Sintaxis y semantica de los lenguajes

Ultima modificacion: June 3, 2021

Simbolo Es una entidad indivisible. Pueden estar formados por varias letras.

Alfabeto Conjunto de simbolos finito y no vacio. Se designa con Σ o una letra mayuscula y un subindice A_1

Cadenas de caracteres Con los simbolos de un alfabeto se pueden formar secuencias, tambien llamadas cadenas o palabras. Una cadena es una secuencia finita de simbolos de un determinado alfabeto.
Una cadena puede ser vacia, es decir no contener ningun simbolo. Por lo tanto pertenece a todos los alfabetos. Se designa con ϵ.

Longitud de una cadena Numero de simbolos de una cadena. Se designa de la siguiente forma $|w|$.

Concatencion de cadenas Sean x e y dos cadenas, xy es su concatenacion. Si $x = a_1a_2...a_i$ y $y = b_1b_2...b_j$ entonces $xy = a_1a_2...a_ib_1b_2...b_j$

- Es una operacion binaria definida sobre el conjunto universo del alfabeto
- El conjunto universo es cerrado bajo la concatenacion
- Es asociativa
- Elemento identidad = Cadena vacia (ϵ)
- No conmutativa
- $\forall$ w,v del mismo universo $|wv| = |w| + |v|$
- Concatenacion de k copias de $w : w^k = ww...w\}k$ veces

Potencia de un alfabeto Es el conjunto de cadenas de longitud k tales que todos los simbolos que las forman pertenecen al alfabeto.

- Potencia 0 = ϵ,
- Potencia n = $|\Sigma|^n$

Universo Conjunto de todas las palabras que se pueden definir sobre un alfabeto Es la union de las infinitas potencias del alfabeto.

Nota: Cadena vacia (ϵ) $\neq$ Conjunto formado por una cadena vacia ($\{\epsilon\}$) $\neq$ Conjunto vacio ($\emptyset$).

Lenguaje formal Conjunto finito o infinito de cadenas de longitud finita, generadas a partir de un conjunto finito de simbolos llamado alfabeto.

- Al ser conjuntos se pueden aplicar operaciones de conjuntos.
- Sea Σ un alfabeto y L un subconjunto del universo Σ^*, entonces L es un lenguaje de Σ .
- Se puede definir a un lenguaje de la siguiente forma: $L = \{a^n b^n, para\, n > 1\}$
- El lenguaje vacio esta incluido en todos los alfabetos
- El lenguaje formado por ϵ esta incluido en todos los alfabetos
- **Concatenacion** $L_1 L_2 = \{\omega/\omega = \alpha\beta, \alpha \in L_1, \beta \in L_2\}$

Potencia de un lenguaje
$L^n = \{\epsilon\}, si\, n = 0;$
$= L.L^{n-1}, si\, n > 0$

Clausura de Kleene
Union de todas las potencias de un lenguaje.
$L^* = \bigcup_{i=0}^{\infty} L^i = L^0 \cup L^1 \cup L^i \cup ... \cup L^{\infty}$

Tipos de Gramaticas
Una gramatica se define como $G = \{N, T, S, P\}$
donde N es el conjunto de simbolos no terminales, T el de los terminales, $S \in N$ el simbolo inicial (el mas general), y P son las reglas de asignacion para formar palabras de la forma $\phi A \rho \longrightarrow \phi \omega \rho$ donde ϕ, ρ son el contexto.

- Tipo 0: Libres de restriccion.
- Tipo 1: Dependientes del contexto, $\omega \in (N \cup T)^+$ (No incluye a ϵ).
- Tipo 2: Independientes del contexto, $\phi = \rho = \epsilon$.
- Tipo 3: Regulares, lineales a la derecha o izquierda. (completar)
- Todos los tipos incluyen las restricciones de los tipos que le preceden.

Lineal a derecha e izquierda respectivamente:
$Nu \longrightarrow 0$
$Nu \longrightarrow 1$
$Nu \longrightarrow B$
$B \longrightarrow 0$
$B \longrightarrow 1$
$B \longrightarrow B0$ -por izquierda
$B \longrightarrow 0B$ -por derecha

Automatas finitos Son un modelo util para construir analizadores lexicos. Son sistemas o componentes con un numero finito de estados, un estado recuerda la parte significativa de la historia del sistema. Sus entradas son palabras del lenguaje que representa.

Un automata tiene tres tipos de estados:

- Estado inicial, permite empezar la ejecucion
- Estados finales, permite realizar la salida de aceptacion
- Estados intermedios

Automatas deterministas

No puede estar en mas de un estado simultaneamente. Se denota con la quintupla $A = \{\Sigma, Q, f, q_0, F\}$ donde:

- Σ es el alfabeto de simbolos de entrada
- Q es el conjunto finito de estados
- $q_0 \in Q$ es el estado inicial
- f es la funcion de transicion, toma como argumentos un estado de Q y un simbolo de entrada de Σ y devuelve un estado de Q
- $F \subseteq Q$ es el conjunto de los estados finales

La funcion de transicion "f" es una funcion binaria que toma un estado

$$\in Q$$

y un simbolo

$$\in \Sigma$$

y produce un nuevo estado:

$$f(q_i, a) = q_j$$

La funcion extendida

$$\hat{f}$$

toma un estado y una cadena formada por simbolos del alfabeto del automata y produce un estado. Esta definida de la siguiente forma:

$$\hat{f}(q, \epsilon) = q$$

$$\hat{f}(q, w) = f(\hat{f}(q, x), a)$$

Donde a es el ultimo elemento de w y x el resto del cuerpo de w.

Un AFD acepta o no una secuencia de simbolos de entrada segun si el automata se encuentra en un estado de aceptacion. Un AFD puede representarse tanto como una tabla de transiciones como un grafo o diagrama.

Automatas no deterministas

Puede estar en varios estados al mismo tiempo.
Se definen de forma similar a los AFD pero su funcion de transicion, que toma un estado y un simbolo de entrada, produce un conjunto de estados como salida en vez de un unico estado como en los AFD. Este conjunto puede ser vacio.
Por lo tanto tambien posee una funcion de transicion extendida diferente: dicha funcion producira tambien un conjunto de estados que se obtiene de la siguiente forma:
Sea $\omega = xa$
y sea $\hat{f}_N(q, x) = \{p_1, p_2, ..., p_k\}$

$$\bigcup_{i=1}^{k} f_N(p_i, a) = \{r_1, r_2, ..., r_m\} = \hat{f}_N(q, \omega)$$

Automatas finitos con transiciones ϵ

Permite transiciones para la cadena vacia.
Equivalencia con un AFD:
$AFN - \epsilon\ E = (Q_E, \Sigma, f_E, q_0, F_E)$, $AFD\ D = (Q_D, \Sigma, f_D, q_D, F_D)$
tal que $L(D) = L(E)$

- $Q_D eselconjuntodesubconjuntosde Q_E$. Aunque solo los subconjuntos cerrados en ϵ seran accesibles por el AFD.
- $q_D = CLAUSURA_\epsilon(q_0)$
- F_D incluye todos los conjuntos de estados de E que incluyen al menos un estado de aceptacion de E.
- Para cada subconjunto de Q_D y para cada simbolo de entrada a en Σ,
  Sea $S = \{p_1, p_2, ..., p_k\}$
  y sea $\bigcup_{i=1}^{k} f_E(p_i, a) = r_1, r_2, ..., r_m$
  $f_D(S, a) = \bigcup_{j=1}^{m} CLAUSURA_\epsilon(r_j)$