I. Fundamentos de la Teoría de Lenguages

1. Definiciones Básicas y Notaciones

1.1 La noción de alfabeto : el conj. Σ

Es un conjunto finito, no vacío. Los elementos de Σ son símbolos indivisibles.

Ejemplo: $\Sigma = \{\emptyset, 1\}$
$\Sigma = \{a, b\}$

Una <u>palabra</u> (o <u>cadena</u>) sobre un alfabeto Σ es una secuencia en Σ de longitud finita.

Ejemplo: $\Sigma = \{\emptyset, 1\}$, $w = \emptyset\emptyset 1 \emptyset\emptyset$

Formalmente, una cadena es de la forma $w = a_1 \dots a_n$, $n \geq \emptyset$, $a_i \in \Sigma$, $1 \leq i \leq n$, donde n es la longitud de w, utilizando la notación $|w| = n$. La palabra vacía se denotará por ε.

1.1.1 Σ^* como monoide.

$$\Sigma^* = \{ a_1 \dots a_n ,\ a_i \in \Sigma,\ 1 \leq i \leq n \} \cup \{\varepsilon\}$$

Sobre Σ^* podemos definir la operación <u>concatenación</u> en la forma siguiente:

$x, y \in \Sigma^* \Rightarrow$: concatenacion$(x,y) := xy$, esto es:

$$\begin{aligned} \Sigma^* \times \Sigma^* &\longrightarrow \Sigma^* \\ (x, y) &\longmapsto xy \end{aligned}$$

Ejercicio: Ver que está bien definida.

* proposición: Sea Σ un alfabeto, entonces la concatenación verifica las propiedades:

a) Asociativa, $x(yz) = (xy)z$, $\forall x,y,z \in \Sigma^*$

b) ε. neutro, $\varepsilon x = x\varepsilon = x$, $\forall x \in \Sigma^*$

c) Σ^* es un monoide con la concatenación

d) Σ^* es ~~monótona~~ por la derecha e izquierda:

i) $zx = zy \Rightarrow x = y$, $\forall x,y,z \in \Sigma^*$

ii) $xz = yz \Rightarrow x = y$, $\forall x,y,z \in \Sigma^*$

e) $|xy| = |x| + |y|$, $\forall x,y \in \Sigma^*$

demostración

a) b) Trivial por def. de concatenación

c) Trivial a partir de a) b)

d) Por inducción en la longitud de z

i)

$|z|=0 \Rightarrow z = \varepsilon \Rightarrow \left.\begin{cases} zx = x \\ zy = y \end{cases} \atop \text{por hipotesis } zx = zy \right\} \Rightarrow x = y$

$|z|=1$ Trivial

$|z| \leq n$ asumimos por inducción que $zx = zy$, $\forall z$, $|z| \leq n$

* $|z| = n+1$ t.q.d $x = y$ donde $\begin{cases} z = a_1 \dots a_n a_{n+1}, \text{ por tanto:} \\ zx = zy \end{cases}$

$a_1(a_2 \dots a_{n+1})x = a_1(a_2 \dots a_{n+1})y \Rightarrow$ (caso $|z|=1$) $\Rightarrow$

$\Rightarrow a_2 \dots a_{n+1}x = a_2 \dots a_{n+1}y \Rightarrow$ (inducción) $\Rightarrow$

$\Rightarrow x = y$.

ii) Idem.

e) Por inducción en la longitud de x (trivial):

$|x|=\emptyset \Rightarrow x=\varepsilon \Rightarrow \left.\begin{array}{l}|xy|=|y| \\ |x|+|y|=\emptyset+|y|\end{array}\right\} \Rightarrow |xy|=|x|+|y|$

$|x|=1$ trivial

$|x|\leq n$ Supuesto cierto

$|x|=n+1$ Entonces $x=a_1a_2 \cdots a_n a_{n+1}$

$$|xy|=|a_1a_2 \cdots a_n a_{n+1}y| \underset{\text{asociativa}}{=} |a_1(a_2\cdots a_n a_{n+1})y| = (\text{inducción}) =$$

$$= |a_1| + |(a_2 \cdots a_n a_{n+1})y| = (\text{inducción}) = |a_1| +$$

$$+ |a_2 \cdots a_n a_{n+1}| + |y| = (\text{inducción}) = |a_1(a_2 \cdots a_{n+1})| +$$

$$+ |y| = |a_1 a_2 \cdots a_n a_{n+1}| + |y| = |x| + |y|$$

demostrado

1.1.2 Semigrupos en Σ^*. Cierre de kleene.

Podemos ahora extender el concepto de concatenación a conjuntos de cadenas de un alfabeto Σ, en la forma:

Sea $X,Y \subseteq \Sigma^*$, $XY := \{xy / x\in X, y \in Y\}$

También definimos dado $X \subseteq \Sigma^*$, $\begin{cases} X^0 := \{\varepsilon\} \\ X^{i+1} := X^i X, \ \forall i \geq \emptyset \end{cases}$

y el cierre de kleene $X^* := \bigcup_{i\geq 0} X^i$ o cierre reflexivo-transitivo

cierre transitivo $X^+ := X^*X = \bigcup_{i\geq 1} X^i$

proposición: Sea Σ un alfabeto y sea $X \subseteq \Sigma^*$, entonces:

a) X^+ es un semigrupo con la concatenación

b) X^* es un monoide con la concatenación

demostración (ejercicio) (semigrupo = op. interna asociativa)

1.2 Gramáticas

1.2.1 Reglas, terminales, variables, axiomas.

Una gramática es un 4-tupla $\mathcal{G} = (N, \Sigma, P, S)$ donde

a) N es un conj. finito de símbolos no terminales (también llamados variables o categorías sintácticas)

b) Σ es un conjunto finito de <u>símbolos terminales</u>, disjunto con N (representará el alfabeto de la gramática)

c) P es un subconj. finito de $(N \cup \Sigma)^* N (N \cup \Sigma)^* \times (N \cup \Sigma)^*$

Un elemento (α, β) en P se escribirá normalmente en la forma $\alpha \to \beta$, estructura que recibe el nombre de <u>regla</u> o <u>producción</u>

d) S es un símbolo distinguido en N, que recibe el nombre de <u>axioma</u> o <u>símbolo inicial</u> de la gramática.

<u>NOTA</u>: A veces, una gramática se denota $\mathcal{G} = (V, \Sigma, P, S)$ donde V es el conjunto de todos los símbolos presentes en la misma, esto es, $V = N \cup \Sigma$.

<u>Ejemplo</u>: $\mathcal{G} = (\{S\}, \{+, *, a,), (\}, P, S)$, donde P viene dado por

$$
\begin{array}{lll}
(1) & S \to & S+S \\
(2) & & \mid S*S \\
(3) & & \mid (S) \\
(4) & & \mid a
\end{array}
$$

<u>NOTACION</u>: En general, y para unificar criterios, utilizaremos en adelante la siguiente notación:

$V := N \cup \Sigma$ el conj. total de símbolos en $\mathcal{G}$.

$a, b, c \ldots \in \Sigma$ los símbolos terminales en $\mathcal{G}$

$A, B, C \ldots \in N$ las variables en $\mathcal{G}$

$X, Y, Z \ldots \in V$ símbolos arbitrarios en $\mathcal{G}$

$\ldots u, v, w, x, y, z \in \Sigma^*$ cadenas determinales

$u_{1..n}$ una cadena $u \in \Sigma^*$ de longitud n

$u_{i:j}$ la subcadena de $u \in \Sigma^*$ que va de la palabra en la posición i a la palabra en la posición j, ambas inclusive.

u_i el caracter de $u \in \Sigma^*$, en la posición i.

$\alpha, \beta, \ldots \in V^*$ cadenas arbitrarias

1.2.2 Derivaciones. Lenguaje engendrado por una gramática.

Sea $\mathcal{G} = (N, \Sigma, P, S)$ una gramática, entonces decimos que $\alpha\beta\gamma$ <u>deriva directamente</u> $\alpha\delta\gamma$ si $\beta \rightarrow \delta \in P$. Utilizaremos la notación $\alpha\beta\gamma \Rightarrow \alpha\delta\gamma$.

Sea $\mathcal{G} = (N, \Sigma, P, S)$ una gramática, diremos que $\alpha\beta\gamma$ <u>deriva indirectamente</u> $\alpha\delta\gamma$ si se verifica uno de los dos casos siguientes:

a) $\beta \Rightarrow \delta_1 \Rightarrow \delta_2 \Rightarrow \cdots \Rightarrow \delta_n \Rightarrow \delta$, que notaremos por $\alpha\beta\gamma \overset{+}{\Rightarrow} \alpha\delta\gamma$

b) $\beta = \delta$ ó $\beta \overset{+}{\Rightarrow} \delta$, en este caso notaremos $\alpha\beta\gamma \overset{*}{\Rightarrow} \alpha\delta\gamma$ ó $\alpha\beta\gamma \overset{k}{\Rightarrow} \alpha\delta\gamma$ cuando queramos expresar un número k de pasos de derivación directa.

<u>Definición</u>: Sea $\mathcal{G} = (N, \Sigma, P, S)$, el <u>lenguaje engendrado por</u> $\mathcal{G}$ es $\mathcal{L}(\mathcal{G}) := \{\alpha \in \Sigma^* / S \overset{*}{\Rightarrow} \alpha\}$