



UNIVERSIDAD DE  
GUADALAJARA

Red Universitaria e Institución Benemérita de Jalisco

**CUCEI**

## ¿Qué es un analizador léxico?

seminario de solución de problemas de traductores de  
lenguaje 2

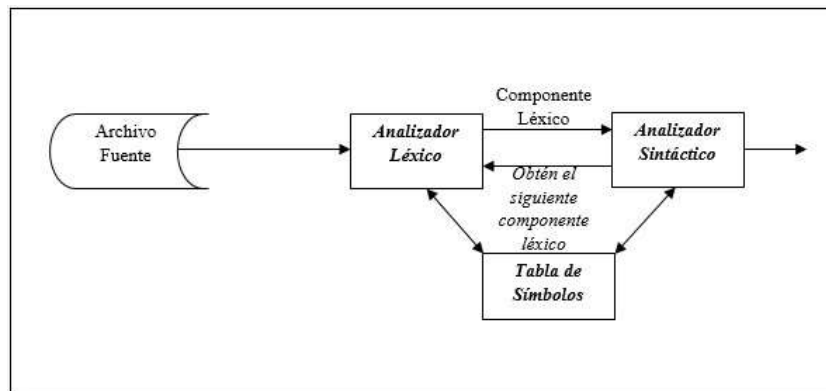
D02 Dr. Michel Emanuel López Franco

Jesús Fernando García Santos

INCO 213623791 Ingeniería en Computación

## ¿Qué es un analizador léxico?

El analizador léxico se encarga de buscar los componentes léxicos o palabras que componen el programa fuente, según unas reglas o patrones. La entrada del analizador léxico podemos definirla como una secuencia de caracteres. tiene que dividir la secuencia de caracteres en palabras con significado propio y después convertirlo a una secuencia de terminales desde el punto de vista del analizador sintáctico, que es la entrada del analizador sintáctico.



El analizador léxico reconoce las palabras en función de una gramática regular de manera que sus NO TERMINALES se convierten en los elementos de entrada de fases posteriores. En LEX, por ejemplo, esta gramática se expresa mediante expresiones regulares.

Es la primera fase de un compilador. Su principal función consiste en leer los caracteres de entrada y elaborar como salida una secuencia de componentes léxicos que utiliza el analizador sintáctico para hacer el análisis. Esta interacción, suele aplicarse convirtiendo al analizador léxico en una subrutina o corrutina del analizador sintáctico. Recibida la orden “Dame el siguiente componente léxico” del analizador sintáctico, el analizador léxico lee los caracteres de entrada hasta que pueda identificar el siguiente componente léxico. Otras funciones que realizan se listan a continuación:

Otras funciones que realiza:

- Eliminar los comentarios del programa.
- Eliminar espacios en blanco, tabuladores, retorno de carro, etc, y en general, todo
- aquello que carezca de significado según la sintaxis del lenguaje.
- Reconocer los identificadores de usuario, números, palabras reservadas del lenguaje, ...,
- y tratarlos correctamente con respecto a la tabla de símbolos (solo en los casos que debe

- de tratar con la tabla de símbolos).
- Llevar la cuenta del número de línea por la que va leyendo, por si se produce algún error,
- dar información sobre donde se ha producido.
- Avisar de errores léxicos. Por ejemplo, si @ no pertenece al lenguaje, avisar de un error.
- Puede hacer funciones de preprocesador

<i>Lexemas</i>	<i>Componente léxico</i>	<i>Patrón</i>
const	<b>const</b>	const
=	<b>relación</b>	< 0 <= 0 = 0 <> 0 > 0 >=
pi	<b>identificador</b>	letra seguida de letras o números
3.1416	<b>número</b>	cualquier literal numérica
"hola mundo"	<b>literal</b>	caracteres entre comillas

Entre las razones por las cuales es necesario separar el analizador léxico del sintáctico se pueden mencionar:

- Un diseño sencillo es quizás la consideración más importante. Separar el análisis léxico del análisis sintáctico a menudo permite simplificar una u otra de dichas fases. El analizador léxico nos permite simplificar el analizador sintáctico.
- Se mejora la eficiencia del compilador. Un analizador léxico independiente permite construir un procesador especializado y potencialmente más eficiente para esa función. Gran parte del tiempo se consume en leer el programa fuente y dividirlo en componentes léxicos. Con técnicas especializadas de manejo de buffers para la lectura de caracteres de entrada y procesamiento de componentes léxicos se puede mejorar significativamente el rendimiento de un compilador.
- Se mejora la portabilidad del compilador. Las peculiaridades del alfabeto de entrada y otras anomalías propias de los dispositivos pueden limitarse al analizador léxico. La representación de símbolos especiales o no estándares pueden ser aisladas en el analizador léxico.
- Para que se centre en el reconocimiento de componentes básicos complejos.