



---

# PRÁCTICA 1

---

Recuperación de la información



17 DE OCTUBRE DE 2018  
VÍCTOR ANDRÉS VIZÁN

# Contenido

1.	Introducción .....	2
2.	Desarrollo .....	3
3.	Pruebas.....	11
4.	Conclusiones y observaciones.....	12
5.	Bibliografía .....	13

# 1. Introducción

Esta primera práctica de la asignatura trata sobre manipular un documento de texto.

El programa se estructura en diversos puntos:

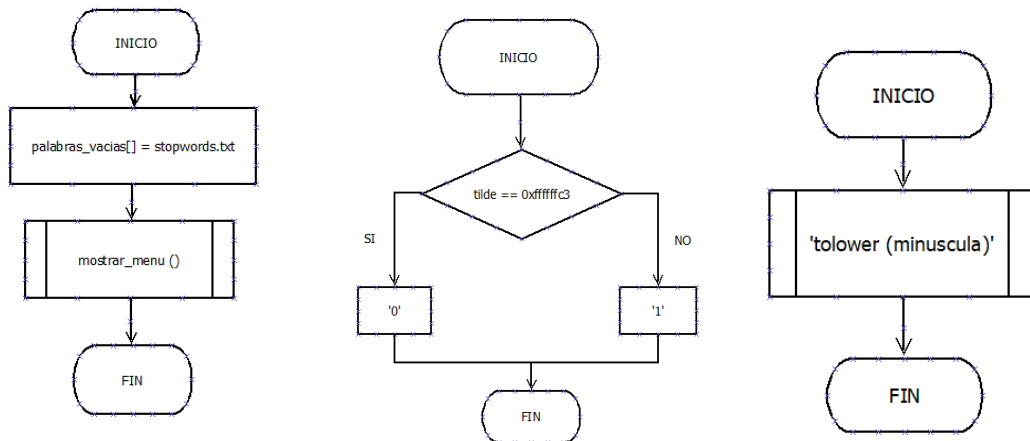
- Seguir un menú con diferentes opciones.
- Mostrar el texto tal cual está en el documento.
- Mostrar el texto sin tildes y usando otro documento de palabras vacías y eliminar los signos de puntuación.

Para este último punto es necesario crear funciones tal que elimine las tildes del texto, para quitar los signos de puntuación, para introducir las palabras vacías, o para crear otro documento para introducir el texto limpio y normalizado.

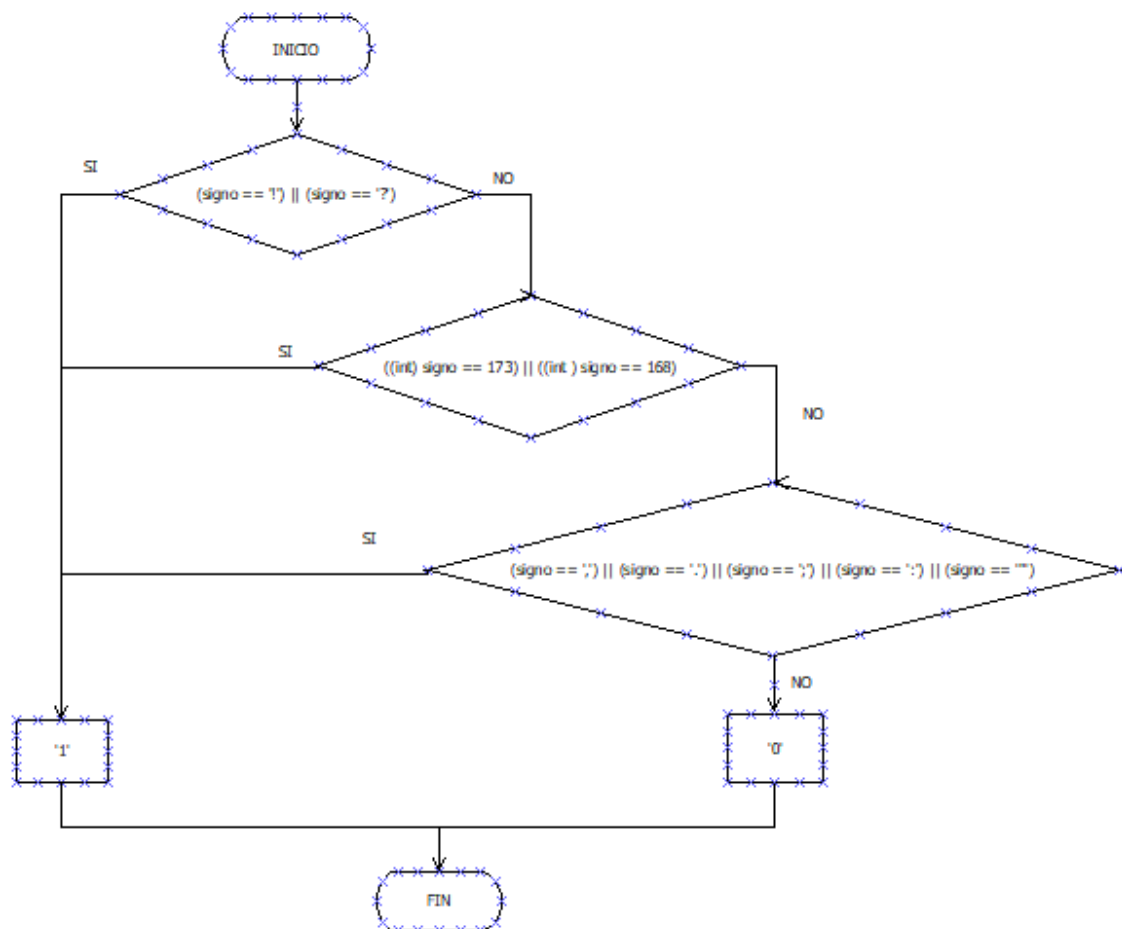
El fin es que el texto en castellano pueda ser leído por otro tipo de máquinas y que no produzcan errores, como, por ejemplo, algunos caracteres extraños.

## 2. Desarrollo

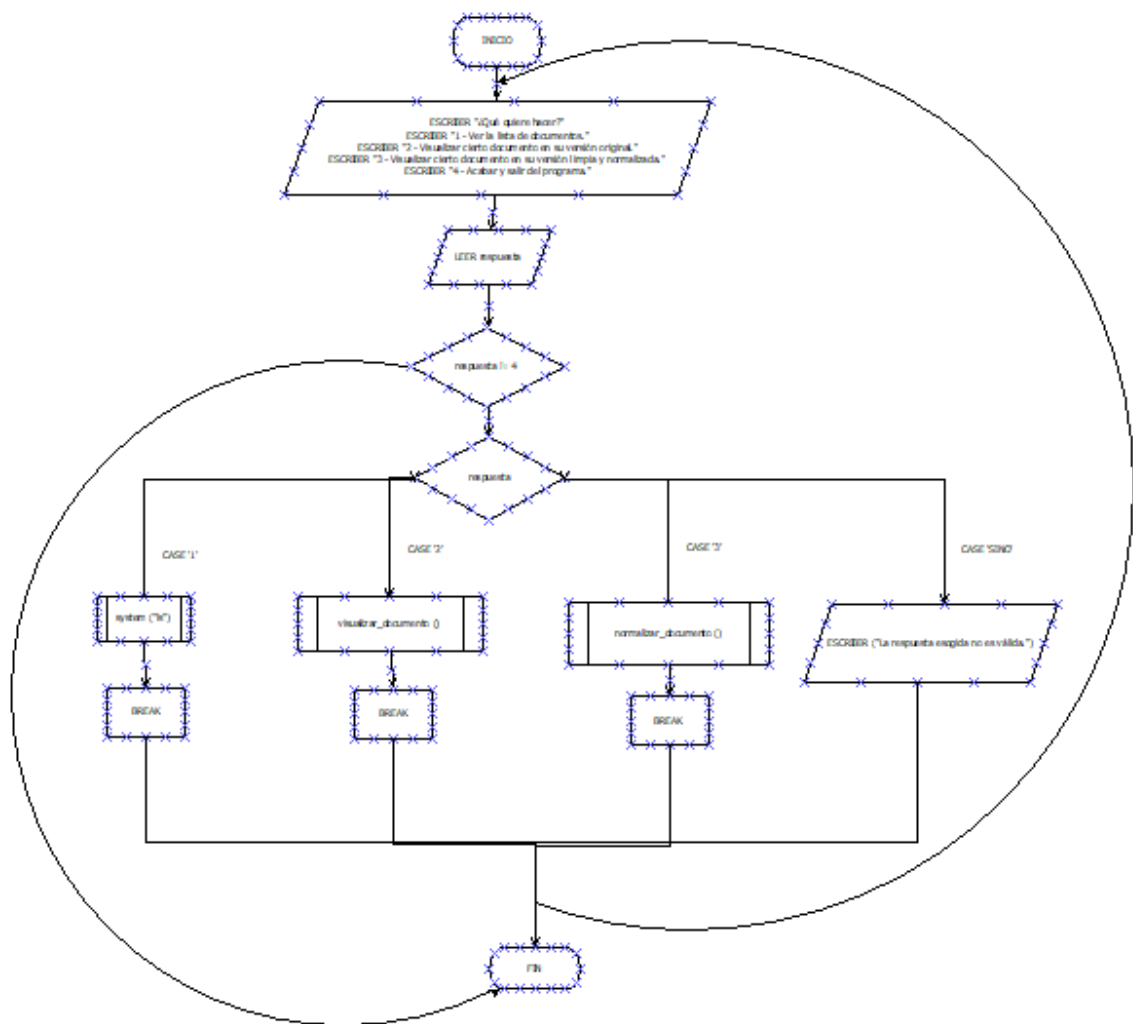
El primer diagrama de flujo es de la función principal **main**. El segundo del módulo **comprobar\_tildes (char tilde)**. Y el tercero del módulo **masyusculas\_a\_minusculas (char minúscula)**.



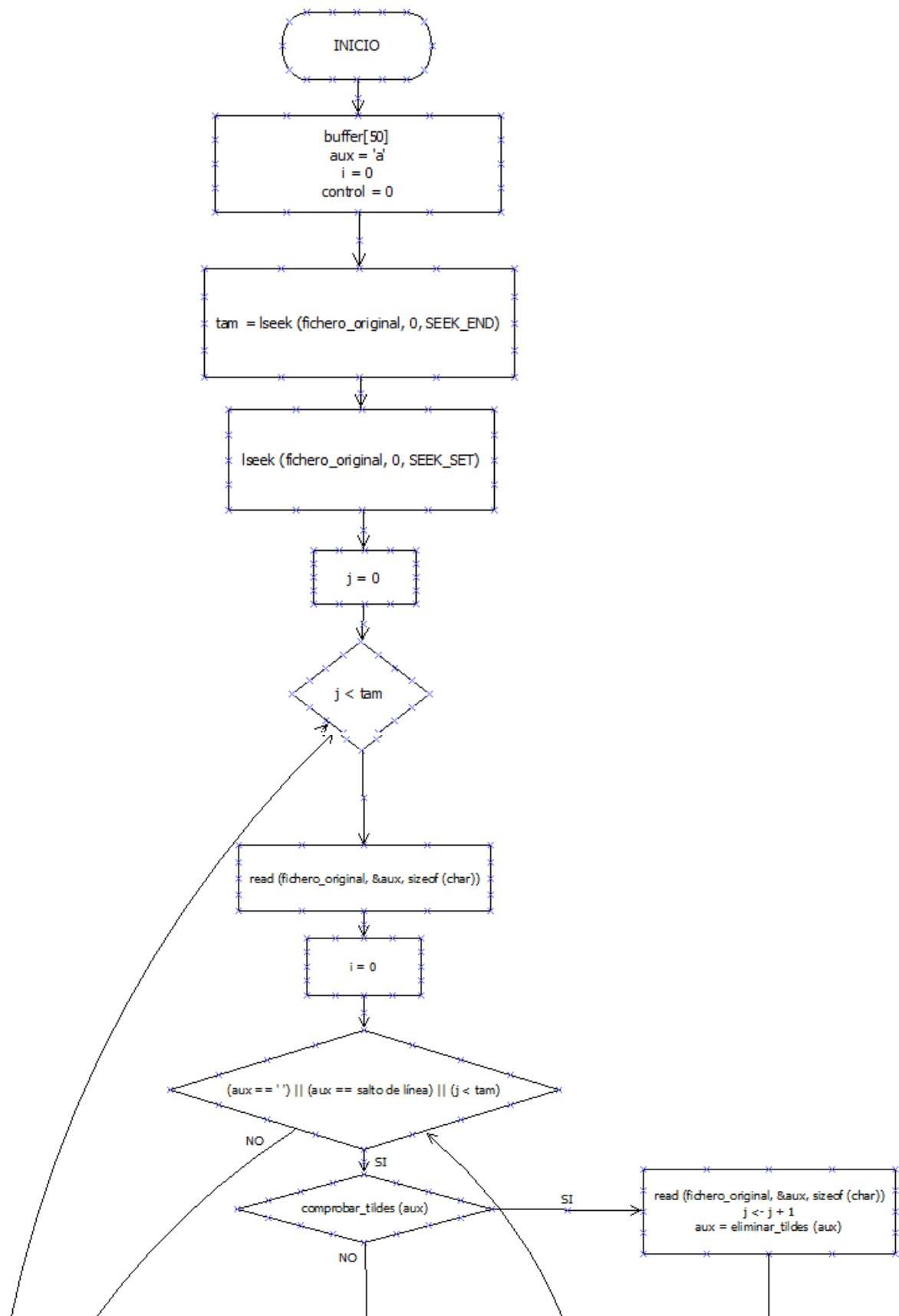
El siguiente diagrama de flujo es de la función **eliminar\_signos (char signo)**.

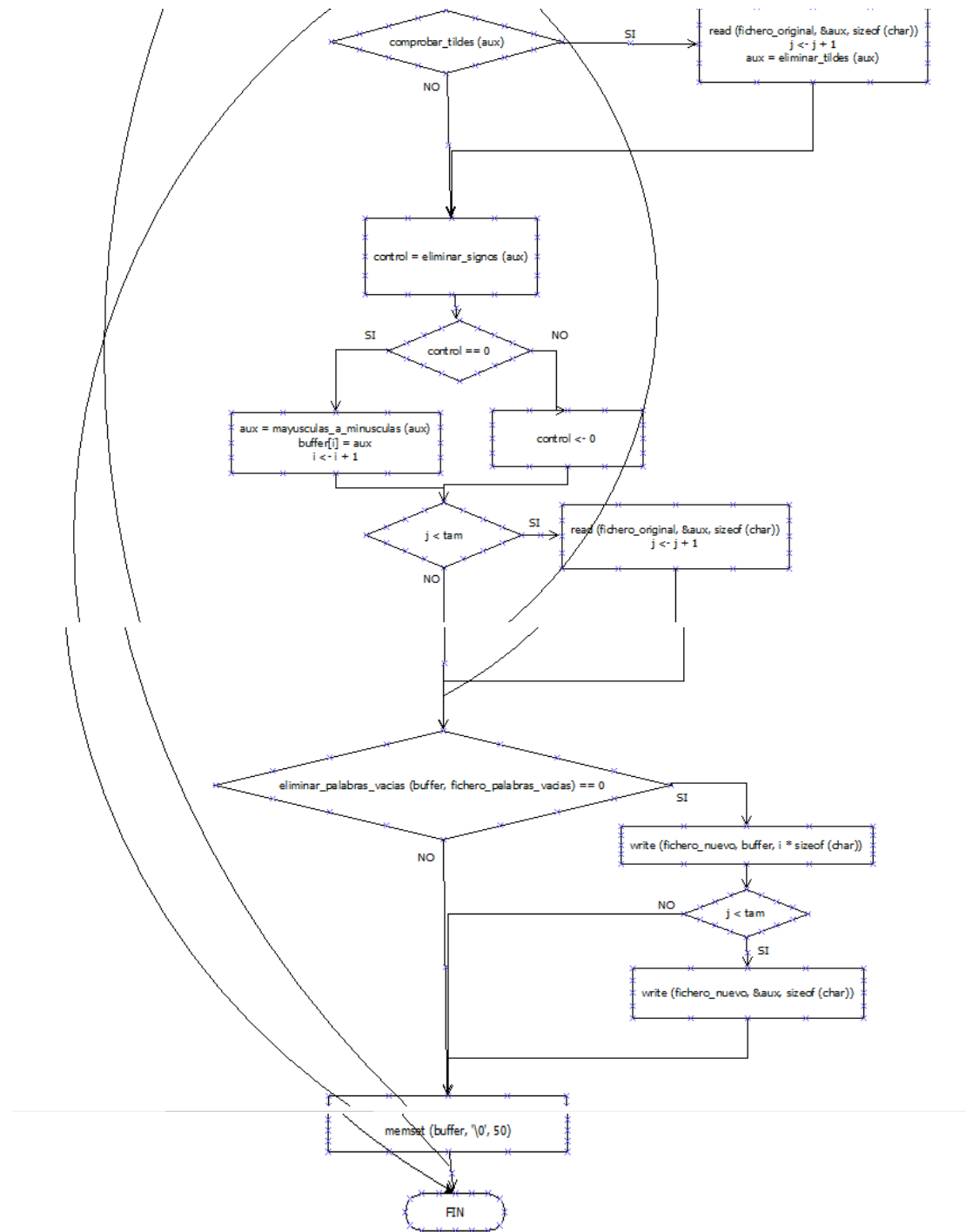


El siguiente de la función `mostrar_menu ()`.

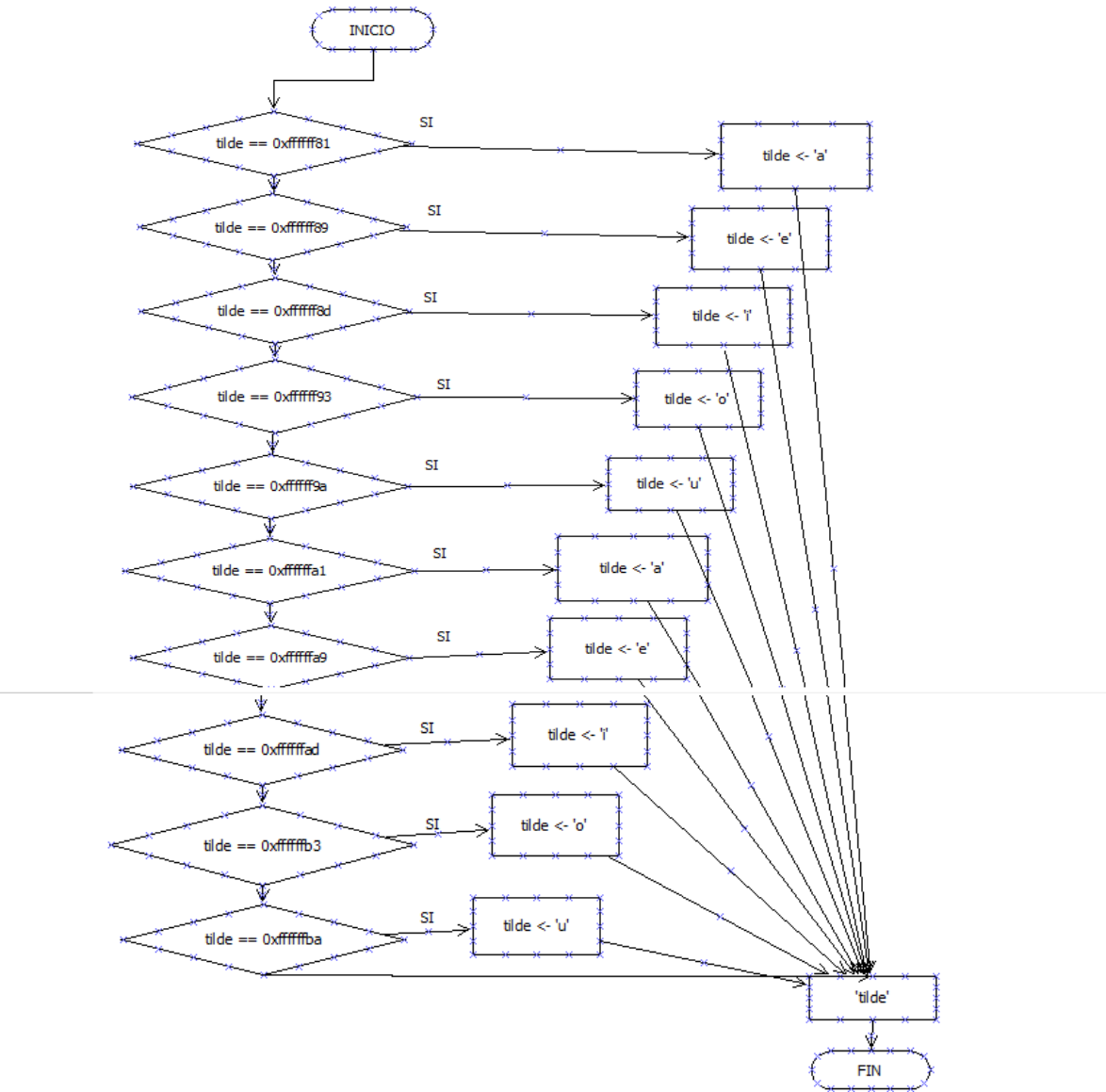


El siguiente del módulo **buscar\_palabras\_vacias** (int fichero\_original, int fichero\_palabras\_vacias, int fichero\_nuevo).



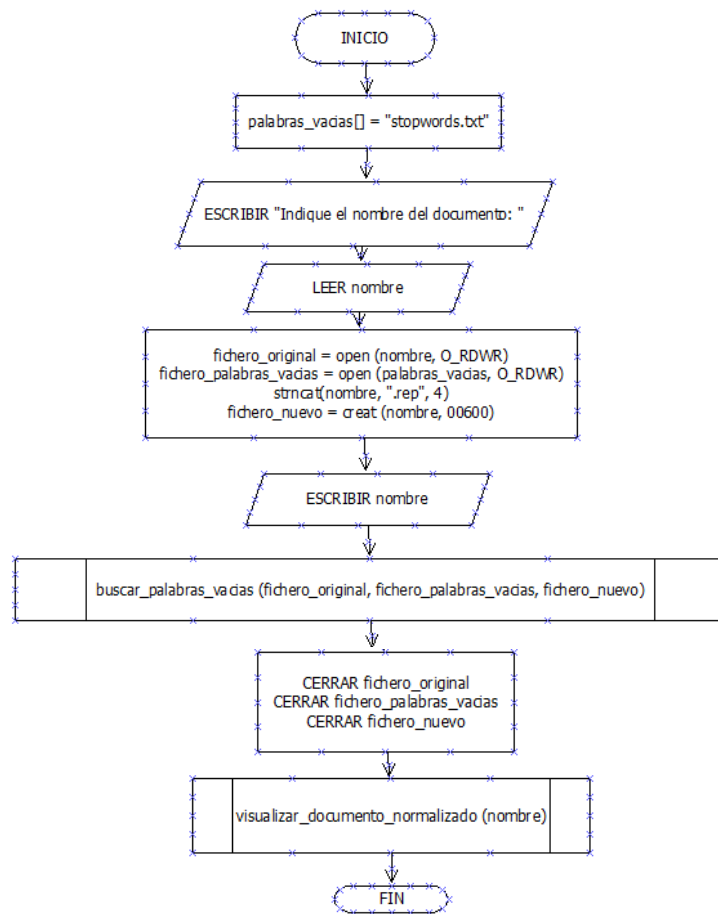


El siguiente diagrama de flujo es de la función **eliminar\_tildes (char tilde)**, que elimina las tildes del texto.

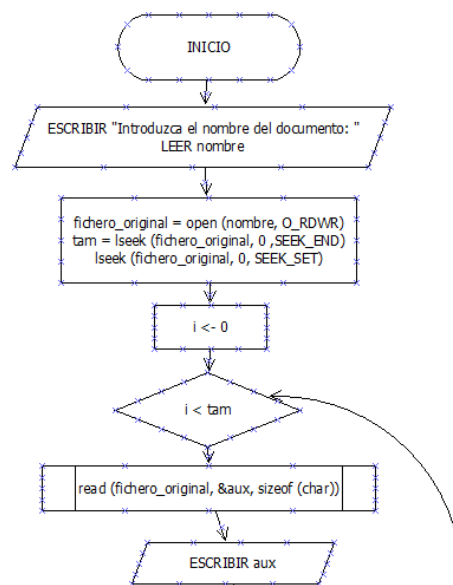


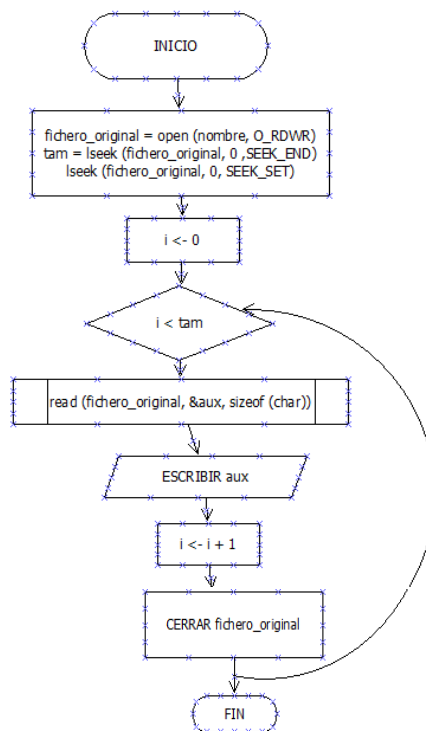
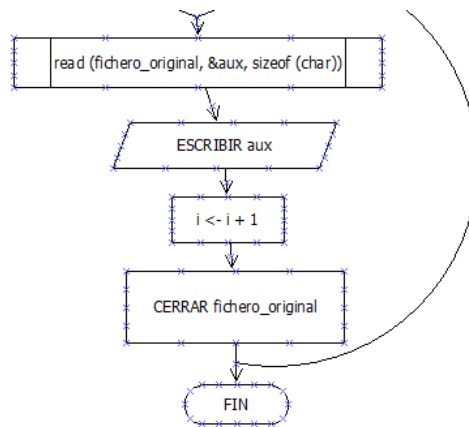


El siguiente diagrama de flujo corresponde al módulo **normalizar\_documento ()**.

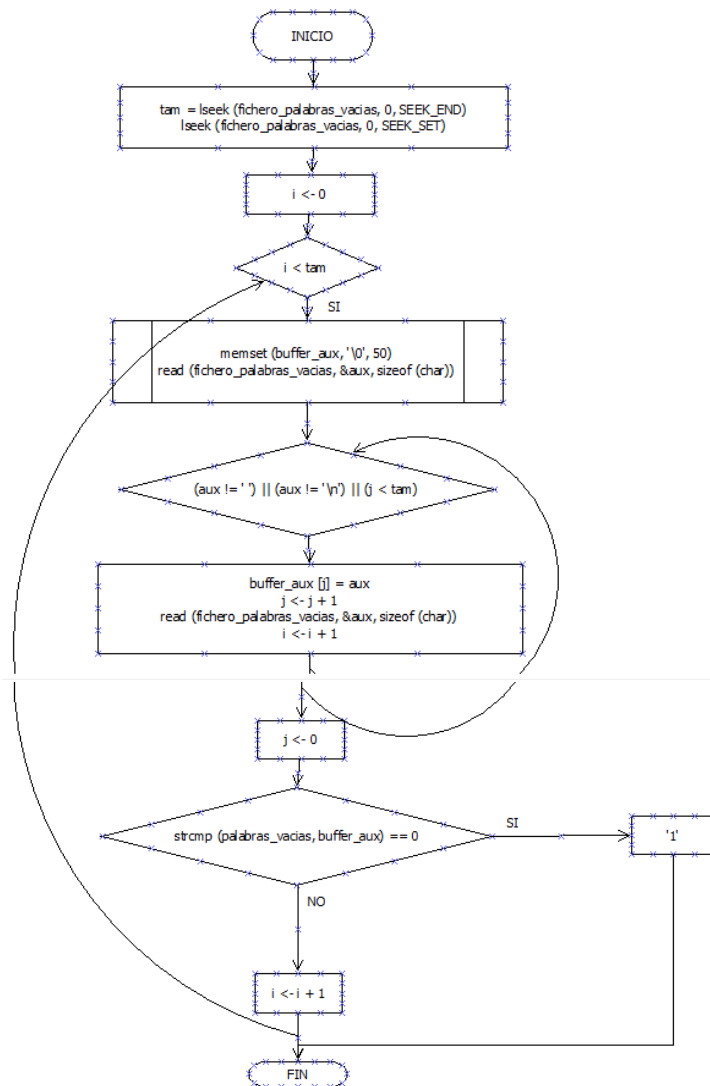


Los siguientes diagramas de flujo corresponden a las funciones **visualizar\_documento ()** y **visualizar\_documento\_normalizado (char \*nombre)** respectivamente.





Y este último diagrama de flujo corresponde a la función del programa **eliminar\_palabras\_vacias (char palabras\_vacias[], int fichero\_palabras\_vacias)**



### 3. Pruebas

Las siguientes capturas de pantalla muestran las diferentes opciones que el programa creado puede llevar a cabo con un ejemplo práctico de un documento .txt.

```
¿Qué quiere hacer?
1 - Ver la lista de documentos.
2 - Visualizar cierto documento en su versión original.
3 - Visualizar cierto documento en su versión limpia y normalizada.
4 - Acabar y salir del programa.

1
a.out parrafos.txt Practica1RI.c primerprograma.cpp prueba.c stopwords.txt
```

```
¿Qué quiere hacer?
1 - Ver la lista de documentos.
2 - Visualizar cierto documento en su versión original.
3 - Visualizar cierto documento en su versión limpia y normalizada.
4 - Acabar y salir del programa.

2
Introduzca el nombre del documento:
parrafos.txt
```

```
Introduzca el nombre del documento:
parrafos.txt
```

Lechuza de campanario enmascarada es el nombre alternativo de esta especie, que es poco conocida más allá de su zona de distribución natural.

Es una lechuza mediana a grande con patas amplias, pico muy corto y oídos no visibles. Carece de mechones largos de plumas a los lados de la parte superior de la cabeza. Los machos adultos miden entre 33 y 41 centímetros de longitud, y las hembras entre 39 y 50 centímetros; por ende, ellas tienden a ser más grandes que sus contrapartes y esto ayuda a diferenciar los sexos. La subespecie de Tasmania es la más grande de todas.

Por otra parte, ellos pesan cerca de 420-800 gramos, y ellas unos 545-1,260 gramos. La envergadura, o distancia entre las puntas de las alas extendidas, registra alrededor de 1,280 centímetros en las hembras del sur.

El rostro posee un disco facial en forma de corazón de color claro en donde destacan dos ojos oscuros. El pico está cubierto casi por completo por las plumas. El pecho y el vientre poseen plumas blancas a rojizas, en tanto el dorso es marrón y la zona superior está decorada con manchones color marrón oscuro. Alrededor del pico y los ojos, el negro predomina. En realidad, existe mucha variación en el color del plumaje; algunos individuos pueden mostrar hasta tonos anaranjados.

La especie vive en Oceanía, en tres regiones: Australia, Indonesia y Papúa Nueva Guinea. El rango total es amplio; está presente en muchísimas zonas no desérticas, por lo que está ausente sobre todo en el centro australiano y en cambio prospera en sus regiones costeras. Nueva Gales del Sur, un estado de Australia, tiene una población de *Tyto novaehollandiae* muy grande: un 90 por ciento de su territorio es su hogar.

Sus hábitats pueden ser bosques secos de eucalipto, campos abiertos, parques y en general, bosques amplios. En Australia, es abundante en los llanos del oeste, a no más de 300 kilómetros al interior del país.

La lechuza australiana mantiene una dieta carnívora en la que predominan presas como zarigüeyas, conejos, bandicuts, murciélagos, roedores (ratas, por ejemplo) y algunas aves, reptiles e insectos. Por lo regular, caza durante las primeras horas de la noche a lo largo de los bosques e incluso a lo largo de carreteras y prefiere animales terrestres, si bien es tan eficaz que puede capturar animales al vuelo.

```
3
Indique el nombre del documento
parrafos.txt
```

```
parrafos.txt.rep
```

```
lechuza campanario enmascarada nombre alternativo especie conocida alla zona distribucion natural
```

```
lechuza mediana grande patas amplias pico corto oidos visibles carece mechones largos plumas lados superior cabeza machos adultos miden 33 41 centimetros longitud hembras 39 50 centimetros ende tienden grandes contrapartes ayuda diferenciar sexos subespecie tasmanita grande todas
```

```
pesan 420-800 gramos 545-1260 gramos envergadura distancia puntas alas extendidas registra 1280 centimetros hembras sur
```

```
rostro posee disco facial forma corazon color destacan ojos oscuros pico cubierto completo plumas pecho vientre poseen plumas blancas rojizas tanto dorso marron zona superior decorada manchones color marron oscuro pico ojos negro predomina realidad existe mucha variacion color plumaje algunos individuos pueden mostrar tonos anaranjados
```

```
especie vive oceania tres regiones australia indonesia papua guinea rango total amplio presente muchisimas zonas deserticas ausente centro australiano cambio prospera regiones costeras gales sur australia poblacion tyto novaehollandiae grande 90 ciento territorio hogar
```

```
habitats pueden bosques secos eucalipto campos abiertos parques bosques amplios australia abundante llanos oeste 300 kilometros interior lechuza australiana mantiene dieta carnivora predominan presas zarigüeyas conejos bandicuts murciélagos roedores (ratas ejemplo) algunas aves reptiles insectos regular caza primeras noche largo bosques largo carreteras prefiere animales terrestres tan eficaz capturar animales vuelo
```

```
¿Qué quiere hacer?
1 - Ver la lista de documentos.
2 - Visualizar cierto documento en su versión original.
3 - Visualizar cierto documento en su versión limpia y normalizada.
4 - Acabar y salir del programa.
```

```
4
victor@Algoritmia:~/Escritorio$
```

## 4. Conclusiones y observaciones

Programar el código para hacer este programa ha sido algo complicado, hay que pensar en muchas funciones, y a la hora de cambiar caracteres extraños es algo complicado, aunque ha estado bien, este tipo de prácticas nos pueden servir en el futuro.

Al elegir la tercera opción del menú, que es la que limpia y normaliza el texto, crea un documento e introduce el texto limpio y normalizado del elegido anteriormente.

## 5. Bibliografía

[1] "C library function tolower()"

[https://www.tutorialspoint.com/c\\_standard\\_library/c\\_function\\_toupper.htm](https://www.tutorialspoint.com/c_standard_library/c_function_toupper.htm)

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

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