# Trabajos de Sintaxis y Semántica de los Lenguajes

Esp. Ing. José María Sola, profesor.

Revisión 3.12.0 2020-04-13

## 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. Uso del Lenguaje C en mi Día a Día	11
4.1. Objetivos	11
4.2. Temas	11
4.3. Tareas	11
4.4. Restricciones	11
4.5. Productos	11
5. Fases de la Traducción y Errores	13
5.1. Objetivos	13
5.2. Temas	13
5.3. Tareas	13
5.3.1. Secuencia de Pasos	13
5.4. Restricciones	15
5.5. Productos	15
6. Interfaces & Makefile — Temperaturas	17
6.1. Objetivos	17
6.2. Temas	17
6.3. Tareas	18
6.4. Restricciones	18
6.5. Productos	18
7. Operaciones de Strings	21

### Trabajos de Sintaxis y Semántica de los Lenguajes

7.1. Objetivos	21
7.2. Temas	22
7.3. Tareas	22
7.4. Restricciones	23
7.5. Productos	23
8. Strings en Go (golang)	25
8.1. Objetivos	25
8.2. Temas	26
8.3. Tareas	26
8.4. Restricciones	26
8.5. Productos	26
9. Máquinas de Estado — Palabras en Líneas	29
9.1. Objetivos	29
9.2. Temas	29
9.3. Tareas	29
9.4. Restricciones	32
9.5. Productos	32
10. Máquinas de Estado — Contador de Palabras	33
10.1. Objetivos	33
10.2. Temas	33
10.3. Tareas	34
10.4. Restricciones	36
10.5. Productos	36
11. Máquinas de Estado — Sin Comentarios	37
11.1. Objetivo	37
11.2. Restricciones	37
11.3. Productos	38
12. Parser Simple	39
12.1. Objetivo	39
12.2. Temas	39
12.3. Tareas	40
12.4. Restricciones	40
12.5. Productos	40
13. Traductor de Declaraciones C a LN	43
13.1. Restricciones	43
14. Traductor de Declaraciones C a LN con Lex	45

### Trabajos de Sintaxis y Semántica de los Lenguajes

15. Traductor de Declaraciones C a LN con Lex & Yacc	47
16. Trabajo #4 — Módulo Stack (?)	49
16.1. Objetivos	49
16.2. Temas	49
16.3. Tareas	50
16.4. Restricciones	51
16.5. Productos	51
16.6. Entrega	51
17. Trabajo #5 — Léxico de la Calculadora Polaca (@)	53
17.1. Objetivos	53
17.2. Temas	53
17.3. Tareas	54
17.4. Restricciones	55
17.5. Productos	56
17.6. Entrega	57
18. Trabajo #7 — Calculadora Polaca con Lex (@)	59
18.1. Objetivo	59
18.2. Restricciones	59
18.3. Productos	59
18.4. Entrega	59
19. Trabajo #8 — Calculadora Infija con RDP (?)	
19.1. Objetivo	61
19.2. Restricciones	
19.3. Entrega	62
20. 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 <sup>1</sup> 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<sup>2</sup> 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.

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

<sup>&</sup>lt;sup>2</sup> https://josemariasola.wordpress.com/ssl/calendar/

## 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 descriptos, 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.
- · Legaio.
- 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.
- archivos tests.<sup>1</sup>

#### 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.

<sup>&</sup>lt;sup>1</sup> Para algunos trabajos, el archivo Makefile y los tests son obligatorios, de ser así, se indica en el enunciado del trabajo.

```
usuario // Usuario GithHub
`-- Asignatura // Repositorio personal público para a la asignatura
    |-- readme.md // Front page del usuario
    |-- 00-Hello // Carperta de resolución de trabajo
         |-- .gitignore
         |-- readme.md // Front page de la resolución
         I-- Makefile
         |-- hello.cpp
         `-- output.txt
    `-- 01-Otro-trabaio
2019-051-02 // Repositorio privado del equipo
|-- redme.md // Front page del equipo
|-- 04-Stack // Carperta de resolución de trabajo
    |-- .gitignore
     |-- readme.md // Front page de la resolución
    l-- 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<sup>2</sup> se publican cuando es la fecha y hora límite de entrega de cada trabajo.

<sup>&</sup>lt;sup>2</sup> https://josemariasola.wordpress.com/ssl/calendar/

## "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 necesarias para resolver los trabajos del curso.

#### 3.4. Restricciones

· Ninguna.

#### 3.5. Tareas

- 1. Si no posee una cuenta GitHub, crearla.
- 2. Crear un repositorio público llamado SSL.

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

#### 3.6. Productos

```
Usuario
`-- SSL
`-- 00-CHelloworld
|-- readme.md
|-- hello.c
`-- output.txt
```

## Uso del Lenguaje C en mi Día a Día

## 4.1. Objetivos

 Identificar tecnologías basadas en el Lenguaje C y que usamos en nuestro día a día para estimar el nivel de adopción de C.

#### 4.2. Temas

· Lenguaje C.

#### 4.3. Tareas

- 1. Listar entre tres y diez tecnologías digitales que usamos en nuestro día a día.
- 2. Indicar para cada tecnología el repositorio público donde se la desarrolla, si es que lo tiene.
- 3. Indicar para cada una de esas tecnologías si se desarrollan en C o no.

#### 4.4. Restricciones

Ninguna.

#### 4.5. Productos

DD-UsoDeC
`-- UsoDeC.md

## Fases de la Traducción y Errores

## 5.1. Objetivos

· Identificar las fases de traducción y errores.

#### 5.2. Temas

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

#### 5.3. 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. Transicribir en readme.md cada comando ejecutado y
  - b. Describir en readme.md el **resultado** u **error** obtenidos para cada paso.

#### 5.3.1. Secuencia de Pasos

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

```
#include <stdio.h>

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

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

```
int printf(const char *s, ...);
int main(void){
  int i=42;
  prontf("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 hellos, 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 hellos.c y generar el ejecutable.
- 12Ejecutar y analizar el resultado.
- 13.Corregir en hello6.c y empezar de nuevo.
- 14Escribir hello.c, una nueva variante:

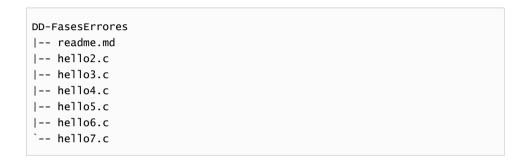
```
int main(void){
  int i=42;
    printf("La respuesta es %d\n", i);
}
```

15Explicar porqué funciona.

### 5.4. Restricciones

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

## 5.5. Productos



## 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 sección1.2 para que use una función de conversión.

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

## 6.1. Objetivos

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

#### 6.2. 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.

#### 6.3. Tareas

- 1. Escribir el Makefile.
- 2. Escribir Conversion.h
- Escribir ConversionTest.h
- 4. Escribir conversion.c
- 5. Escribir TablasDeConversion.c.

#### 6.4. Restricciones

- Las funciones deben llamarse Celsius y Farenheit.
- · Utilizar assert.
- · Utilizar const.
- Utilizar for con declaración (C99).

#### 6.5. Productos

```
DD-Interfaces
|-- readme.md
|-- Makefile
|-- Conversion.h
|-- ConversionTest.c
|-- Conversion.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 que sea genérica, es decir, pueda invocarse desde PrintTablaFahrenheit y desde PrintTablaCelsius. PrintFilas debe invocar a PrintFila.

## **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.
- \* Eiercicio 20 \*

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

## 7.1. Objetivos

- 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 alocación 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.

#### 7.2. Temas

- Strings.
- · Alocación.
- · Tipos.

#### 7.3. 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.

#### 7.4. 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>

#### 7.5. 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 del tipo String con una selección de operaciones de String del lenguaje Go, y de le implementación del ese tipo con representación en memoria igual a la de ese lenguaje, con la facilidad de alocar strings en el heap, y de 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.

## 8.1. Objetivos

- Especificación del tipo String que incluya una selección de operaciones 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.

#### 8.2. Temas

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

#### 8.3. 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 funciópn que recibe un string en GostringExample.c.
- h. Escribir la implementación en Gostring.c.

#### 8.4. Restricciones

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

#### 8.5. Productos

DD-GoStrings

#### Productos

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

# 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.

Problema: Imprimir 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.

# 9.1. Objetivos

- Aplicar máquinas de estado para el procesamiento de texto.
- Implementar máquinas de estado con diferentes métodos.

# 9.2. Temas

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

### 9.3. Tareas

1. Árboles de Expresión

- a. Estudiar el programa del ejemplo las 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 wl.gv la máquina de estado que resuelve el problema planteado.
- b. Formalizar la máquina de estados como una *n-upla*, basarse en el Capítulo #1 del Volumen #3 de [MUCH2012].
- Implementaciones de Máquinas de Estado:
   Las implementaciones varían en los conceptos que utilizan para representaar los estados y las transiciones.
  - a. Implementación #1: Una variable para el estado actual.
    - i. Escribir el programa wl-1-enum-switch.c que siga la Implementación #1, variante enum y switch.
      - Esta implementación es la *regularización* de la implementación de la sección 1.5.4 de [KR1988]. Los *estados* son valores de una variable y las *transiciones* son la selección estructurada y la actualización de esa variable. Esta versión es menos eficiente que la versión de [KR1988], pero su regularidad permite la automatización de la construcción del

programa que implementa la máquina de estados. Además de la regularidad, esta versión debe:

- 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;
- Utilizar switch en vez de if.
- ii. Responder en readme.md: Indicar ventajas y desventajas de la versión de [KR1988] y de esta implementción.
- b. Implementación #2: Sentencias goto (sí, el infame goto)
  - i. Leer la sección 3.8 Goto and labels de [KR1988]
  - ii. Leer Go To Statement Considered Harmful de [DIJ1968].
  - iii. Leer "GOTO Considered Harmful" Considered Harmful de [RUB1987].
  - iv. Responder en readme.md: ¿Tiene alguna aplicación *go to* hoy en día? ¿Algún lenguaje moderno lo utiliza?
  - v. Escribir el programa w1-2-goto.c que siga la Implementación #2. En esta implementación los *estados* son *etiquetas* y las *transiciones* son la selección estructurada y el salto incondicional con la sentencia goto.
- c. Implementación #3: Funciones Recursivas
  - i. Leer la sección 4.10 Recursividad de [KR1988].
  - ii. Responder en readme.md: ¿Es necesario que las funciones accedan a a contadores? Si es así, ¿cómo hacerlo?.
     Leer la sección 1.10 Variables Externas y Alcance y 4.3 Variables Externas de [KR1988].
  - iii. Escribir el programa, w1-3-rec.c que siga la implementación #3. En esta implementación los *estados* son *funciones recursivas* y las *transiciones* son la selección estructurada y la invocación recursiva.
- d. Implementación #X:
  - Es posible diseñar más implementaciones. Por ejemplo, una basada en una tabla que defina las transiciones de la máquina. En ese caso, el programa usaría la tabla para lograr el comportamiento deseado. El

objetivo de este punto es diseñar una implementación **diferente** a las implementaciones #1, #2, y #3.

- i. Diseñar una nueva implementación e indicar en Readme.md cómo esa implementación representa los estados y cómo las transiciones.
- ii. Escribir el programa, w1-x.c que siga la nueva implementación.

### 4. Eficiencia del uso del Tiempo:

Construir una tabla comparativa a modo de *benchmark* que muestre el tiempo de procesamiento para cada una de las cuatro 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.

La tabla tiene en las filas las cuatro implementación, en las columnas los tres archivos, y en la intersección la duración para una implementación para un archivo.

## 9.4. Restricciones

· Ninguna.

## 9.5. Productos

```
DD-wl
|-- readme.md
| |-- Árboles de expresión.
| |-- Respuestas.
| `-- Benchmark.
|-- wl.gv
|-- Makefile
|-- wl-1-enum-switch.c
|-- wl-2-goto.c
|-- wl-3-rec.c
|-- wl-x.c
```

# Máquinas de Estado — Contador de Palabras

Este trabajo está basado en el ejemplo de las 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."

Problema: Determinar la cantidad de líneas, palabra, y caracteres que se reciben por la entrada estándar.

# 10.1. Objetivos

- Aplicar máguinas de estado para el procesamiento de texto.
- Implementar máquinas de estado con diferentes métodos.

## 10.2. Temas

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

### 10.3. Tareas

# 1. Árboles de Expresión

- a. Estudiar el programa del ejemplo las 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 wc.gv la máquina de estado que resuelve el problema planteado.
- b. Formalizar la máquina de estados como una n-upla, basarse en el Capítulo #1 del Volumen #3 de [MUCH2012].
- Implementaciones de Máquinas de Estado:
   Las implementaciones varían en los conceptos que utilizan para representaar los estados y las transiciones.
  - a. Implementación #1: Una variable para el estado actual.
    - i. Escribir el programa wc-1-enum-switch.c que siga la Implementación #1, variante enum y switch.
      - Esta implementación es la *regularización* de la implementación de la sección 1.5.4 de [KR1988]. Los *estados* son valores de una variable y las *transiciones* son la selección estructurada y la actualización de esa variable. Esta versión es menos eficiente que la versión de [KR1988], pero su regularidad permite la automatización de la construcción del programa que implementa la máquina de estados. Además de la regularidad, esta versión debe:
      - 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;
      - Utilizar switch en vez de if.

- ii. Responder en readme.md: Indicar ventajas y desventajas de la versión de [KR1988] y de esta implementción.
- b. Implementación #2: Sentencias goto (sí, el infame goto)
  - i. Leer la sección 3.8 Goto and labels de [KR1988]
  - ii. Leer Go To Statement Considered Harmful de [DIJ1968].
  - iii. Leer "GOTO Considered Harmful" Considered Harmful de [RUB1987].
  - iv. Responder en readme.md: ¿Tiene alguna aplicación go to hoy en día? ¿Algún lenguaje moderno lo utiliza?
  - v. Escribir el programa wc-2-goto.c que siga la Implementación #2. En esta implementación los *estados* son *etiquetas* y las *transiciones* son la selección estructurada y el salto incondicional con la sentencia goto.
- c. Implementación #3: Funciones Recursivas
  - Leer la sección 4.10 Recursividad de [KR1988].
  - ii. Responder en readme.md: ¿Es necesario que las funciones accedan a a contadores? Si es así, ¿cómo hacerlo?. Leer la sección 1.10 Variables Externas y Alcance y 4.3 Variables Externas de [KR1988].
  - iii. Escribir el programa, wc-3-rec.c que siga la implementación #3.
    En esta implementación los estados son funciones recursivas y las transiciones son la selección estructurada y la invocación recursiva.
- d. Implementación #X:
  - Es posible diseñar más implementaciones. Por ejemplo, una basada en una tabla que defina las transiciones de la máquina. En ese caso, el programa usaría la tabla para lograr el comportamiento deseado. El objetivo de este punto es diseñar una implementación **diferente** a las implementaciones #1, #2, y #3.
  - i. Diseñar una nueva implementación e indicar en Readme.md cómo esa implementación representa los estados y cómo las transiciones.
  - ii. Escribir el programa, wc-x.c que siga la nueva implementación.

## 4. Eficiencia del uso del Tiempo:

Construir una tabla comparativa a modo de *benchmark* que muestre el tiempo de procesamiento para cada una de las cuatro 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.

La tabla tiene en las filas las cuatro implementación, en las columnas los tres archivos, y en la intersección la duración para una implementación para un archivo.

### 10.4. Restricciones

· Ninguna.

## 10.5. Productos

```
DD-wc
|-- readme.md
| |-- Árboles de expresión.
| |-- Respuestas.
| `-- Benchmark.
|-- wc.gv
|-- Makefile
|-- wc-1-enum-switch.c
|-- wc-2-goto.c
|-- wc-3-rec.c
`-- wc-x.c
```

# Máquinas de Estado — Sin 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

# 11.1. 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.

## 11.2. 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 la 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.
- Diseñar el archivo Makefile para que construya una, otra o ambas implementaciones, y para que ejecute las pruebas.

### 11.3. Productos

DD-SinComentarios
|-- readme.md
|-- RemoveComments.gv
|-- RemoveComments.c

-- Makefile

# **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, apóstrofos, secuencias de escape, y comentarios. (Este programa es difícil si lo hace en su completa generalidad.)

# 12.1. 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: {[}(])
- · Apóstrofos y comillas, secuencias de escape:
  - Válido: "[("
  - Inválido: "{}

## 12.2. Temas

Autómata de Pila (Push down Automata).

Stacks.

# 12.3. Tareas

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

# 12.4. 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.
- Definición formal del APD.
- Especficación de *PushString* basada en operaciones de cadenas de lenguajes formales.
- 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.
- Crear a mano el archivo de test funcional: Test.c.
- Diseñar y aplciar un módulo stack.
- Diseñar el programa para que pueda invocarse de la siguiente manera: >
   RemoveComments < Test.c | Parse</li>

### 12.5. Productos

```
DD-SimpleParser
```

### Productos

- |-- 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 las sección *5.12 Declaraciones Complicadas* de [KR1988].

# 13.1. 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;
    TokenType type;
    TokenValue val;
};
bool GetNextToken(Token *t /*out*/); // Retorna si pudo leer,
    almacena en t el token leido.
```

14

# Traductor de Declaraciones C a LN con Lex

Este trabajo está basado en el programa dc1 ejemplo de las sección *5.12 Declaraciones Complicadas* de [KR1988].

Esta versión se basa en un scanner generado por Lex.

# Traductor de Declaraciones C a LN con Lex & Yacc

Este trabajo está basado en el programa dc1 ejemplo de las sección *5.12 Declaraciones Complicadas* de [KR1988].

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

# Trabajo #4 — Módulo Stack (?)

# 16.1. Objetivos

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

### 16.2. 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.

### 16.3. 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 stackModuleLinkedDvnamic.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.

### 12Responder:

- 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.

# 16.4. 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.

# 16.5. Productos

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

# 16.6. Entrega

Opcional.

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

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

# 17.1. 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.

### 17.2. Temas

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

# 17.3. Tareas

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

Calculator	StackOfDoublesModule	Scanner
<ul> <li>Qué hace: Procesa entrada y muestra resultado.</li> </ul>	• Qué exporta:  • StackItem	<ul> <li>Qué hace: Obtiene operadores y operandos.</li> </ul>
resultado.  • Qué usa:  • Biblioteca Estándar  • EOF  • printf  • atof  • StackOfDoublesMo  • StackItem  • Push  • Pop  • IsEmpty  • IsFull  • Scanner  • GetNextToken  • TokenType	<ul><li>Push</li><li>Pop</li><li>IsEmpty</li><li>IsFull</li></ul>	operandos.  • Qué usa:  • Biblioteca Estándar  • getchar  • EOF  • isdigit  • ungetc  • Qué exporta:  • GetNextToken  • Token  • TokenType  • TokenValue
<ul><li>TokenValue</li></ul>		

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 l'exeme 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.

### 17.4. 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='*',
    Substraction='-',
    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 leido.
```

- 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

# 17.5. 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

# 17.6. Entrega

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

# Trabajo #7 — Calculadora Polaca con Lex (@)

Este trabajo está es una segunda iteración de Capítulo 17, *Trabajo #5 — Léxico de la Calculadora Polaca* (@), en la cual el *scanner* se implementa con *lex* y no con una máquina de estados.

# 18.1. Objetivo

Aplicar lex para el análisis lexicográfico.

#### 18.2. 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.

### 18.3. Productos

- Sufijo del nombre de la carpeta: PolcalLex.
- Los mismos que Capítulo 17, Trabajo #5—Léxico de la Calculadora Polaca
   (@) con la adición de /scanner.1

# 18.4. Entrega

Sep 6, 13hs.

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

# Trabajo #8 — Calculadora Infija con RDP (?)

Este trabajo es la versión infija de Capítulo 18, *Trabajo #7 — Calculadora Polaca con Lex (@)*; es decir en vez de procesar:

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

el programa debe procesar correctamente:

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

# 19.1. 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).

### 19.2. Restricciones

- Implementar GetnextToken con Lex, basado en el GetnextToken de Capítulo 18, Trabajo #7 — Calculadora Polaca con Lex (@)
- Agregar los tokens LParen y RParen.

# 19.3. Entrega

• Opcional

20

# Trabajo #9 — Calculadora Infija con Yacc (?)

Esta vez, el parser lo construye Yacc por nosotros.

# Bibliografía

- 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
- 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
- Brian W. Kernighan and Dennis Ritchie. *The C Programming Language, 2nd Edition* 1988.
- Jorge Muchnik y Ana María Díaz Bott. SSL, 2da Edición 2012.
- 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