

Trabajos de Sintaxis y Semántica de los Lenguajes

Esp. Ing. José María Sola, profesor.

Revisión 3.9.0

2019-08-05

Tabla de contenidos

1. Introducción	1
2. Requisitos Generales para las Entregas de las Resoluciones	3
2.1. Requisitos de Forma	3
2.1.1. Repositorios	3
2.1.2. Lenguaje de Programación	7
2.1.3. Header Comments (Comentarios Encabezado)	7
2.2. Requisitos de Tiempo	8
3. "Hello, World!" en C	9
3.1. Objetivos	9
3.2. Temas	9
3.3. Problema	9
3.4. Restricciones	9
3.5. Tareas	9
3.6. Productos	10
4. Fases de la Traducción y Errores	11
4.1. Objetivos	11
4.2. Temas	11
4.3. Tareas	11
4.3.1. Secuencia de Pasos	11
4.4. Restricciones	13
4.5. Productos	13
5. Interfaces & Makefile — Temperaturas	15
5.1. Objetivos	15
5.2. Temas	15
5.3. Tareas	16
5.4. Restricciones	16
5.5. Productos	16
6. Operaciones de Strings	19
6.1. Objetivos	19
6.2. Temas	20
6.3. Tareas	20
6.4. Restricciones	21
6.5. Productos	21
7. Strings en Go (<i>golang</i>)	23

7.1. Objetivos	23
7.2. Temas	24
7.3. Tareas	24
7.4. Restricciones	24
7.5. Productos	24
8. Máquinas de Estado — Palabras en Líneas	27
8.1. Objetivos	27
8.2. Temas	27
8.3. Tareas	27
8.4. Restricciones	29
8.5. Productos	29
9. Máquinas de Estado — Contador de Palabras	31
9.1. Objetivos	31
9.2. Temas	31
9.3. Tareas	31
9.4. Restricciones	33
9.5. Productos	33
10. Parser Simple	35
10.1. Objetivo	35
10.2. Temas	35
10.3. Tareas	36
10.4. Restricciones	36
10.5. Productos	36
11. Traductor de Declaraciones C a LN	39
11.1. Restricciones	39
12. Traductor de Declaraciones C a LN con Lex	41
13. Traductor de Declaraciones C a LN con Lex & Yacc	43
14. Trabajo #3 — Removedor de Comentarios	45
14.1. Objetivo	45
14.2. Restricciones	45
14.3. Productos	46
14.4. Entrega	47
15. Trabajo #4 — Módulo Stack (?)	49
15.1. Objetivos	49
15.2. Temas	49
15.3. Tareas	50

15.4. Restricciones	51
15.5. Productos	51
15.6. Entrega	51
16. Trabajo #5 — Léxico de la Calculadora Polaca (@)	53
16.1. Objetivos	53
16.2. Temas	53
16.3. Tareas	54
16.4. Restricciones	55
16.5. Productos	56
16.6. Entrega	57
17. Trabajo #7 — Calculadora Polaca con Lex (@)	59
17.1. Objetivo	59
17.2. Restricciones	59
17.3. Productos	59
17.4. Entrega	59
18. Trabajo #8 — Calculadora Infija con RDP (?)	61
18.1. Objetivo	61
18.2. Restricciones	61
18.3. Entrega	62
19. Trabajo #9 — Calculadora Infija con Yacc (?)	63
Bibliografía	65

Lista de ejemplos

2.1. Nombre de carpeta	5
2.2. Header comments	8

1

Introducción

El objetivo de los trabajos es afianzar los conocimientos y evaluar su comprensión.

En la [sección "Trabajos" de la página del curso](#)¹ se indican cuales de los trabajos acá definidos que son **obligatorios** y cuales **opcionales**, como así también si se deben resolver **individualmente** o en **equipo**.

En el [sección "Calendario" de la página del curso](#)² se establece cuando es la **fecha y hora límite de entrega**,

Hay trabajos opcionales que son introducción a otros trabajos más complejos, también pueden enviar la resolución para que sea evaluada.

Cada trabajo tiene un **número** y un **nombre**, y su enunciado tiene las siguientes secciones:

1. **Objetivos:** Descripción general de los objetivos y requisitos del trabajo.
2. **Temas:** Temas que aborda el trabajo.
3. **Problema:** *Descripción* del problema a resolver, la *definición completa y sin ambigüedades* es parte del trabajo.
4. **Tareas:** Plan de tareas a realizar.
5. **Restricciones:** Restricciones que deben cumplirse.
6. **Productos:** Productos que se deben entregar para la resolución del trabajo.

¹ <https://josemariasola.wordpress.com/ssl/assignments/>

² <https://josemariasola.wordpress.com/ssl/calendar/>

2

Requisitos Generales para las Entregas de las Resoluciones

Cada trabajo tiene sus requisitos particulares de entrega de resoluciones, esta sección indica los requisitos generales, mientras que, cada trabajo define sus requisitos particulares.

Una resolución se considera **entregada** cuando cumple con los **requisitos de tiempo y forma** generales, acá descritos, sumados a los particulares definidos en el enunciado de cada trabajo.

La entrega de cada resolución debe realizarse a través de *GitHub*, por eso, cada estudiante tiene poseer una cuenta en esta plataforma.

2.1. Requisitos de Forma

2.1.1. Repositorios

En el curso usamos repositorios *GitHub*. Uno público y personal y otro privado para del equipo.

Repositorios público y privado.

```
Usuario
`-- Repositorio público personal para la asignatura
   Repositorio privado del equipo
```

Repositorio Personal para Trabajos Individuales

Cada estudiante debe crear un repositorio público dónde publicar las resoluciones de los trabajos individuales. El nombre del repositorio debe ser el de la asignatura. En la raíz del mismo debe publicarse un archivo `readme.md` que actúe como *front page* de la persona. El mismo debe estar escrito en notación *Markdown* y debe contener, como mínimo, la siguiente información:

- Sintaxis y Semántica de los Lenguajes
- Curso.
- Año de cursada, y cuatrimestre si corresponde.
- Legajo.
- Apellido.
- Nombre.

Repositorio personal para la asignatura.

```
Usuario
`-- Repositorio público personal para la asignatura
    |-- readme.md // Front page del usuario
```

Repositorio de Equipo para Trabajos Grupales

A cada equipo se le asigna un **repositorio privado**. En la raíz del mismo debe publicarse un archivo `readme.md` que actúe como *front page* del equipo. El mismo debe estar escrito en notación *Markdown* y debe contener, como mínimo, la siguiente información:

- Sintaxis y Semántica de los Lenguajes
- Curso.
- Año de cursada, y cuatrimestre si corresponde.
- Número de equipo.
- Nombre del equipo (opcional).
- Integrantes del equipo actualizados, ya que, durante el transcurso de la cursada el equipo puede cambiar:

- Usuario *GitHub*.
- Legajo.
- Apellido.
- Nombre.

Repositorio privado del equipo.

```
Repositorio privado del equipo
`-- readme.md // Front page del equipo.
```

Carpetas para cada Resolución

La resolución de cada trabajo debe tener su propia carpeta, ya sea en el repositorio personal, si es un trabajo individual, o en el del equipo, si es un trabajo grupal. El nombre de la carpeta debe seguir el siguiente formato:

DosDígitosNúmeroTrabajo-NombreTrabajo

O en notación *regex*:

```
[0-9]{2}"-"[a-zA-Z]+
```

Ejemplo 2.1. Nombre de carpeta

00-Hello

En los enunciados de cada trabajo, el número de trabajo para utilizar en el nombre de la carpeta está generalizado con "DD", se debe reemplazar por los dos dígitos del trabajo establecidos en el curso.

Adicionalmente a los productos solicitados para la resolución de cada trabajo, la carpeta debe incluir su propio archivo `readme.md` que actúe como *front page* de la resolución. El mismo debe estar escrito en notación *Markdown* y debe contener, como mínimo, la siguiente información:

- Número de equipo.
- Nombre del equipo (opcional).
- Autores de la resolución:
 - Usuario github.
 - Legajo.
 - Apellido.
 - Nombre.
- Número y título del trabajo.
- Transcripción del enunciado.
- Hipótesis de trabajo que surgen luego de leer el enunciado.

Opcionalmente, para facilitar el desarrollo se **recomienda incluir**:

- un archivo `.gitignore`.
- un archivo `Makefile`.¹
- archivos `tests`.¹

Carpeta de resolución de trabajo.

```
Carpeta de resolución de trabajo
|-- .gitignore
|-- Makefile
|-- readme.md // Front page de la resolución
`-- Archivos de resolución
```

Por último, la carpeta **no debe incluir**:

- archivos ejecutables.
- archivos intermedios producto del proceso de compilación o similar.

Ejemplo de Estructura de Repositorios

Ejemplo completo.

¹ Para algunos trabajos, el archivo `Makefile` y los `tests` son obligatorios, de ser así, se indica en el enunciado del trabajo.

```
usuario // Usuario GitHub
`-- Asignatura // Repositorio personal público para a la asignatura
    |-- readme.md // Front page del usuario
    |-- 00-Hello // Carpeta de resolución de trabajo
    |   |-- .gitignore
    |   |-- readme.md // Front page de la resolución
    |   |-- Makefile
    |   |-- hello.cpp
    |   |-- output.txt
    `-- 01-Otro-trabajo
2019-051-02 // Repositorio privado del equipo
|-- readme.md // Front page del equipo
|-- 04-Stack // Carpeta de resolución de trabajo
|   |-- .gitignore
|   |-- readme.md // Front page de la resolución
|   |-- Makefile
|   |-- StackTest.cpp
|   |-- Stack.h
|   |-- Stack.cpp
|   |-- StackApp.cpp
|-- 01-Otro-trabajo
```

2.1.2. Lenguaje de Programación

En el curso se establece la versión del estándar del lenguaje de programación que debe utilizarse en la resolución.

2.1.3. Header Comments (Comentarios Encabezado)

Todo archivo fuente debe comenzar con un comentario que indique el "Qué", "Quiénes", "Cuándo" :

```
/* Qué: Nombre
 * Breve descripción
 * Quiénes: Autores
 * Cuando: Fecha de última modificación
 */
```

Ejemplo 2.2. Header comments

```
/* Stack.h
 * Interface for a stack of ints
 * JMS
 * 20150920
 */
```

2.2. Requisitos de Tiempo

Cada trabajo tiene una **fecha y hora límite de entrega**, los *commits* realizados luego de ese instante no son tomados en cuenta para la evaluación de la resolución del trabajo.

En el [calendario del curso](https://josemariasola.wordpress.com/ssl/calendar/)² se publican cuando es la fecha y hora límite de entrega de cada trabajo.

² <https://josemariasola.wordpress.com/ssl/calendar/>

3

"Hello, World!" en C

3.1. Objetivos

- Demostrar con, un programa simple, que se está en capacidad de editar, compilar, y ejecutar un programa C.
- Contar con las herramientas necesarias para abordar la resolución de los trabajos posteriores.

3.2. Temas

- Sistema de control de versiones.
- Lenguaje de programación C.
- Proceso de compilación.
- Pruebas.

3.3. Problema

Adquirir y preparar los recursos necesarios para resolver los trabajos del curso.

3.4. Restricciones

- Ninguna.

3.5. Tareas

1. Solicitar inscripción al Grupo Yahoo, la aprobación demora un par de días.
2. Si no posee una cuenta *GitHub*, crearla.

3. Crear un repositorio público llamado `ssl`.
4. Escribir el archivo `readme.md` que actúa como *front page* del repositorio personal.
5. Crear la carpeta `00-chelloworld`.
6. Escribir el archivo `readme.md` que actúa como *front page* de la resolución.
7. Seleccionar, instalar, y configurar un compilador **C11** ó **C18**.
8. Indicar en `readme.md` el compilador seleccionado.
9. Probar el compilador con un programa `hello.c` que envíe a `stdout` la línea `hello, world!` o similar.
- 10 Ejecutar el programa, y capturar su salida en un archivo `output.txt`.
11. Publicar en el repositorio personal `ssl` la carpeta `00-chelloworld` con `readme.md`, `hello.c`, y `output.txt`.
- 12 La última tarea es informar por email a UTNFRBASSL@yahoogroups.com¹ el usuario *GitHub*.

3.6. Productos

```

Usuario
|-- ssl
    |-- 00-chelloworld
        |-- readme.md
        |-- hello.c
        |-- output.txt

```

¹ <mailto:UTNFRBASSL@yahoogroups.com>

Fases de la Traducción y Errores

4.1. Objetivos

- Identificar las fases de traducción y errores.

4.2. Temas

- Fases de traducción.
- Preprocesamiento.
- Compilación.
- Ensamblado.
- Vinculación (Link).
- Errores en cada fase.

4.3. Tareas

1. Investigar las funcionalidades y opciones que su compilador presenta para limitar el inicio y fin de las fases de traducción.
2. Para la siguiente secuencia de pasos:
 - a. Transcribir en `readme.md` cada **comando ejecutado** y
 - b. Describir en `readme.md` el **resultado** u **error** obtenidos para cada paso.

4.3.1. Secuencia de Pasos

1. Escribir `hello2.c`, que es una variante de `hello.c`:

```
#include <stdio.h>

int/*medio*/main(void){
    int i=42;
    printf("La respuesta es %d\n");
```

2. Preprocesar hello2.c, no compilar, y generar hello2.i. Analizar su contenido.
3. Escribir hello3.c, una nueva variante:

```
int printf(const char *s, ...);

int main(void){
    int i=42;
    printf("La respuesta es %d\n");
```

4. Investigar la semántica de la primera línea.
5. Preprocesar hello3.c, no compilar, y generar hello3.i. Buscar diferencias entre hello3.c y hello3.i.
6. Compilar el resultado y generar hello3.s, no ensamblar.
7. Corregir en el nuevo archivo hello4.c y empezar de nuevo, generar hello4.s, no ensamblar.
8. Investigar hello4.s.
9. Ensamblar hello4.s en hello4.o, no vincular.
- 10.Vincular hello4.o con la biblioteca estándar y generar el ejecutable.
- 11.Corregir en hello5.c y generar el ejecutable.
- 12Ejecutar y analizar el resultado.
- 13Corregir en hello6.c y empezar de nuevo.
- 14Escribir hello7.c, una nueva variante:

```
int main(void){
    int i=42;
    printf("La respuesta es %d\n", i);
}
```

15 Explicar porqué funciona.

4.4. Restricciones

- El programa ejemplo debe enviar por `stdout` la frase `La respuesta es 42`, el valor 42 debe surgir de una variable.

4.5. Productos

```
DD-FasesErrores
|-- readme.md
|-- hello2.c
|-- hello3.c
|-- hello4.c
|-- hello5.c
|-- hello6.c
`-- hello7.c
```

5

Interfaces & Makefile — Temperaturas

Este trabajo está basado en los ejercicios 1-4 y 1-15 de [\[KR1988\]](#):

1-4. Escriba un programa para imprimir la tabla correspondiente de Celsius a Fahrenheit

1-15. Reescriba el programa de conversión de temperatura de la sección 1.2 para que use una función de conversión.

Desarrollar un programa que imprima dos tablas de conversión, una de Fahrenheit a Celsius y otra de Celsius a Fahrenheit.

5.1. Objetivos

- Realizar el primer trabajo en equipo en el repositorio privado del equipo en **GitHub**.
- Aplicar el uso de interfaces y de `makefile`.

5.2. Temas

- `makefile`.
- Archivos header (`.h`).

- Tipo de dato `double`.
- Funciones.
- Pruebas unitarias.
- `assert`.



La comparación de los tipos flotantes puede ser no trivial debido a su representación y precisión.

- Interfaces e Implementación.

5.3. Tareas

1. Escribir `makefile`.
2. Escribir `conversion.h`
3. Escribir `conversionTest.h`
4. Escribir `conversion.c`
5. Escribir `tablasDeConversion.c`.

5.4. Restricciones

- Las funciones deben llamarse `celsius` y `fahrenheit`.
- Utilizar `assert`.
- Utilizar `const`.
- Utilizar `for` con declaración (C99).

5.5. Productos

```
DD-Interfaces
|-- readme.md
|-- Makefile
|-- conversion.h
|-- conversionTest.c
|-- conversion.c
`-- tablasDeConversion.c.
```




Crédito extra

Desarrolle `TablasDeConversion.c` para que use funciones del estilo `PrintTablas`, `PrintTablaCelsius`, `PrintTablaFahrenheit`, `PrintFilas`, `PrintFila`.



Crédito extra

Desarrollar la función `PrintFilas` para que sea genérica, es decir, pueda invocarse desde `PrintTablaFahrenheit` y desde `PrintTablaCelsius`. `PrintFilas` debe invocar a `PrintFila`.

6

Operaciones de Strings

Este trabajo tiene dos partes, una de análisis comparativo y otra de desarrollo.

El análisis comparativo es sobre el tipo de dato String en el lenguaje de programación C versus otro lenguaje de programación a elección; mientras que el desarrollo está basado en los ejercicios 20 y 21 del capítulo 1 del volumen 1 de [MUCH2012], que a continuación transcribe:

Investigue y construya, en LENGUAJE C, la función que realiza cada operación solicitada:

* Ejercicio 20 *

- (a) Calcula la longitud de una cadena;
- (b) Determina si una cadena dada es vacía.
- (c) Concatena dos cadenas.

* Ejercicio 20 *

Construya un programa de testeo para cada función del ejercicio anterior.

6.1. Objetivos

1. Parte I — Análisis Comparativo del tipo String en Lenguajes de Programación: Realizar un análisis comparativo de dato String en el lenguaje C versus un lenguaje de programación a elección. El análisis debe contener, por lo menos, los siguientes ítems:

- a. ¿El tipo es parte del lenguaje en algún nivel?

- b. ¿El tipo es parte de la biblioteca?
- c. ¿Qué alfabeto usa?
- d. ¿Cómo se resuelve la alocaación de memoria?
- e. ¿El tipo tiene mutabilidad o es inmutable?
- f. ¿El tipo es un *first class citizen*?
- g. ¿Cuál es la mecánica para ese tipo cuando se los pasa como argumentos?
- h. ¿Y cuando son retornados por una función?

Las anteriores preguntas son disparadores para realizar una análisis profundo.

2. Parte II — Biblioteca para el Tipo String: Desarrollar una biblioteca con las siguientes operaciones de strings:

- a. *GetLength* ó *GetLongitud*
- b. *IsEmpty* ó *IsVacía*
- c. *Power* ó *Potenciar*
- d. Una operación a definir libremente.

Notar que en vez de la operación concatenar que propone [MUCH2012] se debe desarrollar *Power* ó *Potenciar* que repite un string n veces.

La parte pública de la biblioteca se desarrolla en el header "string.h", el cual no debe incluir <string.h>. El programa que prueba la biblioteca, por supuesto, incluye a "string.h", pero sí puede incluir <string.h> para facilitar las comparaciones.

6.2. Temas

- Strings.
- Alocaación.
- Tipos.

6.3. Tareas

1. Parte I

- a. Escribir el `AnálisisComparativo.md` con la comparación de strings en C versus otro lenguaje de programación a elección.

2. Parte II

- a. Para cada operación, escribir en `strings.md` la especificación matemática de la operación, con conjuntos de salida y de llegada, y con especificación de la operación.
- b. Escribir el `makefile`.
- c. Por cada operación:
 - i. Escribir las pruebas en `stringsTest.c`.
 - ii. Escribir los prototipos en `string.h`.
 - iii. Escribir en `String.h` comentarios con las precondiciones y poscondiciones de cada función, arriba de cada prototipo.
 - iv. Escribir las implementaciones en `strings.c`.

6.4. Restricciones

- Las pruebas deben utilizar `assert`.
- Los proptotipos de utilizar `const` cuando corresponde.
- Por lo menos una operación debe implementarse con recursividad.
- Las implementaciones no deben utilizar funciones estándar, declaradas en `<string.h>`

6.5. Productos

```
DD-Strings
|-- readme.md
|-- AnálisisComparativo.md
|-- String.md
|-- Makefile
|-- StringTest.c
|-- String.h
`-- String.c.
```


Strings en Go (*golang*)

El trabajo consta de la especificación e implementación del tipo string según Go, con representación en memoria igual a la de ese lenguaje, con la facilidad de aloca strings en el heap, y para liberar esa memoria reservada ante el pedido del *garbage collector*.

Las operaciones del tipo son:

- Len
- Count
- New
- At
- Delete

Recordar que el tipo string es inmutable en Go.

El trabajo incluye también un ejemplo el uso del tipo en un programa que haga uso de las de las operaciones y el desarrollo de una función que reciba un string como parámetro.

7.1. Objetivos

1. Especificación del tipo String de Go.
2. Programa ejemplo en Go.
3. Desarrollo del tipo String de Go en C.
4. Programa ejemplo en C que usa el tipo String de Go.

7.2. Temas

- Strings.
- Alocación.
- Tipos.
- Heap
- Garbage Collector.

7.3. Tareas

- a. Especificar el tipo en `goString.md`.
- b. Escribir y ejecutar un programa Go ejemplo de uso con una función que recibe un string en `goStringExample.go`.
- c. Escribir el `makefile`.
- d. Escribir las pruebas en `goStringTest.c`.
- e. Escribir las declaraciones públicas en `goString.h`.
- f. Escribir en `goString.h` comentarios con las precondiciones, poscondiciones e invariantes.
- g. Escribir un ejemplo de uso de tipo, que incluya una función que recibe un string en `goStringExample.c`.
- h. Escribir la implementación en `goString.c`.

7.4. Restricciones

- Las pruebas deben utilizar `assert`.
- Los prototipos de utilizar `const` cuando corresponde.
- La operación `at` debe implementar el mismo compartamiento `panic` que tienen Go; para eso debe desarrollarse la función `panic` que es invocada por `at` cuando el índice es inválido.

7.5. Productos

```
DD-GoStrings
|-- readme.md
```



```
|-- GoString.md  
|-- GoStringExample.go  
|-- Makefile  
|-- GoStringTest.c  
|-- GoString.h  
|-- GoStringExample.c  
`-- GoString.c
```

8

Máquinas de Estado — Palabras en Líneas

Este trabajo está basado en el ejercicio 1-12 de [\[KR1988\]](#):

1-12. Escriba un programa que imprima su entrada una palabra por línea.

Desarrollar un programa que imprima cada palabra de la entrada en su propia línea. La cantidad de líneas en la salida coincide con la cantidad de palabras en la entrada. Cada línea tiene solo una palabra.

8.1. Objetivos

- Aplicar máquinas de estado para el procesamiento de texto.

8.2. Temas

- Árboles de expresión.
- Representación de máquinas de estado.
- Implementación de máquinas de estado.

8.3. Tareas

1. Árboles de Expresión

- a. Estudiar el programa del ejemplo la sección *1.5.4 Conteo de Palabras* de [\[KR1988\]](#).

- b. Dibujar el árbol de expresión para la inicialización de los contadores: `n1 = nw = nc = 0`.
 - c. Dibujar el árbol de expresión para la expresión de control del segundo if: `c == ' ' || c == '\n' || c == '\t'`.
2. Máquina de Estado:
- a. Describir en lenguaje dot [\[DOT2015\]](#) y dentro del archivo `p1.gv` la máquina de estado que resuelve el problema planteado.
 - b. Formalizar la máquina de estados como una *n-upla*.
3. Implementación #2: Sentencias goto (sí, el infame *goto*)
- a. Leer la sección 3.8 *Goto and labels* de [\[KR1988\]](#)
 - b. Leer *Go To Statement Considered Harmful* de [\[DIJ1968\]](#).
 - c. Leer *"GOTO Considered Harmful" Considered Harmful* de [\[RUB1987\]](#).
 - d. Responder: ¿Tiene alguna aplicación *go to* hoy en día? ¿Algún lenguaje moderno lo utiliza?
 - e. Escribir una segunda versión del programa, `p1-goto.c`, que, en vez de variable o funciones, utilice etiquetas para representar los estados y sentencias *goto* para las transiciones.
4. Implementación #3: Funciones Recursivas
- a. Leer la sección 4.10 *Recursividad* de [\[KR1988\]](#).
 - b. Escribir una tercera versión del programa, `p1-rec.c`, que, en vez de una variable, utilice funciones recursivas para representar los estados.
5. Implementación #X:
- a. Diseñar un modelo de implementación **diferente** a las implementaciones #1, #2, y #3.
 - b. Responder sobre el nuevo modelo: ¿Cómo implementa los estados este programa? ¿Y las transiciones?
 - c. Escribir una cuarta versión del programa, `p1-x.c`, que, represente los estados y las transiciones tal como se responda en la pregunta anterior.

8.4. Restricciones

- Ninguna.

8.5. Productos

```
DD-Palabras-En-Líneas
|-- readme.md
|   |-- Árboles de expresión.
|   `-- Respuestas.
|-- pl.gv
|-- Makefile
|-- pl-goto.c
|-- pl-rec.c
`-- pl-x.c
```

9

Máquinas de Estado — Contador de Palabras

Este trabajo está basado en el ejemplo de la sección *1.5.4 Conteo de Palabras* de [\[KR1988\]](#):

"... cuenta líneas, palabras, y caracteres, con la definición ligera que una palabra es cualquier secuencia de caracteres que no contienen un blanco, tabulado o nueva línea."

9.1. Objetivos

- Aplicar máquinas de estado para el procesamiento de texto.

9.2. Temas

- Árboles de expresión.
- Representación de máquinas de estado.
- Implementación de máquinas de estado.

9.3. Tareas

1. Árboles de Expresión

- a. Estudiar el programa del ejemplo la sección *1.5.4 Conteo de Palabras* de [\[KR1988\]](#).

- b. Dibujar el árbol de expresión para la inicialización de los contadores: `n1 = nw = nc = 0`.
 - c. Dibujar el árbol de expresión para la expresión de control del segundo `if`:
`c == ' ' || c == '\n' || c == '\t'`.
 - d. Máquina de Estado:
 - i. Describir en lenguaje dot [\[DOT2015\]](#) y dentro del archivo `wc.gv` la máquina de estado que implementa el programa ejemplo.
 - ii. Formalizar la máquina de estados como una *n-upla*.
2. Implementación #1: `enum` y `switch`
- a. Escribir una segunda versión del programa, `wc-enum-switch.c`, que:
 - i. Utilizar `typedef` y `enum` en vez de `define`, de tal modo que la variable estado se pueda declarar de la siguiente manera: `state s = out;`
 - ii. Utilizar `switch` en vez de `if`.
 - b. Responder: ¿Cómo implementa los estados este programa? ¿Y las transiciones?
3. Implementación #2: Sentencias `goto` (sí, el infame *goto*)
- a. Leer la sección 3.8 *Goto and labels* de [\[KR1988\]](#)
 - b. Leer *Go To Statement Considered Harmful* de [\[DIJ1968\]](#).
 - c. Leer *"GOTO Considered Harmful" Considered Harmful* de [\[RUB1987\]](#).
 - d. Responder: ¿Tiene alguna aplicación *go to* hoy en día? ¿Algún lenguaje moderno lo utiliza?
 - e. Escribir una tercera versión del programa, `wc-goto.c`, que, en vez de variable o funciones, utilice etiquetas para representar los estados y sentencias `goto` para las transiciones.

4. Implementación #3: Funciones Recursivas

- a. Leer la sección 4.10 *Recursividad* de [KR1988].
- b. Responder: ¿Cómo pueden las funciones actualizar los contadores?
- c. Leer la sección 1.10 *Variables Externas y Alcance* y 4.3 *Variables Externas* de [KR1988].
- d. Escribir una cuarta versión del programa, `wc-rec.c`, que, en vez de una variable, utilice funciones recursivas para representar los estados.

5. Eficiencia

- a. Construir una tabla comparativa a modo de *benchmark* que muestre el tiempo de procesamiento para cada una de las tres implementaciones, para tres archivos diferentes de tamaños diferentes, el primero en el orden de los kilobytes, el segundo en el orden de los megabytes, y el tercero en el orden de los gigabytes.

9.4. Restricciones

- Ninguna.

9.5. Productos

```
DD-Contador-de-Palabras
|-- readme.md
|   |-- Árboles de expresión.
|   |-- Respuestas.
|   |-- Tabla de mediciones.
|-- wc.gv
|-- Makefile
|-- wc-enum-switch.c
|-- wc-goto.c
|-- wc-rec.c
```

10

Parser Simple

Este trabajo está basado en el ejercicio 1-24 de [\[KR1988\]](#):

Escriba un programa para verificar errores sintácticos rudimentarios de un programa C, como paréntesis, corchetes, y llaves sin par. No se olvide de las comillas, apóstrofes, secuencias de escape, y comentarios. (Este programa es difícil si lo hace en su completa generalidad.)

10.1. Objetivo

El objetivo es diseñar e implementar un autómata de pila (APD) que verifique el balanceo de los paréntesis, corchetes, y llaves; en un programa C pueden estar anidados. La solución debe validar:

- Paréntesis, corchetes y llaves desbalanceados:
 - Válido: `{[()]}`
 - Inválido: `{[]}()`
- Apóstrofes y comillas, secuencias de escape:
 - Válido: `"[{"`
 - Inválido: `"{}`

10.2. Temas

- Autómata de Pila (Push down Automata).

- Stacks.

10.3. Tareas

- Primero diseñar y el APD y luego derivar una implementación.

10.4. Restricciones

- Primero diseñar y el APD y luego derivar una implementación.
- Utilizar el lenguaje dot para dibujar el digrafo.
- Resolver con máquina de estados, para eso leer Capítulo #2 del Volumen #2 de [\[MUCH2012\]](#).
- Utilizar el símbolo \$ para la pila vacío.
- Considerar las variantes no comunes de literales carácter y de literales cadenas que son parte del estándar de C.
- Diseñar `pushString("xyz")` para que sea equivalente a `push('z')`, `push('y')`, `push('x')`
- Para la implementación indicar cómo se representan los estados y cómo las transiciones.
- Respetar la máquina de estado especificada, en la implementación utilizar los mismos nombres de estado y cantidad de transiciones.
- En el caso que sea necesario, utilizar enum, y no define.
- En el caso que sea necesario, utilizar switch, y no if.
- Crear a mano el archivo de test funcional: `Test.c`.
- Buscar la mejor forma para reutilizar la implemenación de *stack* de problemas anteriores.



Crédito extra

Diseñar el programa para que pueda invocarse de la siguiente manera: `> RemoveComments < Test.c | Parse`

10.5. Productos

- Sufijo del nombre de la carpeta: `SimpleParser`.

- /Readme.md
 - Definición formal del APD.
 - Especificación de *PushString* basada en operaciones de cadenas de lenguajes formales.
- /StackOfCharsModule.h
- /StackOfCharsModule.c
- /Parser.gv
- /Parser.c
- /Makefile

Traductor de Declaraciones C a LN

Este trabajo está basado en el programa `dc1` ejemplo de la sección 5.12 *Declaraciones Complicadas* de [\[KR1988\]](#).

11.1. Restricciones

- Aplicar los conceptos de modularización, componentes, e interfaces.
- Codificar `scanner.h` y `scanner.c`, para declarar y definir las siguientes declaraciones:

```
enum TokenType {
    ...,
    LexError
};
typedef enum TokenType TokenType;
typedef ... TokenValue;
struct Token{
    TokenType type;
    TokenValue val;
};
bool GetNextToken(Token *t /*out*/); // Retorna si pudo leer,
    almacena en t el token leído.
```

12

Traductor de Declaraciones C a LN con Lex

Este trabajo está basado en el programa `dc1` ejemplo de la sección 5.12 *Declaraciones Complicadas* de [\[KR1988\]](#).

Esta versión se basa en un scanner generado por Lex.

13

Traductor de Declaraciones C a LN con Lex & Yacc

Este trabajo está basado en el programa `dc1` ejemplo de la sección 5.12 *Declaraciones Complicadas* de [\[KR1988\]](#).

Esta versión se basa en un scanner generado por Lex y un parser generado por Yacc.

14

Trabajo #3 — Removedor de Comentarios

Este trabajo está basado en el ejercicio 1-23 de [\[KR1988\]](#):

Escriba un programa que remueva todos los comentarios de un programa C. Los comentarios en C no se anidan. No se olvide de tratar correctamente las cadenas y los caracteres literales

14.1. Objetivo

El objetivo es diseñar una máquina de estado que remueva comentarios, implementar dos versiones, e informar cual es la más eficiente mediante un benchmark.

14.2. Restricciones

- Primero diseñar y especificar la máquina de estado y luego derivar dos implementaciones.
- Utilizar el lenguaje `dot` para dibujar los digrafos.
- Incluir comentarios de una sola línea (`//`).
- Considerar las variantes no comunes de literales carácter y de literales cadenas que son parte del estándar de C.
- Diseñar el programa para que pueda invocarse de la siguiente manera:
`RemoveComments < Test.c > NoComments.c`

- Ninguna de las implementaciones debe ser la *Implementación #1: estado como variable y transiciones con selección estructurada*.
- Indicar para cada implementación cómo se representan los estados y cómo las transiciones.
- Respetar la máquina de estado especificada, en cada implementación utilizar los mismos nombres de estado y cantidad de transiciones.
- En el caso que sea necesario, utilizar `enum`, y no `define`.
- En el caso que sea necesario, utilizar `switch`, y no `if`.
- Realizar una prueba funcional y tres pruebas de volumen.
- Construir una tabla comparativa a modo de *benchmark* que muestre el tiempo de procesamiento para cada una de las dos implementaciones, para tres archivos diferentes de tamaños diferentes, el primero en el orden de los kilobytes, el segundo en el orden de los megabytes, y el tercero en el orden de los gigabytes.
- Crear a mano el archivo de test funcional: `Test.c`.
- Construir el programa `GenerateTest.c` que genere automáticamente los tres archivos para pruebas de volumen: `Testkilo.c`, `Testmega.c`, y `Testgiga.c`.
- No incorporar al repositorio los archivos de prueba de volumen, sí el de prueba funcional.
- Diseñar el archivo `Makefile` para que construya una, otra o ambas implementaciones, y para que ejecute las pruebas.

14.3. Productos

- Sufijo del nombre de la carpeta: `sincomentarios`.
- `/Readme.md`
 - Autómata finito para cada lenguaje.
 - Diagrama de transiciones.
 - Definición Formal.
 - Expresión regular para cada lenguaje.
 - Máquina de Estados del programa.

- Descripción de la implementación A: rc-a.c.
- Descripción de la implementación B: rc-a.c.
- Benchmark.
- /rc-a.c
- /rc-b.c
- /tests/Test.c
- /tests/GenerateTest.c
- /Makefile

14.4. Entrega

- Jun 5, 13hs

15

Trabajo #4 — Módulo Stack (?)

15.1. Objetivos

Construir dos implementaciones del Módulo Stack de `int`s`.

15.2. Temas

- Módulos.
- Interfaz.
- Stack.
- Unit tests.
- `assert`
- Reserva estática de memoria.
- Ocultamiento de información.
- Encapsulamiento.
- Precondiciones.
- Poscondiciones.
- Call stack.
- heap.
- Reserva dinámica de memoria.
- Punteros.
- `malloc`.
- `free`.

15.3. Tareas

1. Analizar el stack de la sección 4.3 de [\[KR1988\]](#).
2. Codificar la interfaz `stackModule.h` para que incluya las operaciones:
 - a. Push.
 - b. Pop.
 - c. IsEmpty.
 - d. IsFull.
3. Escribir en la interfaz `stackModule.h` comentarios que incluya *especificaciones* y *pre* y *poscondiciones* de las operaciones.
4. Codificar los unit tests en `stackModuleTest.c`.
5. Codificar una implementación contigua y estática en `stackModuleContiguousStatic.c`.
6. Probar `stackModuleContiguousStatic.c` con `stackModuleTest.c`.
7. Codificar una implementación enlazada y dinámica en `stackModuleLinkedDynamic.c`.
8. Probar `stackModuleLinkedDynamic.c` con `stackModuleTest.c`.
9. Probar `StackDynamic.c` con `StackTest`.
10. Construir una tabla comparativa a modo de *benchmark* que muestre el tiempo de procesamiento para cada una de las dos implementaciones.
11. Diseñar el archivo `makefile` para que construya una, otra o ambas implementaciones, y para que ejecute las pruebas.
12. Responder:
 - a. ¿Cuál es la mejor implementación? Justifique.
 - b. ¿Qué cambios haría para que no haya precondiciones? ¿Qué implicancia tiene el cambio?
 - c. ¿Qué cambios haría en el diseño para que el stack sea genérico, es decir permita elementos de otros tipos que no sean `int`? ¿Qué implicancia tiene el cambio?
 - d. Proponga un nuevo diseño para que el módulo pase a ser un *tipo de dato*, es decir, permita a un programa utilizar más de un stack.

15.4. Restricciones

- En `stackModule.h`:
 - Aplicar guardas de inclusión.
 - Declarar `typedef int StackItem;`
- En `stackModuleTest.c` incluir `assert.h` y aplicar `assert`.
- En ambas implementaciones utilizar `static` para aplicar encapsulamiento.
- En la implementación contigua y estática:
 - No utilizar índices, sí aritmética punteros.
 - Aplicar el *idiom* para stacks.
- En la implementación enlazada y dinámica:
 - Invocar a `malloc` y a `free`.
 - No utilizar el operador `sizeof(tipo)`, sí `sizeof expresión`.

15.5. Productos

- Sufijo del nombre de la carpeta: `stackModule`.
- `/Readme.md`
 - Benchmark.
 - Preguntas y Respuestas.
- `/StackModule.h`.
- `/StackModuleTest.c`
- `/StackModuleContiguousStatic.c`
- `/StackModuleLinkedDynamic.c`
- `/Makefile`

15.6. Entrega

Opcional.

16

Trabajo #5 — Léxico de la Calculadora Polaca (@)

Este trabajo está basado en el la sección 4.3 de [\[KR1988\]](#): *Calculadora con notación polaca inversa*.

16.1. Objetivos

- Estudiar los fundamentos de los scanner aplicados a una calculadora con notación polaca inversa que utiliza un stack.
- Implementar modularización mediante los módulos `calculator`, `stackOfDoublesModule`, y `Scanner`.

16.2. Temas

- Módulos.
- Interfaz.
- Stack.
- Ocultamiento de información.
- Encapsulamiento.
- Análisis léxico.
- Lexema.
- Token.
- Scanner.
- `enum`.

16.3. Tareas

1. Estudiar la implementación de la sección 4.3 de [\[KR1988\]](#).
2. Construir los siguientes componentes, con las siguientes entidades públicas:

Calculator	StackOfDoublesModule	Scanner
<ul style="list-style-type: none"> • Qué hace: Procesa entrada y muestra resultado. • Qué usa: <ul style="list-style-type: none"> ◦ Biblioteca Estándar <ul style="list-style-type: none"> ▪ EOF ▪ printf ▪ atof ◦ StackOfDoublesModule <ul style="list-style-type: none"> ▪ StackItem ▪ Push ▪ Pop ▪ IsEmpty ▪ IsFull ◦ Scanner <ul style="list-style-type: none"> ▪ GetNextToken ▪ Token ▪ TokenType ▪ TokenValue 	<ul style="list-style-type: none"> • Qué exporta: <ul style="list-style-type: none"> ◦ StackItem ◦ Push ◦ Pop ◦ IsEmpty ◦ IsFull 	<ul style="list-style-type: none"> • Qué hace: Obtiene operadores y operandos. • Qué usa: <ul style="list-style-type: none"> ◦ Biblioteca Estándar <ul style="list-style-type: none"> ▪ getchar ▪ EOF ▪ isdigit ▪ ungetc • Qué exporta: <ul style="list-style-type: none"> ◦ GetNextToken ◦ Token ◦ TokenType ◦ TokenValue

1. Diagramar en *Dot* las dependencias entre los componentes e interfaces.

2. Definir formalmente y con digrafo en *Dot* la máquina de estados que implementa `getNextToken`, utilizar estados finales para diferentes para cada clase de tokens.
3. Escribir un archivo `expresiones.txt` para probar la calculadora.
4. Construir el programa `calculator`.
5. Ejecutar `calculator < expresiones.txt`.
6. Responder:
 - a. ¿Es necesario modificar `stackModule.h`? ¿Por qué?
 - b. ¿Es necesario recompilar la implementación de `Stack`? ¿Por qué?
 - c. ¿Es necesario que `calculator` muestre el lexema que originó el error léxico? Justifique su decisión.
 - i. Si decide hacerlo, ¿de qué forma debería exponerse el lexema?
Algunas opciones:
 - Tercer componente `lexeme` en `Token` ¿De qué tipo de dato es aplicable?
 - Cambiar el tipo de `val` para que sea un `union` que pueda representar el valor para `Number` y valor `LexError`.
 - ii. Implemente la solución según su decisión.

16.4. Restricciones

- Aplicar los conceptos de modularización, componentes, e interfaces.
- En `calculator.c` la variable `token` del tipo `Token`, que es asignada por `getNextToken`.
- Codificar `stackOfDoublesModule.h` a partir de la implementación contigua y estática de `stackModule`, `stackModuleContiguousStatic.c`, del trabajo #4, y modificar `stackItem`.
- Codificar `scanner.h` y `scanner.c`, para que usen las siguientes declaraciones:

```
enum TokenType {  
    Number,  
    Addition='+',
```

```

    Multiplication='*',
    Subtraction='-',
    Division='/',
    PopResult='\n',
    LexError
};
typedef enum TokenType TokenType;
typedef double TokenValue;
struct Token{
    TokenType type;
    TokenValue val;
};
bool GetNextToken(Token *t /*out*/); // Retorna si pudo leer,
    almacena en t el token leído.

```

- GetNextToken debe usar una variable llamada lexeme para almacenar el lexema leído.
- Usar las siguientes entidades de la biblioteca estándar:
 - `stdio.h`
 - `getchar`
 - `EOF`
 - `stdin`
 - `printf`
 - `stdout`
 - `getchar`
 - `ungetc`
 - `ctype.h`
 - `isdigit`
 - `stdlib.h`
 - `atof`

16.5. Productos

- Sufijo del nombre de la carpeta: `PolCalc`.
- `/Readme.md`

- Preguntas y Respuestas.
- /expresiones.txt
- /Dependencias.gv
- /Calculator.c
- /StackOfDoublesModule.h
- /StackOfDoublesModule.c
- /Scanner.gv
- /Scanner.h
- /Scanner.c
- /Makefile

16.6. Entrega

- Jul 3, 13hs.
 - Preentrega:
 - stackofDoublesModule.h
 - stackofDoublesModule.c
 - Scanner.h
 - Scanner.gv
- Jul 31, 13hs
 - Entrega final completa.

17

Trabajo #7 — Calculadora Polaca con Lex (@)

Este trabajo está es una segunda iteración de [Capítulo 16, Trabajo #5 — Léxico de la Calculadora Polaca \(@\)](#), en la cual el *scanner* se implementa con *lex* y no con una máquina de estados.

17.1. Objetivo

Aplicar *lex* para el análisis lexicográfico.

17.2. Restricciones

- No cambiar `scanner.h`, implica recompilar solo `scanner.c` y volver a vincular.
- Utilizar *make* para construir el hacer uso de *lex*.
- La única diferencia está en `scanner.c`, en el cual la función `getNextToken` debe invocar a la función `yylex`.

17.3. Productos

- Sufijo del nombre de la carpeta: `Polca1Lex`.
- Los mismos que [Capítulo 16, Trabajo #5 — Léxico de la Calculadora Polaca \(@\)](#) con la adición de `/Scanner.1`

17.4. Entrega

- Sep 6, 13hs.

- Preentrega: `scanner.1` con `main` que informa por `stdout` los tokens encontrados en `stdin`
- Sep 11, 13hs
 - Entrega final completa.

18

Trabajo #8 — Calculadora Infija con RDP (?)

Este trabajo es la versión infija de [Capítulo 17, Trabajo #7 — Calculadora Polaca con Lex \(@\)](#); es decir en vez de procesar:

```
1 2 - 4 5 + *  
-9
```

el programa debe procesar correctamente:

```
(1 - 2) * (4 + 5)  
-9
```

18.1. Objetivo

- Diseñar una gramática independiente de contexto que represente la asociatividad y precedencia de las operaciones.
- Las operaciones son: + - * / ().
- Implementar un Parser Descendente Recursivo (RDP).

18.2. Restricciones

- Implementar getNextToken con Lex, basado en el getNextToken de [Capítulo 17, Trabajo #7 — Calculadora Polaca con Lex \(@\)](#)
- Agregar los tokens LParen y RParen.

18.3. Entrega

- Opcional

19

Trabajo #9 — Calculadora Infija con Yacc (?)

Esta vez, el parser lo construye Yacc por nosotros.

Bibliografía

- [DIJ1968] Edsger W. Dijkstra. *Go To Statement Considered Harmful*. Reprinted from Communications of the ACM, Vol. 11, No. 3, March 1968, pp. 147-148. <http://homepages.cwi.nl/~storm/teaching/reader/Dijkstra68.pdf>
- [RUB1987] Frank Rubin. *"Go To Statement Considered Harmful" Considered Harmful*. Reprinted from Communications of the ACM, Vol. 30, No. 3, March 1987, pp. 195-196. <http://web.archive.org/web/20090320002214/http://www.ecn.purdue.edu/ParaMount/papers/rubin87goto.pdf>
- [KR1988] Brian W. Kernighan and Dennis Ritchie. *The C Programming Language, 2nd Edition* 1988.
- [MUCH2012] Jorge Muchnik y Ana María Díaz Bott. *SSL, 2da Edición* 2012.
- [DOT2015] Emden R. Gansner and Eleftherios Koutsofios and Stephen North. *Drawing graphs with dot*. Retrived 2018-06-19 from <http://www.graphviz.org/pdf/dotguide.pdf>

