ACTIVIDAD 16 – ANALISIS COMPRENSIVO DE VIDEO



MARLE ALEXA RAMIREZ CABRAL #21-0890

EMILIO DE LOS SANTOS #21-0843

TEORÍA DE AUTÓMATAS Y COMPILADORES

PROF. RINA MARIA FAMILIA

UNIVERSIDAD IBEROAMERICANA

15/10/2022

**Función del analizador léxico por medio de autómatas**

Es la primera fase de un compilador. Su principal función consiste en leer la

secuencia de caracteres del programa fuente, carácter a carácter, y elaborar como

salida la secuencia de componentes léxicos que utiliza el analizador sintáctico. El

analizador sintáctico emite la orden al analizador léxico para que agrupe los

caracteres y forme unidades con significado propio llamados componentes léxicos

(tokens). El analizador léxico opera bajo petición del analizador sintáctico devolviendo un

componente léxico conforme el analizador sintáctico lo va necesitando para avanzar

en la gramática. Los componentes léxicos son los símbolos terminales de la

gramática.

Suele implementarse como una subrutina del analizador sintáctico. Cuando recibe la

orden "obtén el siguiente componente léxico", el analizador léxico lee los caracteres

de entrada, hasta identificar el siguiente componente léxico.

Un componente léxico es la secuencia lógica y coherente de caracteres relativo a una categoría identificadora, palabra reservada, literales (cadena/numérica), operador o carácter de puntuación, además de que un componente léxico puede tener uno o varios lexemas. El analizador léxico recoge información sobre los componentes léxicos en sus atributos asociados. Los tokens influyen en las decisiones del análisis sintáctico, y los atributos, en la traducción de los tokens. En la práctica los componentes léxicos suelen tener solo un atributo. Para efectos de diagnóstico, puede considerarse tanto el lexema para un identificador como el número de línea en el que se encontró por primera vez. Esta información puede ser almacenada en la tabla de símbolos para el identificador (estructura de datos).  
  
Las expresiones regulares son una notación importante para especificar patrones.

Cada patrón concuerda con una serie de cadenas que, de modo que las

expresiones regulares servirán como nombres para conjuntos de cadenas.

El término alfabeto denota cualquier conjunto finito de símbolos, por ejemplo, el

alfabeto binario {0,1}. Una cadena es una secuencia finita de símbolos de un

alfabeto y un lenguaje se refiere al conjunto de cadenas de un alfabeto fijo.

Hay veces en que el analizador léxico necesita analizar previamente varios

caracteres, además del lexema para un patrón, antes de poder anunciar una

concordancia. Se pueden emplear muchos esquemas de manejos de buffer, pero,

aquí tan solo se señalan algunos principios básicos.

Un reconocedor de lenguajes es un programa que toma como entrada una cadena x

e indica si dicha cadena pertenece a una frase del programa. Se compila una

expresión regular en un reconocedor construyendo un diagrama de transiciones

generalizado llamado autómata finito.

Un autómata finito puede ser determinista o no determinista, dependiendo del

número de transiciones para un mismo símbolo de entrada. Tanto los autómatas

finitos deterministas como los no deterministas pueden reconocer con precisión a

los conjuntos regulares.

**Algoritmo de Thompson**

es un método para transformar una expresión regular en un autómata finito no determinista equivalente (NFA). [2] Este NFA se puede utilizar para hacer coincidir cadenas con la expresión regular. Este algoritmo se le atribuye a Ken Thompson.  
  
Las expresiones regulares y los autómatas finitos no deterministas son dos representaciones de lenguajes formales. Por ejemplo, las utilidades de procesamiento de texto utilizan expresiones regulares para describir patrones de búsqueda avanzada, pero las NFA son más adecuadas para su ejecución en una computadora. Por lo tanto, este algoritmo es de interés práctico, ya que puede compilar expresiones regulares en NFA. Desde un punto de vista teórico, este algoritmo es parte de la prueba de que ambos aceptan exactamente los mismos lenguajes, es decir, los lenguajes regulares. Un NFA puede hacerse determinista mediante la construcción del conjunto de potencias y luego minimizarse para obtener un autómata óptimo correspondiente a la expresión regular dada. Sin embargo, una NFA también se puede interpretar directamente.

Para decidir si dos expresiones regulares dadas describen el mismo lenguaje, cada una se puede convertir en un autómata finito determinista mínimo equivalente mediante la construcción de Thompson, la construcción del conjunto de potencias y la minimización de DFA . Si, y solo si, los autómatas resultantes aceptan el cambio de nombre de los estados, los lenguajes de las expresiones regulares están de acuerdo.

