

PRÁCTICA 3

Recuperación de la información



12 DE NOVIEMBRE DEL 2018
VÍCTOR ANDRÉS VIZÁN

Contenido

1.	Introducción	2
2.	Desarrollo	3
3.	Pruebas.....	7
4.	Conclusiones y observaciones.....	9
5.	Bibliografía	10

1. Introducción

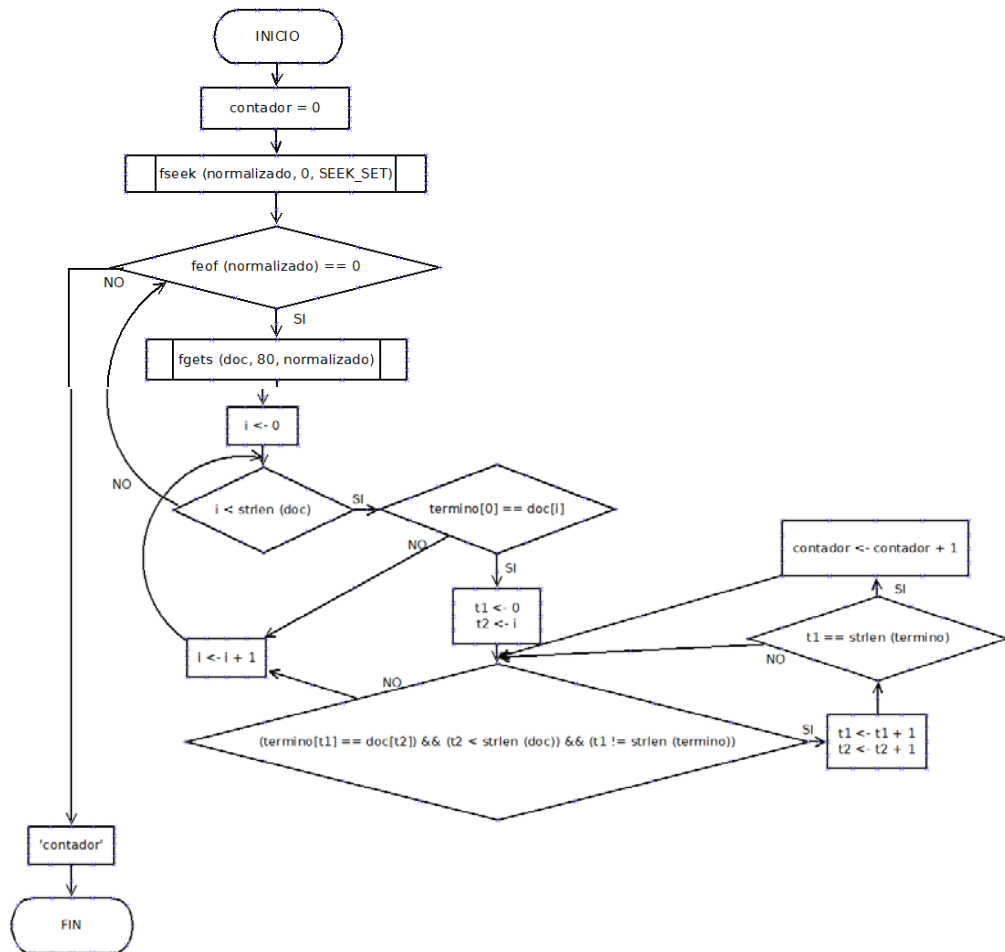
Esta tercera práctica de la asignatura usa los códigos de la anterior práctica para hacer una consulta al programa mediante el método vectorial.

El programa se estructura en diversos puntos:

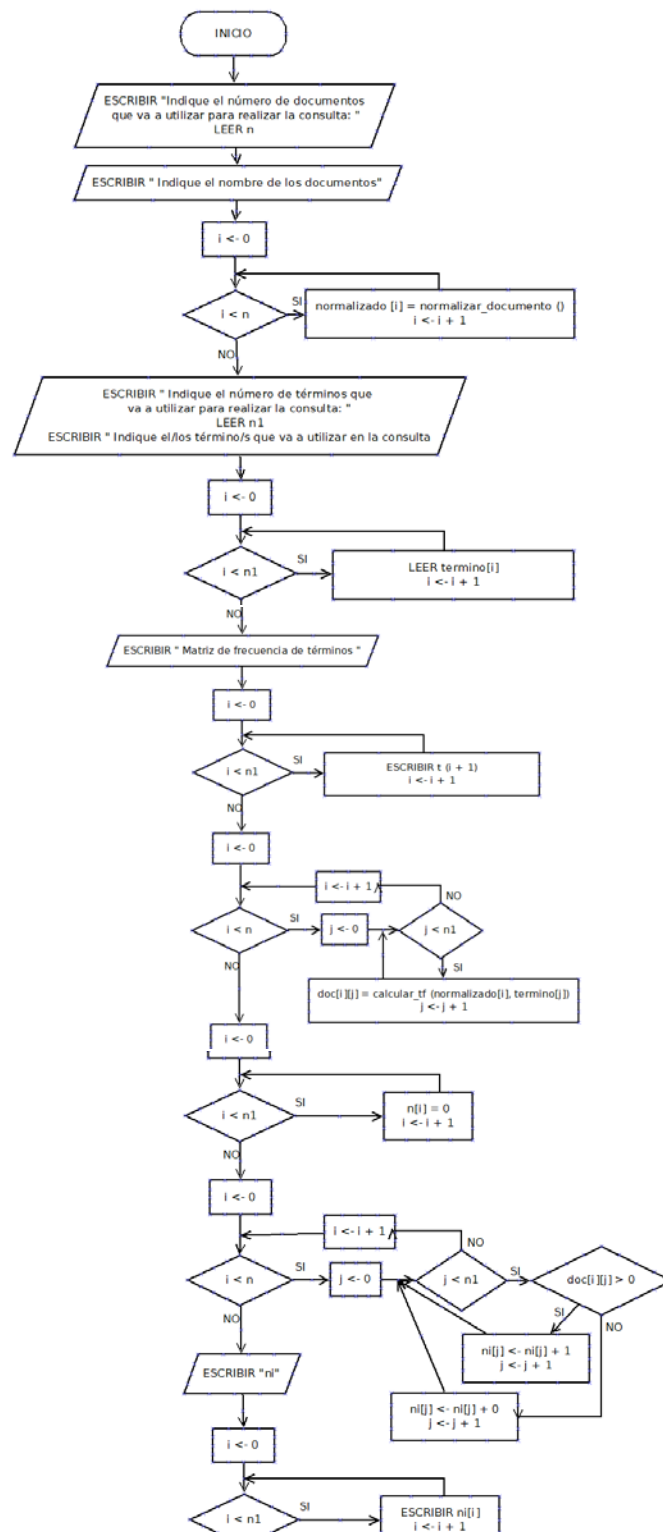
- Seguir un menú con diferentes opciones.
- La primera opción del menú te permite normalizar documento como en la práctica 1 y 2.
- La segunda opción te permite hacer una consulta mediante el método vectorial introduciendo la consulta y el nombre de los documentos por teclado.
- La tercera opción te permite salir del programa.

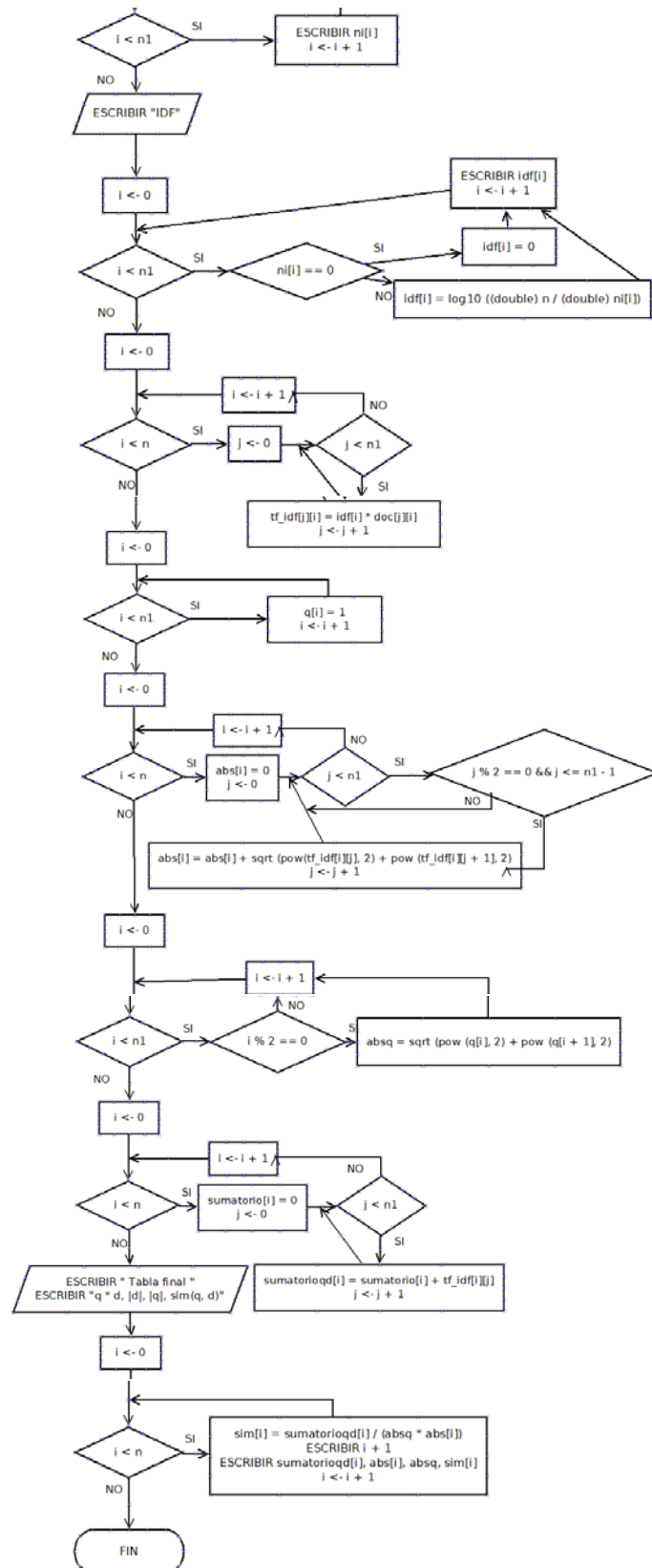
2. Desarrollo

Función que busca un término introducido por parámetro en un documento (buscar_termino).

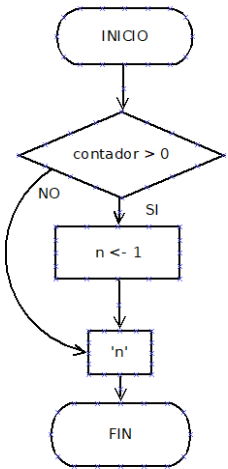


Módulo que hace una consulta





Módulo que calcula el ni, que se utiliza en el módulo anterior consulta ().



3. Pruebas

Las siguientes capturas de pantalla muestran las diferentes opciones que el programa creado

En esta primera captura de pantalla se puede ver como al escoger la opción 1 del programa normaliza el documento introducido.

```
victor@Algoritmia:~/Escritorio$ gcc Practica3RI.c -lm
victor@Algoritmia:~/Escritorio$ ./a.out

Menu
1. Normalizar documento.
2. Hacer consulta.
3. Salir del programa.
Respuesta: 1

Indique el nombre del documento
documento1.txt

Documento normalizado.
```

En estas tres capturas siguientes se muestra como introducir la información para que el programa haga la consulta. En la segunda imagen te muestra la información que el programa muestra. Y la tercera muestra el resultado de la consulta, la información obtenida.

```
Menu
1. Normalizar documento.
2. Hacer consulta.
3. Salir del programa.
Respuesta: 2

Indique el número de documentos que va a utilizar para realizar la consulta: 4

Indique el nombre de los documentos
documento1.txt
documento2.txt
documento3.txt
documento4.txt

Indique el número de términos que va a utilizar para realizar la consulta: 5

Indique el término/s que va a utilizar en la consulta
barco
casa
coche
gato
perro
```

Matriz de frecuencia de términos					
	t1	t2	t3	t4	t5
d1	6	0	0	4	0
d2	0	6	4	2	2
d3	0	0	2	0	4
d4	2	4	4	2	0
ni	2	2	3	3	2
IDF	0.301030	0.301030	0.124939	0.124939	0.301030

Tabla final				
	$q * d$	$ d $	$ q $	$\sin(q, d)$
d1	2.305935	2.305935	1.000000	1.000000
d2	3.157872	2.966983	1.000000	1.064338
d3	1.453997	1.596125	1.000000	0.910955
d4	2.555812	1.904990	1.000000	1.341641

Y en esta última captura de pantalla se muestra la tercera opción, que es la de salir del programa.

```
Menu
1. Normalizar documento.
2. Hacer consulta.
3. Salir del programa.
Respuesta: 3
victor@Algoritmia:~/Escritorio$
```

4. Conclusiones y observaciones

Casi todo el código es reutilizado de la práctica 2. La diferencia está en el módulo consultar, en el de buscar_término y y en el de calcular_ni. Esto hace la diferencia entre las prácticas, la segunda práctica el modelo booleano, y este, vectorial.

Se ha reutilizado casi todo el código de la primera y segunda práctica. Ha habido que cambiar un par de cosas al módulo normalizar_documento (), para que en vez de devolver nada devolviera el fichero normalizado para poder utilizarlo luego en otra función.

Al final con estas prácticas además de seguir la asignatura, aprendes C y te pones a pensar en cómo hacer un programa de forma autodidacta.

5. Bibliografía

[1] "C library function tolower()"

https://www.tutorialspoint.com/c_standard_library/c_function_tolower.htm

[2] "Dia" <http://dia-installer.de/index.html.es>

[3] "C | Mi Word Press" <https://pasky.wordpress.com/category/programacion/c/>