

4.1 Tabla de Tokens para el proyecto de Analizador Léxico

Nombre del lenguaje: **HTML-es**

Definición del lenguaje

1. Palabras clave:

- Las palabras clave en HTML-es son las etiquetas que se utilizan para estructurar y dar formato al contenido de una página web
 - **<html>**: Define el inicio del documento HTML.
 - **<encabezado>**: Contiene información sobre el documento como el título, enlaces a hojas de estilo y scripts, etc.
 - **<cuerpo>**: Contiene todo el contenido visible de la página.
 - **<título>**: Define el título del documento que aparece en la barra de título del navegador.
 - **<p>**: Define un párrafo de texto.
 - **<bloque>** y **<contenedor>**: Elementos de bloque y en línea respectivamente, que se utilizan para agrupar y estilizar el contenido.
 - **<enlace>**: Define un enlace a otra página web.
 - ****: Inserta una imagen en la página.
 - ****, ****, ****: Se utilizan para crear listas sin orden y ordenadas, y elementos de lista respectivamente.

2. Operadores:

- En HTML-es, los operadores principales son los símbolos de las etiquetas de apertura y cierre.
 - **<**: Símbolo de etiqueta de apertura.
 - **>**: Símbolo de etiqueta de cierre.
 - **/**: Utilizado para cerrar etiquetas.
 - **=**: Utilizado para asignar valores a los atributos.

3. Símbolos especiales:

- En HTML-es, los símbolos especiales son entidades que representan caracteres especiales que tienen un significado especial
 - **<**: Representa el símbolo **<**.
 - **>**: Representa el símbolo **>**.
 - **&**: Representa el símbolo **&**.

- **";** Representa el símbolo ".
- ** ;** Representa un espacio en blanco no rompible.

4. Comentarios:

- Los comentarios en HTML-es se escriben entre `<!--` y `-->`. Por ejemplo:
`<!-- Este es un comentario en HTML-es -->`

5. Atributos:

- Los atributos proporcionan información adicional sobre un elemento HTML-es y se escriben dentro de las etiquetas
 - **urle**: Define la URL de destino de un enlace.
 - **urlrec**: Define la URL de origen de un recurso, como una imagen.
 - **id**: Define un identificador único para el elemento.
 - **clase**: Asigna una o más clases CSS al elemento para aplicar estilos.

6. Entidades HTML:

- Se utilizan para representar caracteres especiales que no se pueden escribir directamente en HTML-es.
 - **©;** Representa el símbolo de copyright.
 - **€;** Representa el símbolo del euro.
 - **—;** Representa un guion largo.

7. Elementos de formulario:

- Incluyen etiquetas como `<input>`, `<textarea>`, `<select>`, `<button>`, etc. Cada uno tiene sus propios atributos y valores asociados, como **type**, **name**, **value**, etc.

8. DOCTYPE:

- Define la versión de HTML-es que se está utilizando en el documento.
`<!DOCTYPE html>`

Especificacion de Tokens

- **Palabras clave:**
 - `<html>`, `<encabezado>`, `<cuerpo>`, `<titulo>`, `<p>`, `<bloque>`, `<contenedor>`, `<enlace>`, ``, ``, ``, ``
- **Identificadores:**
 - `[a-zA-Z][a-zA-Z0-9_]*`.
- **Números:**
 - `\d+(\.\d+)?`.
- **Operadores:**
 - `<`, `>`, `/`, `=`
- **Símbolos especiales:**
 - `&`, `<`, `>`, `"`
- **Comentarios:**
 - `<!--.*-->`.

Patrones regulares de cada tipo de token.

- **Palabras clave:**
 - `<html>`, `<encabezado>`, `<cuerpo>`, `<titulo>`, `<p>`, `<bloque>`, `<contenedor>`, `<enlace>`, ``, ``, ``, ``
- **Identificadores:**
 - `[a-zA-Z][a-zA-Z0-9_]*`
- **Números:**
 - `\d+(\.\d+)?`
- **Operadores:**
 - Los operadores son principalmente los símbolos de las etiquetas de apertura y cierre, por lo que no necesitan patrones regulares adicionales.
- **Símbolos especiales:**
 - `&`, `<`, `>`, `"`, etc.
- **Comentarios:**
 - `<!--.*-->`

Tabla de Tokens

Token	Tipo	Identificador
HTML	Palabra clave	página
HEAD	Palabra clave	encabezado
TITLE	Palabra clave	título
BODY	Palabra clave	cuerpo
H1	Palabra clave	e1
H2	Palabra clave	e2
H3	Palabra clave	e3
H4	Palabra clave	e4
H5	Palabra clave	e5
H6	Palabra clave	e6
P	Palabra clave	P
DIV	Palabra clave	bloque
SPAN	Palabra clave	contenedor
A	Palabra clave	enlace
IMG	Palabra clave	imagen
SRC	Atributo	Src
HREF	Atributo	Href
ALT	Atributo	Alt
ID	Atributo	Id
CLASS	Atributo	Clase
STYLE	Atributo	Estilo
WIDTH	Atributo	Ancho

Token	Tipo	Identificador
HEIGHT	Atributo	Alto
TITLE	Atributo	Título
EQUALS	Operador	=
QUOTE	Símbolo	"
DOUBLE_QUOTE	Símbolo	""
SEMICOLON	Símbolo	;
LEFT_ANGLE	Símbolo	<
RIGHT_ANGLE	Símbolo	>
COMMENT	Comentario	<!--