

Universidad Rafael Landívar

Facultad de Ingeniería

Lenguajes Formales y Autómatas

Sección: 02

Ing. Campos Gonzáles Vivian Damaris

Proyecto 2 - Fase 1

Expresiones Regulares

Zane Christian Bran Lockridge - 1373724
Esteban Josué Maldonado Velasco - 1053424

Guatemala 22 de Octubre del 2025

Expresiones Regulares

Las Expresiones Regulares se crearon basándose en el archivo de prueba y los resultados esperados que la ingeniera a cargo publicó en el portal de clase.

Token	Expresión Regular	Ejemplo
OPE_A	<Operacion=\s*(SUMA RESTA MULTIPLICACION DIVISION POTENCIA RAIZ INVERSO MOD)\s*>	<Operacion= SUMA>
NUM	<Numero>\s*(\d+(\.\d+)?)\s*</Numero>	<Numero> 12.75 </Numero>
NUM_P	<P>\s*\d+\s*</P>	<P> 3 </P>
NUM_R	<R>\s*\d+\s*</R>	<R> 5 </R>
OPE_C	</Operacion>	</Operacion>
ERROR	Cualquier otro símbolo u orden	4..5, .99, 78;91, etc

Los nombres asignados a los tokens se eligieron como abreviaciones representativas de su función dentro del analizador léxico. A continuación, se detalla el significado de cada uno:

- OPE_A: Corresponde a la función “Operación Abrir”, encargada de identificar el inicio de una operación mediante la etiqueta <Operacion= ...>
- NUM: Hace referencia a los valores numéricos que serán utilizados en las diferentes operaciones.
- NUM_P: Se utiliza cuando la operación es POTENCIA, representando el número que actuará como exponente dentro de la operación.
- NUM_R: Similar al caso anterior, pero se aplica cuando la operación es RAIZ, representando el valor base sobre el cual se calculará la raíz.
- OPE_C: Corresponde a la función “Operación Cerrar”, encargada de reconocer el final de una operación mediante la etiqueta </Operacion>
- ERROR: Detecta y reporta las entradas inválidas sin detener el scanner.

Se utilizó el patrón \s* para representar los posibles espacios en blanco entre los caracteres. Esto se debe a que, en los ejemplos proporcionados, se observa la presencia de espacios entre los símbolos < > y los valores numéricos empleados en las distintas funciones.

Por tal motivo, se consideró necesario permitir de cero a múltiples espacios sin que esto afecte la interpretación de las operaciones, ya que no se indicó que dichos espacios deban tratarse como errores léxicos.

No obstante, en caso de que durante la retroalimentación de la Fase 1 se determine que esto debe restringirse, el patrón será ajustado en la versión siguiente del proyecto Fase 2.

En los casos donde se deben leer valores numéricos, se empleó el símbolo `\d`, ya que tras realizar una investigación se determinó que esta es la forma más formal y estándar para representar dígitos del 0 al 9 dentro de las expresiones regulares. Esta notación permite abarcar todos los números posibles sin necesidad de enumerarlos explícitamente. De igual manera, se utilizó `\.` para representar un punto literal, el cual se utiliza para separar la parte entera de la parte decimal en los números flotantes.

Cabe mencionar que, aunque podría emplearse la letra N como una representación simbólica de los números, el uso de `\d` resulta más apropiado y formal, pues se ajusta a la notación reconocida en la teoría y práctica de expresiones regulares.