", stderr);
33
      return NULL;
34
35
     //Get file size
37
     fseek (file , 0 , SEEK_END);
38
39
     fileSize = ftell (file);
     rewind (file);
40
41
     //Allocate memory for the whole file
42
     buffer = (char *) malloc ((sizeof(char) * fileSize) + 1);
43
45
     if (buffer == NULL)
46
      fputs ("There is no enoght memory to allocate the file in memory \n", stderr);
      return NULL;
48
49
50
     //Copy the file to the buffer
51
52
     result = fread (buffer, 1, fileSize, file);
           buffer[(sizeof(char) * fileSize)] = '\0';
53
54
     if (result != fileSize)
55
56
      fputs ("Fail to read the whole file \n", stderr);
57
      return NULL;
58
59
     /* the whole file is now loaded in the memory buffer. */
61
```

```
fclose (file);
63
64
65
      return buffer;
66 }
67
69 // Funcion que separa las palabras de un texto en unidades.
70 // PRE: 'text' es una cadena que contiene el texto; 'result' es un puntero en donde se
71 // insertan las palabras del texto.
72 // POST: en 'result' se almacenaron las palabras.
int to_words(char * text, char *** result)
74 {
75
     char ** words;
76
     char ** reallocatedWords = NULL;
77
     bool isNewWord = true;
78
     size_t textSize = strlen(text);
     int wordsIndex = 0;
     size_t bufferSize = INITIAL_BUFFER_SIZE * sizeof(char*);
80
     size_t spaceForWords = INITIAL_BUFFER_SIZE;
81
82
     //initialy allocate memory for 100 words
83
     words = (char **) malloc ( INITIAL_BUFFER_SIZE * sizeof(char*) );
84
85
     if(words == NULL)
86
87
       fputs ("Not enough memory to allocate the words \n", stderr);
88
89
       return -1;
90
91
     for (int i = 0; i < textSize; i++)</pre>
92
93
       if ( !isEndOfWord (text[i]) )
94
95
       {
         if(isNewWord)
96
97
           words[wordsIndex] = &text[i];
98
           isNewWord = false;
99
100
           if(wordsIndex >= spaceForWords-1)
101
102
              //Reallocate for more space and copy the old array to the new one
              bufferSize *= 2;
104
              spaceForWords *= 2;
105
              reallocatedWords = realloc(words, bufferSize);
106
107
108
             if(reallocatedWords != NULL)
              {
109
                words = reallocatedWords;
110
                reallocatedWords = NULL;
              }
              else
113
114
              {
                fputs ("Not enough memory to allocate the words \n", stderr);
116
                return -1;
             }
117
           }
118
119
           wordsIndex++;
120
         }
122
       else
124
         text[i] = '\0';
125
         isNewWord = true;
126
127
     }
128
129
     *result = words;
130
131
132
     return wordsIndex;
133 }
134
136 // Funcion que verifica si un caracter es un fin de palabra o un separador.
137 // PRE: 'letter' es el caracter a procesar.
138 // POST: devuelve true si el caracter es un fin de palabra o separador, o
```

B. Implementación completa de Shellsort en assembly MIPS

B.1. *tp0.c.* Implementación del main del programa

```
Código 13: "tp0.c"
* PROGRAMA DE ORDENAMIENTO DE PALABRAS
  * Facultad de Ingenieria - UBA
   * 66.20 Organizacion de computadoras
   * Trabajo Practico: Conjunto de instrucciones MIPS
  * ALUMNOS:
10
  * Facundo Rossi (86707) - frossi85@gmail.com
   * Federico Colangelo (89869) - federico.colangelo@semperti.com
  * Federico Martin Rossi (92086) - federicomrossi@gmail.com
15
  * Programa que permite aplicar metodos de ordenamiento a un texto o textos
  * especificados por el usuario. Estos textos pueden ser ingresados a traves de
  * la entrada estandar o bien a traves de la especificacion de los nombres de
  * archivo que se desean procesar. Cabe resaltar que al especificar varios
  * nombres de archivos, los metodos seran aplicados al conjunto total de los
  * textos contenidos en estos. Es decir, el resultado del ordenamiento sera
  * unico ya que contendra la mezcla de palabras de todos los archivos.
  * Los metodos de ordenamiento soportados por la presente version del programa
26
27
       - Shellsort
   */
29
31 include <stdio.h>
32 #include <ctype.h>
33 #include <stdlib.h>
34 #include <string.h>
35 #include <unistd.h>
#include "fileloader.h"
39 extern int shellsort(char **, size_t);
  int main (int argc, char** argv)
41
42 {
    char* texto;
43
    char** words;
   size_t size;
45
   int i;
    texto=file_loader((charargv[1]);
48
    size=to_words(texto,&words);
50
    shellsort(words, size);
51
    for(i=0;i<size;i++)</pre>
53
     puts(words[i]);
```

```
55 }

56 
57     return 0;
58 }
```

B.2. shellsort.S. Implementación del algoritmo Shellsort

```
Código 14: "shellsort.S"
#ifndef USE_MIPS_ASSEMBLY
2 #define USE_MIPS_ASSEMBLY
  #include <mips/regdef.h>
#include <sys/syscall.h>
6 .text
7 .align 2
8 .globl shellsort
9 .extern swap
10 .extern compare
12
13 shellsort: subu sp,sp,24 # creo stack frame, es de 24 porque necesito guardar ra
      sw ra,16(sp)
14
15
      sw $fp,12(sp)
16
      sw gp,8(sp)
      move $fp,sp
17
18
19
      #Guardo los parametros recibidos en fp
      sw a0,24($fp) # words = char **
20
21
      sw a1,28($fp)
                         # size
22
      #Guardo los parametros cargados en fp a variables temporales
23
24
      lw s6, 24($fp)
                         #en 24 esta el words
      lw s7, 28($fp)
                          #en 28 esta el size
25
26
      #Iniciaizo variables locales
      srl s0, s7, 1 # intervalo = size/2
28
                  # k
29
      li s1, 0
      li s2, 0
                   # j
30
      li s3, 0
                   # i
31
32
      li s4, 0
                   # sAux
33
34
35
      #add v0, s0, zero
      #j fin
36
37
38
      #while(intervalo > 0)
39 loop1: bgt s0, zero, initloop2 #branch grater than, salta si s0 es mayor a 0
41
      #for(i = intervalo-1; i < arraysize; i++)</pre>
42
43 initloop2: addi s3, s0, -1  # i = interlvalo - 1
  loop2: bge s3, s7, endloop2 # si i >= size fin de for, bge=branch greater than or equal
    sub s2, s3, s0 #j = i - intervalo;
45
47
      \#while(j >= 0)
48
  loop3: bge s2, zero, loop3task
      addi s3, s3, 1
50
51
       j loop2
52
53
56 loop3task: add s1, s2, s0
                                \#k = j + intervalo;
57
       #Para comparar cadenas, la forma mas sencilla es iterar sobre cada caracter de cada cadena, y restarlos. \hookleftarrow
58
           Si el resultado es 0, son
_{59} #iguales. Si no, a continuacion, si el resultado es> 0, entonces la primera cadena es mayor de que la otra \leftrightarrow
       cadena, de lo contrario la segunda cadena es inferior y #habria que intercambiarlas. Si te quedas sin \hookleftarrow
```

```
alguna de las cadenas antes que el otro, y son iguales hasta llegar a ese punto, la cadena mas corta es \leftrightarrow
       menor que el que #mas.
       #CARGO LAS CADENAS A COMPARAR
60
61
       #words[j]
       ##Guardo en variables auxiliares las direcciones de memoria a intercambiar
62
       #words = s6
       #Guardo i y j por q las voy a modificar y no las quiero perder
64
       sw s1, 0($fp)
65
       sw s2, 4($fp)
66
67
68
69
       sll s1, s1, 2
70
71
       sll s2, s2, 2
       addu s1, s1, s6
72
       addu s2, s2, s6
       #Cargo los strings a comparar en ax que luego compare utiliza para trabajar
75
       lw a0, 0(s1)
                        #OPTIMIZAR CODIGO PARA ISAR MENOS ti
76
77
       lw a1, 0(s2)
       jal compare
                      #llamo a compare
78
79
                        #Cargo el valor de k en stack
80
       lw s1, 0($fp)
       lw s2, 4($fp)
                        #Cargo el valor de j en stack
81
       bltz v0, swapping #swap si compare > 0
83
84
       add s2, zero, zero #j=0
       j endif
85
86
       # poner el valor a utilizar en los registros correctos y llamar a swap
88 swapping: add a0, s6,0
       add a1, s1, 0
89
       add a2, s2, 0
       jal swap
91
92
             sub s2, s2, s0
                                #j = j - intervalo;
93
       j loop3
94
95
96
   endloop2: srl s0, s0, 1 #intervalo = intervalo / 2;
97
       j loop1
99
100
101
       ##Comienzo retorno
102
103 fin:
           move sp, $fp
       lw gp, 8(sp)
104
       lw $fp, 12(sp)
105
       lw ra, 16(sp)
       addiu sp, sp, 24
107
                 #ra = return address es la direccion a la q quiero volver luego de llamar la funcion
108
       jr ra
110 #endif
```

B.3. compare.S. Implementación del comparador de palabras

```
Código 15: "compare.S"

#ifndef USE_MIPS_ASSEMBLY
#define USE_MIPS_ASSEMBLY
#include <mips/regdef.h>
#include <sys/syscall.h>

text
align 2
globl compare

# t1 = posicion caracter actual en cadena 1
# t2 = posicion caracter actual en cadena 2
```

```
14
15 compare:
                       # creo stack frame
16
      subu sp,sp,16
       sw $fp,12(sp)
17
18
       sw gp,8(sp)
      move $fp,sp
19
20
21
       add t0, zero, zero # result =0
       add t1, zero, a0
22
       add t2, zero, a1
23
24
25 nextchar:
       #Cargo los elementos a comparar en t4 y t5
       lb t3, 0(t1)
27
      lb t4, 0(t2)
28
      #Voy a pasar todas las mayusculas a minusculas para que compare sea case insenstive como en C
30
31
32
       addi t5, t3, -91
       bgtz t5, aMinusculas2
33
34
       addi t5, t3, -64
       bltz t5, aMinusculas2
35
       addi t3, t3, 32
36
37
38 aMinusculas2: addi t5, t4, -91
39
       bgtz t5, comparacion
       addi t5, t4, -64
40
       bltz t5, comparacion
41
42
       addi t4, t4, 32
43
44
       #Comparo caracter para ver si es final del string
46 comparacion: beqz t3, fincadena1
47 beqz t4, fincadena2
49 continue: sub t0, t3, t4
                               # cadena1[i] - cadena2[i]
       add t1, t1, 1 #NO VA sll x,y,2: por q char es un byte
       add t2, t2, 1
51
52
53
       beqz t0, nextchar
       j fin
54
55
56
57 fincadenal: beqz t4, fin
58
       add t0, zero, -1
       j fin
59
61 fincadena2: add t0, zero, 1
      j fin
62
63
           ##Comienzo retorno
      addi v0, t0, 0
65
66
       move sp, $fp
67
       lw gp, 8(sp)
68
       lw $fp, 12(sp)
       addiu sp, sp, 16
70
                 #ra = return address es la direccion a la q quiero volver luego de llamar la funcion
73 #endif
```

B.4. swap.S. Implementación del comparador de palabras

```
Código 16: "swap.S"

#ifndef USE_MIPS_ASSEMBLY
#define USE_MIPS_ASSEMBLY
#include <mips/regdef.h>
#include <sys/syscall.h>
```

```
6 .text
7 .align 2
8 .globl swap
11
12 swap:
13
      subu sp,sp,16 # creo stack frame
      sw $fp,12(sp)
14
15
      sw gp,8(sp)
16
      move $fp,sp
17
      #Guardo los parametros recibidos en fp
      sw a0,16($fp)
                      # guardo los primeros 4 parametros con la llamo a checksum
19
                        # 20-16 = 4Bytes=32bits
      sw a1,20($fp)
20
      sw a2,24($fp)
22
      ##Guardo en variables auxiliares las direcciones de memoria a intercambiar
23
24
      lw t0, 16($fp)
                        #en 16 esta el words
      lw t1, 20($fp)
lw t2, 24($fp)
                         #en 20 esta el i
25
                        #en 20 esta el j
26
27
      ##Realizo el intercambio
28
      \#t3 = direction de fp+(i*4)
      \#t4 = direction de fp+(j+4)
30
31
      sll t1, t1, 2
      sll t2, t2, 2
32
      addu t1, t1, t0
33
34
      addu t2, t2, t0
35
36
37
      #Cargo los elementos a intercambiar en t3 y t4
      lw t3, 0(t1)
38
      lw t4, 0(t2)
39
40
      #Guardo los elementos de manera intercambiada
41
      sw t4, 0(t1)
42
      sw t3, 0(t2)
43
44
45
      lw v0, 0(t2)
46
47
48
      ##Comienzo retorno
      move sp, $fp
49
50
      lw gp, 8(sp)
      lw $fp, 12(sp)
51
      addiu sp, sp, 16
52
                #ra = return address es la direccion a la q quiero volver luego de llamar la funcion
54
55 #endif
```