Sintaxis y semántica de los lenguajes

Trabajo Práctico Integrador: Diseño e implementación de Lexer y Parser

Ciclo lectivo: 2023 Primer cuatrimestre

Integrantes Grupo 19 ISI B:

Acosta Quintana, Lautaro

Galeano, Martín

Niveyro, Iván

Morel, Francisco

Localidad:

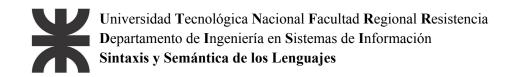
Resistencia, Chaco

Fecha primera entrega:

30/04/23

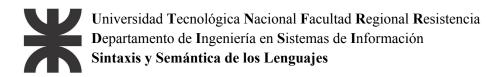
Universidad:

Universidad Tecnológica Nacional, Facultad Regional Resistencia



Contenido

Contenido	2
1. Introducción	3
2. Pseudocódigo	3



1. Introducción

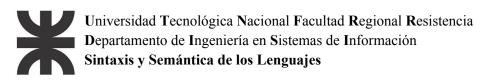
El trabajo consiste en la construcción de un analizador léxico y un analizador sintáctico con el fin de analizar un archivo de entrada, de formato y extensión .docbook, y validarlo para posteriormente traducirlo a otro archivo con formato y extensión .HTML.

Para lograrlo, en esta entrega mostraremos un diseño en primera instancia de las distintas reglas de construcción con las que se evaluará la entrada del programa. Este diseño está realizado en pseudocódigo y representa cómo se comportaría el analizador sintáctico.

2. Pseudocódigo

En esta sección se presenta el pseudocódigo inicial en el que se basará el analizador sintáctico. Los terminales están escritos en minúscula y los no terminales en *CamelCase*. Con la intención de facilitar la lectura, los no terminales están coloreados en azul y las producciones gramaticales están divididas en secciones como en el pdf de guía.

Tabla de Tokens		
Token	Contenido	
texto	Letras, números, signos de puntuación y caracteres especiales	
numero	Números enteros	
extension	Extensión de un archivo	
<article>; </article>	Etiqueta de inicio y final de un bloque articulo	
<info>; </info>	Etiqueta de inicio y final de un bloque info	
<title>; </title>	Etiqueta de inicio y final de un bloque title	
<abstract>; </abstract>	Etiqueta de inicio y final de un bloque abstract	
<para>; </para>	Etiqueta de inicio y final de un bloque para	
<author>; </author>	Etiqueta de inicio y final de un author	
<personname>; </personname>	Etiqueta de inicio y final de un nombre completo de persona	
<firstname>; </firstname>	Etiqueta de inicio y final del primer nombre de una persona	
<surname>; </surname>	Etiqueta de inicio y final del segundo	



	nombre de una persona
<date>; </date>	Etiqueta de inicio y final para una fecha
<section>; </section>	Etiqueta de inicio y final de una sección
<simplesect>;</simplesect>	Etiqueta de inicio y final de una sección simple
<copyright>; </copyright>	Etiqueta de inicio y final para bloque de copyright
<address>; </address>	Etiqueta de inicio y final de una dirección
<city>; </city>	Etiqueta de inicio y final del nombre de una ciudad
<state>; </state>	Etiqueta de inicio y final del nombre de una provincia/estado
<postcode>; </postcode>	Etiqueta de inicio y final de un código postal
<street>; </street>	Etiqueta de inicio y final del nombre de una calle
<email>; </email>	Etiqueta de inicio y final de una dirección de correo
<phone>; </phone>	Etiqueta de inicio y final de un número de correo
<itemizedlist>; </itemizedlist>	Etiqueta de inicio y final para una lista itemizada
	Etiqueta de inicio y final para una lista
<emphasis>; </emphasis>	Etiqueta de inicio y final de bloque emphasis
<holder>; </holder>	Etiqueta de inicio y final de bloque holder
<simpara>; </simpara>	Etiqueta de inicio y final de bloque simpara
<year>; </year>	Etiqueta de inicio y final para un número de año
<comment>; </comment>	Etiqueta de inicio y final para realizar comentarios
<important>; </important>	Etiqueta de inicio y final de bloque

	important
k>; 	Etiqueta de inicio y final del bloque link
<videodata ;="" fileref="; /></td><td>Etiqueta para adjuntar un archivo del tipo video</td></tr><tr><td><imagedata fileref="></videodata>	Etiqueta para adjuntar un archivo del tipo imagen
<informaltable>; </informaltable>	Etiqueta de inicio y final de bloque informaltable
<tgroup>; </tgroup>	Etiqueta de inicio y final de bloque tgroup
<row>; </row>	Etiqueta de inicio y final para indicar el número de filas de una tabla
;	Etiqueta de inicio y final para crear una tabla
<thead>; </thead> <tfoot>; </tfoot> ;	Etiquetas de inicio y final para la dimensiones de una tabla
<entrytbl>; </entrytbl>	Etiqueta de inicio y final de bloque entrytbl
<pre><link ;="" xlink:href:=""/></pre>	Etiqueta para URL's

Etiquetas estructurales:

```
Article → <article>Info Title Content Section</article>
<article>Info Title Content SimSection</article>
| <article>Info Title Content</article>
| <article>Info Content Section</article>
| <article>Info Content SimSection</article>
| <article>Info Content</article>
| <article>Title Content Section</article>
<article>Title Content SimSection</article>
| <article>Title Cuerpo</article>
| <article>Content Section</article>
| <article>Content SimSection</article>
<article>Content</article>
Content → ItemizedList Content
                                 | ItemizedList
| Important Content
                                 | Important
| Para Content
                                 Para
```

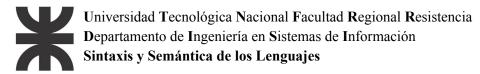
SimPara Content	SimPara
Address Content	Address
MediaObject Content	MediaObject
InformalTable Content	InformalTable
Comment Content	Comment
Abstract Content	Abstract

Secciones:

```
Section >> <section> Info Title Content SimSection
| <section>Info Title Content Section
| <section>Info Content Section
| <section>Info Content SimSection
| <section>Title Content Section
| <section>Title Content SimSection
| <section>Title Content SimSection
| <section>Content Section
| <section>Content SimSection
| <section>Content SimSection
| <section>Content 
| <section>Content 
| <simplesect>Info Title Content
| <simplesect>Info Content
| <simplesect>Title Content
| <simplesect>Title Content
| <simplesect>Content
| <simplesect>Content
```

Etiquetas básicas de párrafo:

```
InfoContent → Title InfoContent | Title
| InfoContent MediaObject | MediaObject
                             Abstract
InfoContent Abstract
InfoContent Address
                             Address
InfoContent Author
                              Author
InfoContent Date
                              Date
InfoContent Copyright
                              Copyright
Info → <info> InfoContent </info>
AbstractContent → Para AbstractContent | Para
| SimPara AbstractContent | SimPara
Abstract → <abstract> Title AbstractContent </abstract>
<abstract> AbstractContent </abstract>
AddressContent → texto AddressContent | texto
| Street AddressContent | Street
| City AddressContent | City
| State AddressContent | State
```



```
| Phone AddressContent | Phone
| Email AddressContent | Email
Address → <address>AddressContent</address>
AuthorContent → Firstname AuthorContent
| Surname AuthorContent
Author → <author> AuthorContent </author>
CopyrightYearContent → Year CopyrightYearContent
Year CopyrightHolderContent
Year
CopyrightHolderContent → Holder CopyrightHolderContent | Holder
Copyright → <copyright>CopyrightYearContent</copyright>
TitleContent → Emphasis TitleContent | Emphasis
| Link TitleContent
                            Link
| Email TitleContent
                            | Email
| texto TitleContent
                            texto
Title → <title>TitleContent</title>
SimParaContent → Emphasis | texto
                 Email
Link
Author
                 Comment
SimPara → <simpara>SimPara SimParaContent</simpara>
<simpara>SimParaContent</simpara>
Emphasis → <emphasis>Emphasis SimParaContent</emphasis>
<emphasis>SimParaContent</emphasis>
Comment → <comment>Comment SimParaContent</comment>
<comment>SimParaContent</comment>
Link → k>Link SimParaContent</link>
<link>SimParaContent</link>
ParaContent → Emphasis
                       Link
| Email | Author
                      Comment
                                      | ItemizedList
| Important | Address
                      | MediaObject | InformalTable
```

```
Para → <para>Para ParaContent</para>
Important → <important>Title Content</important>
<important>Content</important>
SharedContent → SharedContent Comment | Comment
| SharedContent Emphasis | Emphasis
| SharedContent Link
                           Link
| SharedContent texto
                           texto
          → <firstname>SharedContent</firstname>
FirstName
Surname
          → <surname>SharedContent</surname>
          → <street>SharedContent</street>
Street
          → <city>SharedContent</city>
City
          → <phone>SharedContent</phone>
Phone
Email
          → <email>SharedContent</email>
          → <date>SharedContent</date>
Date
          → <year>SharedContent</year>
Year
Holder
          → <holder>SharedContent</holder>
          → <state>SharedContent</state>
State
```

Imágenes y multimedia:

```
MediaObjectContent → VideoObject MediaObjectContent | VideoObject
| ImageObject MediaObjectContent | ImageObject
MediaObject → < mediaobject>Info VideoObject
MediaObjectContent</mediaobject>
</pre
| <mediaobject>ImageObject MediaObjectContent</mediaobject>
| <mediaobject>VideoObject MediaObjectContent</mediaobject>
<mediaobject>VideoObject</mediaobject>
| <mediaobject>VideoObject</mediaobject>
ImageObject→<imageobject> Info ImageData</imageoobject>
|<imageobject>ImageData</imageoobject>
VideoObject→<videoobject>Info VideoData</videoobject>
<videoobject>VideoData</videoobject>
VideoData → <videodata fileref ="Directorio" />
ImageData → <imagedata fileref="Directorio" />
Directorio → /texto Directorio/ | texto.extension
```

Listas:

```
ItemizedList→ <itemizedlist>ListItem</itemizedlist>
  | <itemizedlist>ListItem ItemizedList</itemizedlist>

ListItem → listitem>ListItem Content</listitem>
  | | listitem>Content</listitem>
```

Tablas:

```
InformalTableContent → MediaObject InformalTableContent | MediaObject
| Tgroup InformalTableContent | Tgroup
InformalTable →<informaltable>InformalTableContent</informaltable>
Tgroup → <tgroup>Thead Tbody Tfoot</tgroup>
<tgroup>Thead Tfoot</tgroup>
<tgroup>Tbody Tfoot</tgroup>
TableContent → Row TableContent | Row
Thead→<thead>TableContent</thead>
Tbody→TableContent
Tfoot→<tfoot>TableContent</tfoot>
RowContent → Entry RowContent | Entry
| EntryTbl RowContent
                            | EntryTbl
Row → <row>RowContent</row>
EntryContent → texto EntryContent | texto
| ItemizedList EntryContent | ItemizedList
Important EntryContent
                                   | Important
| Para EntryContent
                                   Para
| SimPara EntryContent
                             SimPara
MediaObject EntryContent
                            MediaObject
Comment EntryContent
                             Comment
Abstract EntryContent
                                   Abstract
Entry→<entry>EntryContent</entry>
Entrytbl→<entrytbl>Thead Tbody</entrytbl>
Entrytbl→<entrytbl>Tbody</entrytbl>
```

Enlaces:

```
Link → <link xlink:href: Url />
Url → Protocolo://Dominio:Puerto/Directorio#LocalizadorInterno
| Protocolo://Dominio:Puerto/Directorio
| Protocolo://Dominio:Puerto
```

```
| Protocolo://Dominio/Directorio # LocalizadorInterno
| Protocolo://Dominio/Directorio
| Protocolo://Dominio#LocalizadorInterno
| Protocolo://Dominio
| Protocolo → http | https | ftp | ftps
| Dominio → texto
| Puerto → numero
| LocalizadorInterno → texto
```