

Compiladores
Práctica 1
Construcción de una Analizador Léxico
Nombre _____
Grupo _____
Calificación _____

José Sánchez Juárez

23 de febrero de 2016

1. Objetivo

Diseño e implementación de un analizador léxico.

2. Teoría

Los autómatas se definen de manera formal como una quintúpla: La cual esta formada de los elementos tales como, el estado inicial que se representa como s_0 , el conjunto de estados que se representa por s donde se cumple $s_0 \in s$, el conjunto de transiciones que se determina como δ , el conjunto de estados finales o de aceptación que se representan como s_A y el alfabeto que contiene los caracteres que acepta el autómata y que se representa por medio de Σ . La representación formal del autómata es de la siguiente forma: $A(\Sigma, \delta, s, s_0, s_A)$. Si el conjunto de estados es un conjunto finito entonces el autómata es finito y se abrevia como AF (Autómata Finito).

El autómata es una herramienta que sirve para reconocer series de caracteres que tienen un estructura regular como por ejemplo una serie de a 's o una serie de b 's. O la serie de dígitos para representar un número entero. O los caracteres que representan un identificador. Cuando las estructuras regulares se hacen más complejas el autómata que las reconoce se hace también más complejo.

Los autómatas se clasifican en autómatas finitos determinísticos (AFD) los cuales no aceptan la transición de caracter vacío ni dos transiciones con el mismo caracter a estados diferentes y los autómatas finitos no determinísticos (AFN) si aceptan transiciones ϵ y transiciones a dos estados diferentes con el mismo símbolo.

Los analizadores léxicos reconocen lenguajes regulares, los cuales se representan por medio de expresiones regulares. Así que el paso lógico es convertir la expresión regular a autómata AFN. La conversión de expresión regular a AFN se hace por medio de la construcción de Thompson.

Por lo que el siguiente paso es convertir los AFN a AFD, aplicando el algoritmo de subconjuntos. El algoritmo de subconjuntos se hace aplicando al estado inicial del AFN la cerradura- ϵ y los siguientes estados se construyen aplicando la cerradura- ϵ a los movimientos de todos y cada uno de los estado que resulten con transición de todos y cada uno de los símbolos del alfabeto. Al aplicar el algoritmo de subconjuntos se obtiene el AFD no mínimo (AFD_{NM}).

Todavía se debe aplicar el algoritmo de minimización por consistencia para generar el AFD mínimo (AFD_M). El AFD_M se debe convertir en tabla, la cual se representa como $\delta[Estado, Car]$, donde Estado y Car son variables, la primera que representa los estados del AFD_M y Car que representa la lectura del código fuente, caracter por caracter.

Los analizadores léxicos se contruyen de tres formas: por manejo de tabla, de manera artesanal y por codificación directa. En esta práctica se implementará el analizador léxico por codificación directa. De cualquier manera para comenzar a construirlo se debe presentar la expresión regular que representa las palabras a reconocer.

El siguiente pseudocódigo muestra como se programa un analizador léxico por codificación directa:

Algorithm .1: Escaner

```

1 begin
2  $Estado \leftarrow s_0$  ;
3 while  $Estado \neq s_e$  do
4    $Car \leftarrow SigCar()$  ;
5    $Estado \leftarrow \delta[Estado, Car]$  ;
6   if  $Estado == s_A$  then
7     break ;
8 if  $Estado == s_e$  then
9   EsEnPan(Palabra No Aceptada) ;
10 else
11   EsEnPan(Palabra Aceptada) ;
12 end
```

3. Actividades

Realice las siguientes actividades:

1. Transformar las ER's: $ab \mid (ab) * c$ y $(a \mid b) * a(a \mid b)(a \mid b)$ en AFN's aplicando la construcción de Thompson.

2. Transformar el AFN's en $AFD'_{NM}s$ aplicando el algoritmo de subconjuntos.
3. Transformar el $AFD'_{NM}s$ en $AFD'_M s$ aplicando el algoritmo de minimización por consistencia.
4. Programar en cualquier lenguaje de alto nivel el analizador léxico que reconozca las palabras que aceptan las ER's del primer punto.

4. Resultados

Entregar en el laboratorio:

1. Todo el proceso de cálculo de los $AFD'_M s$.
2. Los programas de los escaners que reconozcan las palabras que aceptan las ER's del punto uno de actividades.

Entregar todo en el laboratorio, en la fecha que solicite el Profesor.