

Carlos Fernando Ramos Mena

A01197622

Actividad 1.7 Tokenizacion (primer avance de la situación problema)

Proyecto: <https://replit.com/@a01197622/Tokenizacion?v=1#main.py>

Tipo de Lexema	Expresion Regular	Cadenas Probadas
Identificadores	<code>r'[a-zA-Z_]+[a-zA-Z\d\w]*'</code>	Var1 _var1 var12_ok
<div><div>code.txt</div><div>1 Var1 2 _var1 3 var12_ok</div><div>PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS</div><div>● PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py ['identificador'] ['identificador'] ['identificador'] ○ PS C:\Users\Ferna\Desktop\IMC&gt; </div></div>		
Operadores	<code>r'\+'</code> <code>r'\-'</code> <code>r'\*'</code> <code>r'\/'</code> <code>r'\^'</code> <code>r'\&lt;=</code> <code>r'\&gt;=</code> <code>r'\!='</code> <code>r'\&lt;'</code> <code>r'\&gt;'</code> <code>r'\='</code>	+ - * / ^ <= >= != < > =

Carlos Fernando Ramos Mena

A01197622

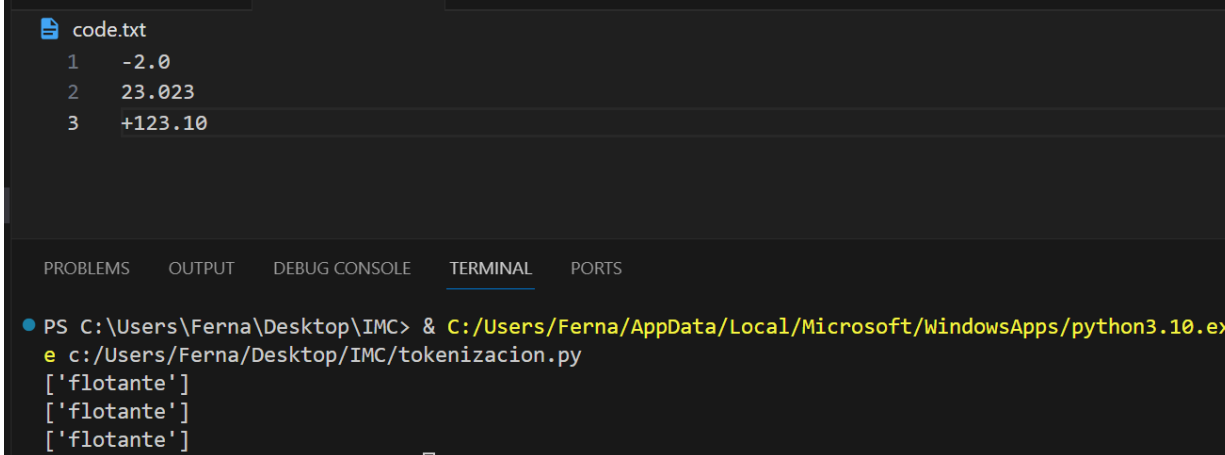
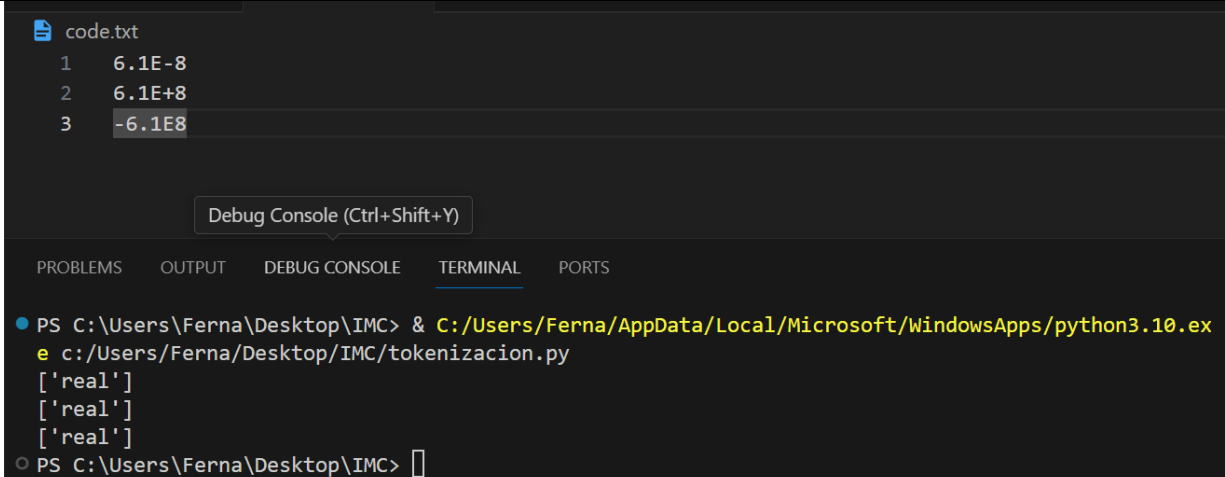
Actividad 1.7 Tokenizacion (primer avance de la situación problema)

<pre>1 + 2 - 3 * 4 / 5 ^ 6 &lt;= 7 &gt;= 8 != 9 &lt; 10 &gt; 11 =</pre> <p>PROBLEMS OUTPUT DEBUG CONSOLE <u>TERMINAL</u> PORTS</p> <pre>e c:/Users/Ferna/Desktop/IMC/tokenizacion.py ['suma'] ['resta'] ['multiplicación'] ['division'] ['potencia'] ['menor_igual'] ['mayor_igual'] ['diferente'] ['menor_que'] ['mayor_que'] ['asignacion']</pre>		
Enteros	<code>r'[+-]?\\d+'</code>	--2,023 +123 0123
<p>code.txt</p> <pre>1 -2023 2 +123 3 0123</pre> <p>PROBLEMS OUTPUT DEBUG CONSOLE <u>TERMINAL</u> PORTS</p> <pre>PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe e c:/Users/Ferna/Desktop/IMC/tokenizacion.py ['entero'] ['entero'] ['entero'] PS C:\Users\Ferna\Desktop\IMC&gt; </pre>		
Flotantes	<code>r'[+-]?\\d+\\.\\d+'</code>	-2.0 23.023 +123.10

Carlos Fernando Ramos Mena

A01197622

## Actividad 1.7 Tokenizacion (primer avance de la situación problema)

		
Números reales	<code>r'[+-]?d+(\.\d+)?[Ee][+-]?d+'</code>	6.1E-8 6.1E+8 -6.1E-8
		
Cadenas	<code>r'"[^"]*"'</code>	"hola" "hola mundo" "" "hola_mundo2" "h0l4 t0d0s"

Carlos Fernando Ramos Mena

A01197622

Actividad 1.7 Tokenizacion (primer avance de la situación problema)

<div><div>code.txt</div><div><div>1"hol</div><div>2"hol mundo</div><div>3"</div><div>4"hol_mundo2"</div><div>5"h0l4 t0d0s"</div></div></div> <div><div>PROBLEMSOUTPUTDEBUG CONSOLETERMINALPORTS</div><div>kfra8p0\lib\encodings\cp1252.py", line 23, in decode return codecs.charmap_decode(input,self.errors,decoding_table)[0] UnicodeDecodeError: 'charmap' codec can't decode byte 0x9d in position 9: character maps to &lt;undefined&gt; PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py ['cadena'] ['cadena'] ['cadena'] ['cadena'] ['cadena']</div></div>		
Reservadas	Se usa el mismo de los identificadores: <code>r'[a-zA-Z_]+[a-zA-Z\d\w]*'</code>  luego se comprueba si este identificador se encuentra en el array keyword:  if kind == 'identificador' and value in keywords: kind = "reservada"	for if else
<div><div>code.txt</div><div><div>1for</div><div>2if</div><div>3else</div></div></div> <div><div>PROBLEMSOUTPUTDEBUG CONSOLETERMINALPORTS</div><div>PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py ['reservada'] ['reservada'] ['reservada'] PS C:\Users\Ferna\Desktop\IMC&gt; </div></div>		
Comentarios	<code>r'//[^\n]*'</code>	// hola mundo // // h0l4 tod0s_

Carlos Fernando Ramos Mena

A01197622

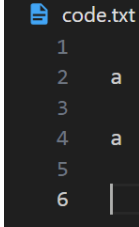
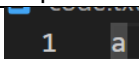
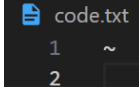
## Actividad 1.7 Tokenizacion (primer avance de la situación problema)

<pre>1 // hola mundo 2 // 3 // h0l4 tod0s_</pre> <div>PROBLEMS OUTPUT DEBUG CONSOLE <u>TERMINAL</u> PORTS</div> <pre>PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py ['comentario'] ['comentario'] ['comentario'] PS C:\Users\Ferna\Desktop\IMC&gt; </pre>		
Puntuación	<pre>r'\(' r'\)' r'\{' r'\}' r'\[' r'\]'</pre>	<pre>( ) { } [ ]</pre>
<div>code.txt</div> <pre>1 ( 2 ) 3 { 4 } 5 [ 6 ] 7 </pre> <div>PROBLEMS OUTPUT DEBUG CONSOLE <u>TERMINAL</u> PORTS</div> <pre>PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py ○ ['a_parentesis']   ['c_parentesis']   ['a_corchete']   ['c_corchete']   ['a_corchete_box']   ['c_corchete_box']</pre>		
newline (ignorar)	<pre>r'\n'</pre>	<pre>a a</pre>

Carlos Fernando Ramos Mena

A01197622

Actividad 1.7 Tokenizacion (primer avance de la situación problema)

 <pre>code.txt 1 2   a 3 4   a 5 6    </pre>	<p>PROBLEMS OUTPUT DEBUG CONSOLE <u>TERMINAL</u> PORTS</p> <pre>PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py ['identificador'] ['identificador']</pre>	
<p>skip</p>  <pre>code.txt 1  a  a</pre>	<p>r'[\t]+'</p>	<p>a a</p>
<p>mismatch</p>	<p>r'</p>	<p>~ %   °</p>
 <pre>code.txt 1  ~ 2</pre> <p>PROBLEMS OUTPUT DEBUG CONSOLE <u>TERMINAL</u> PORTS</p> <pre>PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py Traceback (most recent call last):   File "c:\Users\Ferna\Desktop\IMC\tokenizacion.py", line 64, in &lt;module&gt;     for token in tokens:   File "c:\Users\Ferna\Desktop\IMC\tokenizacion.py", line 54, in tokenize     raise RuntimeError(f'{value!r} unexpected on line {line_num}') RuntimeError: '~' unexpected on line 1</pre>		

Carlos Fernando Ramos Mena

A01197622

Actividad 1.7 Tokenizacion (primer avance de la situación problema)

<div><div>code.txt</div><div>1 % 2</div><div>PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS</div><div>PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py Traceback (most recent call last):   File "c:\Users\Ferna\Desktop\IMC\tokenizacion.py", line 64, in &lt;module&gt;     for token in tokens:   File "c:\Users\Ferna\Desktop\IMC\tokenizacion.py", line 54, in tokenize     raise RuntimeError(f'{value!r} unexpected on line {line_num}') RuntimeError: '%' unexpected on line 1 PS C:\Users\Ferna\Desktop\IMC&gt;</div></div>		
<div><div>code.txt</div><div>1    2</div><div>PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS</div><div>PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py Traceback (most recent call last):   File "c:\Users\Ferna\Desktop\IMC\tokenizacion.py", line 64, in &lt;module&gt;     for token in tokens:   File "c:\Users\Ferna\Desktop\IMC\tokenizacion.py", line 54, in tokenize     raise RuntimeError(f'{value!r} unexpected on line {line_num}') RuntimeError: ' ' unexpected on line 1 PS C:\Users\Ferna\Desktop\IMC&gt;</div></div>		
<div><div>code.txt</div><div>1 º 2</div><div>PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS</div><div>PS C:\Users\Ferna\Desktop\IMC&gt; &amp; C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py Traceback (most recent call last):   File "c:\Users\Ferna\Desktop\IMC\tokenizacion.py", line 64, in &lt;module&gt;     for token in tokens:   File "c:\Users\Ferna\Desktop\IMC\tokenizacion.py", line 54, in tokenize     raise RuntimeError(f'{value!r} unexpected on line {line_num}') RuntimeError: 'Â' unexpected on line 1 PS C:\Users\Ferna\Desktop\IMC&gt;</div></div>		
Ejemplo	Todas las anteriores	<div>b = 7  a = 32.4 * (-8.6 - b) / 6.1E-8  d = a ^ b // Esto es un comentario</div>

Carlos Fernando Ramos Mena

A01197622

Actividad 1.7 Tokenizacion (primer avance de la situación problema)

code.txt

```
1  b = 7
2
3  a = 32.4 * (-8.6 - b) / 6.1E-8
4
5  d = a ^ b // Esto es un comentario
```

PROBLEMS

OUTPUT

DEBUG CONSOLE

TERMINAL

PORTS

```
PS C:\Users\Ferna\Desktop\IMC> & C:/Users/Ferna/AppData/Local/Microsoft/WindowsApps/python3.10.exe c:/Users/Ferna/Desktop/IMC/tokenizacion.py
['identificador', 'asignacion', 'entero']
['identificador', 'asignacion', 'flotante', 'multiplicación', 'a_parentesis', 'flotante', 'resta', 'identificador', 'c_parentesis', 'division', 'real']
['identificador', 'asignacion', 'identificador', 'potencia', 'identificador', 'comentario']
PS C:\Users\Ferna\Desktop\IMC> 
```