



TAREA 2.1

Expresión Regular

INSTITUTO TECNOLÓGICO DE PACHUCA

INGENIERÍA EN SISTEMAS COMPUTACIONALES

MATERIA:

Lenguajes Autómatas

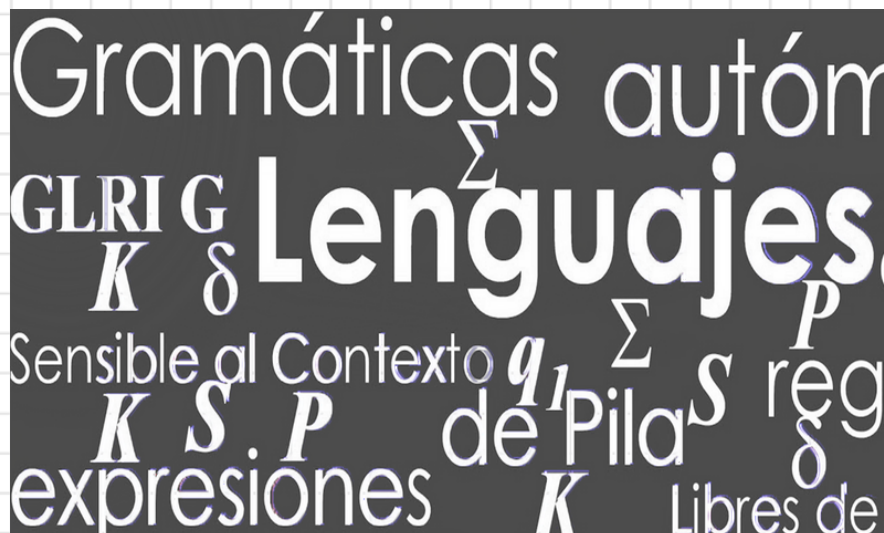
PROFESOR:

Rodolfo Baume Lazcano

NOMBRE:

Emanuel Tolentino Santander

21200635

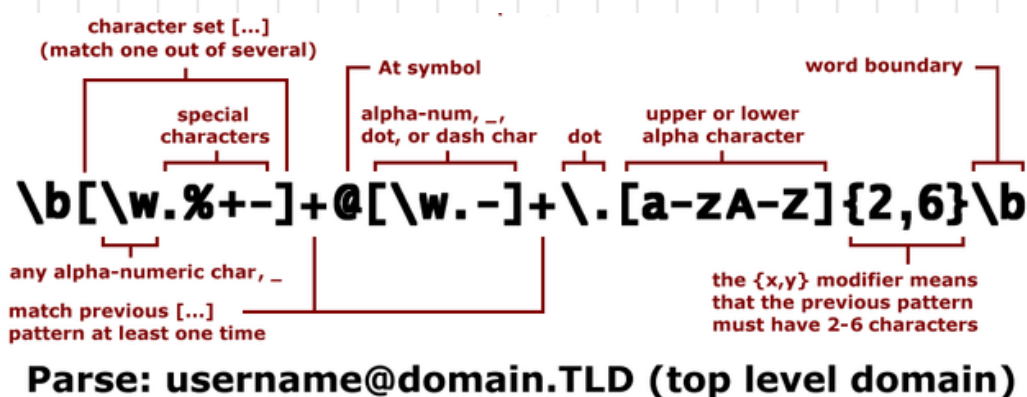


INTRODUCCION

Una expresión regular es una secuencia de caracteres que define un patrón de búsqueda en un texto. Estos patrones pueden ser utilizados para buscar, validar o manipular cadenas de texto de manera eficiente.

Las expresiones regulares se componen de varios elementos, incluyendo caracteres literales, metacaracteres y cuantificadores, que permiten definir patrones complejos.

Son una herramienta fundamental para el procesamiento y manipulación de texto, y son ampliamente utilizadas en la programación y la informática en general.



¿QUÉ ES UNA EXPRESION REGULAR?

Es una secuencia de caracteres que conforma un patrón de búsqueda. Se utilizan principalmente para la búsqueda de patrones de cadenas de caracteres u operaciones de sustituciones.

Las expresiones regulares son patrones utilizados para encontrar una determinada combinación de caracteres dentro de una cadena de texto. Las expresiones regulares proporcionan una manera muy flexible de buscar o reconocer cadenas de texto.

IMPORTANCIA

- **Búsqueda de patrones:** Permiten encontrar patrones específicos dentro de cadenas de texto de manera rápida y precisa. Esto es útil en tareas como la extracción de información, la validación de formatos de datos, la búsqueda de palabras clave, etc.
- **Manipulación de texto:** Facilitan la manipulación y transformación de cadenas de texto. Puedes reemplazar, eliminar, o modificar partes de un texto basado en un patrón definido por una expresión regular.
- **Validación de datos:** Son fundamentales para validar la entrada de datos en aplicaciones. Por ejemplo, puedes usar expresiones regulares para verificar si una dirección de correo electrónico, un número de teléfono o una contraseña cumplen con ciertos criterios de formato.

- **Automatización de tareas:** Permiten automatizar tareas que implican procesamiento de texto. Por ejemplo, en la extracción de datos de documentos, en el procesamiento de registros de archivos de registro, en la normalización de datos, entre otros.
- **Flexibilidad y portabilidad:** Las expresiones regulares son independientes del lenguaje de programación o del entorno de desarrollo. Puedes utilizar las mismas expresiones regulares en diferentes lenguajes y herramientas, lo que las hace portátiles y reutilizables.
- **Eficiencia:** Son altamente eficientes en términos de tiempo de procesamiento y recursos. Al utilizar patrones predefinidos, las expresiones regulares pueden realizar búsquedas y manipulaciones de texto de manera rápida y escalable.

CASO DE USO

- **Validación de Formularios:** Al desarrollar aplicaciones web, se pueden utilizar expresiones regulares para validar la entrada de datos en formularios. Por ejemplo, puedes verificar si un campo de correo electrónico tiene un formato válido, si un número de teléfono es válido, o si una contraseña cumple con ciertos requisitos de complejidad.
- **Extracción de Información:** Si necesitas extraer información específica de un conjunto de datos, como números de teléfono, direcciones de correo electrónico o URLs, puedes utilizar expresiones regulares para buscar y extraer esa información de manera eficiente.

- **Análisis de Logs:** En el análisis de archivos de registro (logs) de servidores web u otras aplicaciones, las expresiones regulares pueden ayudar a extraer información relevante, como direcciones IP, tipos de solicitud, códigos de respuesta, etc.
- **Transformación de Datos:** Puedes utilizar expresiones regulares para transformar datos en diferentes formatos. Por ejemplo, reformatear fechas, eliminar caracteres no deseados, o normalizar nombres y direcciones.
- **Búsqueda y Reemplazo de Texto:** Al editar documentos de texto o archivos de código fuente, puedes utilizar expresiones regulares para buscar y reemplazar texto de manera eficiente. Esto es especialmente útil para hacer cambios en múltiples lugares de manera simultánea.
- **Análisis de Texto:** En el procesamiento de lenguaje natural (NLP) y en la minería de texto, las expresiones regulares pueden utilizarse para identificar patrones lingüísticos específicos, como nombres de entidades, fechas, números, etc.
- **Web Scraping:** Al extraer datos de páginas web, las expresiones regulares pueden identificar y extraer información específica de manera precisa y eficiente.
- **Manipulación de Cadenas de Texto:** En tareas de manipulación de cadenas de texto en general, como la limpieza de datos, la normalización de texto, la tokenización, etc.

EJEMPLOS

- Validación de Correos Electrónicos:
 - Expresión regular: `^[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}$`
 - Esta expresión regular valida si una cadena es un correo electrónico válido, verificando que tenga un formato adecuado.
- Extracción de Números de Teléfono:
 - Expresión regular: `\b\d{3}[-.]?\d{3}[-.]?\d{4}\b`
 - Esta expresión regular encuentra números de teléfono en formato común de 10 dígitos, con o sin guiones o puntos.
- Búsqueda de URLs en un Texto:
 - Expresión regular: `https?:\/\/(?:www\.)?\w+\.[\w]{2,3}(?:\/\S*)?`
 - Esta expresión regular busca URLs en un texto, incluyendo tanto http como https, y permitiendo rutas opcionales después del dominio.
- Validación de Números de Tarjeta de Crédito:
 - Expresión regular: `^(?:4[0-9]{12}(?:[0-9]{3})?|5[1-5][0-9]{14}|6(?:011|5[0-9][0-9])[0-9]{12}|3[47][0-9]{13}|3(?:0[0-5]|[68][0-9])[0-9]{11}|(?:2131|1800|35\d{3})\d{11})$`
 - Esta expresión regular valida varios tipos de números de tarjeta de crédito, incluyendo Visa, MasterCard, American Express, Discover, entre otros.

- Extracción de Etiquetas HTML:
 - Expresión regular: <[>]+>
 - Esta expresión regular encuentra todas las etiquetas HTML en un texto, incluyendo cualquier atributo que puedan tener.
- Búsqueda de Palabras Clave:
 - Expresión regular: \b(programación|desarrollo web|Python|JavaScript)\b
 - Esta expresión regular busca palabras clave específicas en un texto, como "programación", "desarrollo web", "Python" o "JavaScript".

Expresión Regular	Equivalencia
a.b	a x b a a b a b b a S b a # b ...
a..b	a x x b a a a b a b b b a 4 \$b ...
[abc]	a b c (cadenas de un caracter)
[aA]	a A (cadenas de un caracter)
[aA][bB]	ab Ab aB AB (cadenas de dos caracteres)
[0123456789]	0 1 2 3 4 5 6 7 8 9
[0-9]	0 1 2 3 4 5 6 7 8 9
[A-Za-z]	A B C ... Z a b c ... z
[0-9][0-9][0-9]	000 001 .. 009 010 .. 019 100 .. 999
[0-9]*	cadena_vacia 0 1 9 00 99 123 456 999 9999 ...
[0-9][0-9]*	0 1 9 00 99 123 456 999 9999 99999 999999 ...
^.*\$	cualquier línea completa

CONCLUSION

Las expresiones regulares son una herramienta poderosa y versátil para el procesamiento y manipulación de texto en una amplia variedad de aplicaciones, desde la programación hasta el análisis de datos y la administración de sistemas. Su capacidad para buscar y manipular texto basado en patrones definidos las convierte en una parte integral del kit de herramientas de cualquier desarrollador o profesional de la informática que trabaje con datos de texto.

