

ST0270 Lenguajes Formales y Compiladores

Teoría de los lenguajes formales: Definiciones y operaciones básicas

Andrés Sicard Ramírez

Universidad EAFIT

Semestre 2024-1

(Última actualización: 25 de julio de 2025)

Preliminares

Referencias

La referencia principal para estas diapositivas es (Kozen 2012, Lectura 2).

Convenciones

- (i) Los números asignados a los definiciones, ejemplos, ejercicios, páginas y teoremas en estas diapositivas corresponden a los números asignados en (Kozen 2012).
- (ii) Los números naturales incluyen el cero, es decir, $\mathbb{N} = \{0, 1, 2, \dots\}$.
- (iii) El conjunto potencia de un conjunto A es denotado 2^A .
- (iv) Los términos «definición inductiva» y «definición recursiva» se usan como sinónimos.

Agradecimientos

Agradezco a mi colega Sergio Steven Ramírez Rico por señalarme algunas correcciones en una versión anterior de estas diapositivas.

Problemas de decisión

Definición

Un **problema de decisión** sobre un conjunto A (el dominio del problema) es una función

$$f : A \rightarrow \{0, 1\}.$$

Observación

Abstracción: Usualmente solo será necesario considerar los problemas de decisión.

Definiciones básicas

Definiciones 2.1

- Un **alfabeto** es un conjunto **finito**, no vacío, de símbolos indivisibles.

Definiciones básicas

Definiciones 2.1

- Un **alfabeto** es un conjunto **finito**, no vacío, de símbolos indivisibles.
- Una **palabra** (*string*) es una secuencia **finita** de símbolos de un alfabeto. La palabra vacía es denotada ϵ .

Definiciones básicas

Definiciones 2.1

- Un **alfabeto** es un conjunto **finito**, no vacío, de símbolos indivisibles.
- Una **palabra** (*string*) es una secuencia **finita** de símbolos de un alfabeto. La palabra vacía es denotada ϵ .
- La palabra vacía, denotada ϵ , es la palabra con cero ocurrencias de símbolos.

Definiciones básicas

Definiciones 2.1

- Un **alfabeto** es un conjunto **finito**, no vacío, de símbolos indivisibles.
- Una **palabra** (*string*) es una secuencia **finita** de símbolos de un alfabeto. La palabra vacía es denotada ϵ .
- La palabra vacía, denotada ϵ , es la palabra con cero ocurrencias de símbolos.

Convenciones

- Los alfabetos se denotarán por Σ y Γ .
- Los símbolos se denotarán por las primeras letras minúsculas $a, b, c, \dots$
- Las palabras se denotarán por las últimas letras minúsculas $w, x, y, z, \dots$

Definiciones básicas

Definición

El **conjunto de todas las palabras** (incluyendo la palabra vacía) sobre un alfabeto Σ es denotado Σ^* . Este conjunto puede ser inductivamente definido por:

(i) Definición tradicional:

- a) Paso base: $\epsilon \in \Sigma^*$,
- b) Paso inductivo: Si $x \in \Sigma^*$ y $a \in \Sigma$ entonces $xa \in \Sigma^*$.

Definiciones básicas

Definición

El **conjunto de todas las palabras** (incluyendo la palabra vacía) sobre un alfabeto Σ es denotado Σ^* . Este conjunto puede ser inductivamente definido por:

(i) Definición tradicional:

a) Paso base: $\epsilon \in \Sigma^*$,

b) Paso inductivo: Si $x \in \Sigma^*$ y $a \in \Sigma$ entonces $xa \in \Sigma^*$.

(ii) Definición usando reglas de inferencia:

$$\frac{}{\epsilon \in \Sigma^*},$$

$$\frac{x \in \Sigma^* \quad a \in \Sigma}{xa \in \Sigma^*}.$$

Definiciones básicas

Definición

Un **lenguaje** sobre un alfabeto Σ es **cualquier** subconjunto de Σ^* .

Definiciones básicas

Definición

Un **lenguaje** sobre un alfabeto Σ es **cualquier** subconjunto de Σ^* .

Observación

Kozen (2012) no define el término «lenguaje» explícitamente. En su lugar el texto habla de conjuntos.

Definiciones básicas

Definición

Un **lenguaje** sobre un alfabeto Σ es **cualquier** subconjunto de Σ^* .

Observación

Kozen (2012) no define el término «lenguaje» explícitamente. En su lugar el texto habla de conjuntos.

Observación

Note que $\emptyset \neq \{\epsilon\} \neq \epsilon$.

Operaciones sobre palabras y símbolos

Definición

La **longitud** de una palabra x sobre un alfabeto Σ , denotada $|x|$, es el número de símbolos en x . Esta función es definida recursivamente por:

$$|\cdot| : \Sigma^* \rightarrow \mathbb{N}$$

$$|\epsilon| \stackrel{\text{def}}{=} 0,$$

$$|xa| \stackrel{\text{def}}{=} 1 + |x|.$$

Operaciones sobre palabras y símbolos

Definición

Sea a un símbolo de un alfabeto Σ . La **potencias de a** , para $n \geq 0$, denotadas a^n , es la palabra formada por n repeticiones del símbolo a . Esta operación es definida recursivamente por:

$$(\cdot)^{(\cdot)} : \Sigma \times \mathbb{N} \rightarrow \Sigma^*$$

$$a^0 \stackrel{\text{def}}{=} \epsilon,$$

$$a^{n+1} \stackrel{\text{def}}{=} a^n a.$$

Operaciones sobre palabras y símbolos

Definición

Sean x y y dos palabras sobre un alfabeto Σ . La **concatenación** de x y y , denotada xy , es la palabra que se obtiene al yuxtaponer y después de x . Esta operación es definida por

$$xy \stackrel{\text{def}}{=} \text{concat}(x, y),$$

donde $\text{concat} : \Sigma^* \times \Sigma^* \rightarrow \Sigma^*$ es definida recursivamente por:

$$\begin{aligned}\text{concat}(x, \epsilon) &\stackrel{\text{def}}{=} x, \\ \text{concat}(x, ya) &\stackrel{\text{def}}{=} \text{concat}(x, y)a.\end{aligned}$$

Es decir, sean $x = a_1a_2 \dots a_n$ y $y = b_1b_2 \dots b_n$ dos palabras, entonces

$$xy = a_1a_2 \dots a_nb_1b_2 \dots b_n.$$

Operaciones sobre palabras y símbolos

Algunas propiedades de la concatenación

Sean x , y y z palabras.

- (i) Concatenación es asociativa, es decir, $x(yz) = (xy)z$.
- (ii) La palabra vacía es el elemento neutro para la concatenación, es decir, $x\epsilon = \epsilon x = x$.
- (iii) Concatenación no es conmutativa, es decir $xy \neq yx$.
- (iv) $|xy| = |x| + |y|$.

Operaciones sobre lenguajes

Definición

Sean L , L_1 y L_2 lenguajes sobre un alfabeto Σ .

(i) **Unión** de lenguajes:

$$\cup : 2^{\Sigma^*} \times 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L_1 \cup L_2 \stackrel{\text{def}}{=} \{ x \mid x \in L_1 \text{ o } x \in L_2 \}.$$

Operaciones sobre lenguajes

Definición

Sean L , L_1 y L_2 lenguajes sobre un alfabeto Σ .

(i) **Unión** de lenguajes:

$$\cup : 2^{\Sigma^*} \times 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L_1 \cup L_2 \stackrel{\text{def}}{=} \{ x \mid x \in L_1 \text{ o } x \in L_2 \}.$$

(ii) **Intersección** de lenguajes:

$$\cap : 2^{\Sigma^*} \times 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L_1 \cap L_2 \stackrel{\text{def}}{=} \{ x \mid x \in L_1 \text{ y } x \in L_2 \}.$$

Operaciones sobre lenguajes

Definición

Sean L , L_1 y L_2 lenguajes sobre un alfabeto Σ .

(i) **Unión** de lenguajes:

$$\cup : 2^{\Sigma^*} \times 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L_1 \cup L_2 \stackrel{\text{def}}{=} \{x \mid x \in L_1 \text{ o } x \in L_2\}.$$

(iii) **Complemento** de un lenguaje:

$$\sim : 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$\sim L \stackrel{\text{def}}{=} \{x \in \Sigma^* \mid x \notin L\}.$$

(ii) **Intersección** de lenguajes:

$$\cap : 2^{\Sigma^*} \times 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L_1 \cap L_2 \stackrel{\text{def}}{=} \{x \mid x \in L_1 \text{ y } x \in L_2\}.$$

Operaciones sobre lenguajes

Definición

Sean L , L_1 y L_2 lenguajes sobre un alfabeto Σ .

(i) **Unión** de lenguajes:

$$\cup : 2^{\Sigma^*} \times 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L_1 \cup L_2 \stackrel{\text{def}}{=} \{x \mid x \in L_1 \text{ o } x \in L_2\}.$$

(iii) **Complemento** de un lenguaje:

$$\sim : 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$\sim L \stackrel{\text{def}}{=} \{x \in \Sigma^* \mid x \notin L\}.$$

(ii) **Intersección** de lenguajes:

$$\cap : 2^{\Sigma^*} \times 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L_1 \cap L_2 \stackrel{\text{def}}{=} \{x \mid x \in L_1 \text{ y } x \in L_2\}.$$

(iv) **Concatenación** de lenguajes:

$$(\cdot)(\cdot) : 2^{\Sigma^*} \times 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L_1 L_2 \stackrel{\text{def}}{=} \{xy \mid x \in L_1 \text{ y } y \in L_2\}.$$

(continúa en la próxima diapositiva)

Operaciones sobre lenguajes

Definición (continuación)

(v) **Potencias** de un lenguaje, para $n \geq 0$:

$$(\cdot)^{(\cdot)} : 2^{\Sigma^*} \times \mathbb{N} \rightarrow 2^{\Sigma^*}$$

$$L^0 \stackrel{\text{def}}{=} \{\epsilon\},$$

$$L^{n+1} \stackrel{\text{def}}{=} LL^n.$$

Operaciones sobre lenguajes

Definición (continuación)

(v) **Potencias** de un lenguaje, para $n \geq 0$:

$$(\cdot)^{(\cdot)} : 2^{\Sigma^*} \times \mathbb{N} \rightarrow 2^{\Sigma^*}$$

$$L^0 \stackrel{\text{def}}{=} \{\epsilon\},$$

$$L^{n+1} \stackrel{\text{def}}{=} LL^n.$$

(vi) **La clausura de Kleene** de un lenguaje es la unión de la potencias finitas del lenguaje:

$$(\cdot)^* : 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L^* \stackrel{\text{def}}{=} \bigcup_{n \geq 0} L^n.$$

Operaciones sobre lenguajes

Definición (continuación)

- (v) **Potencias** de un lenguaje, para $n \geq 0$: (vii) **La clausura positiva de Kleene** de un lenguaje es la unión de las potencias positivas del lenguaje:

$$(\cdot)^{(\cdot)} : 2^{\Sigma^*} \times \mathbb{N} \rightarrow 2^{\Sigma^*}$$

$$L^0 \stackrel{\text{def}}{=} \{\epsilon\},$$

$$L^{n+1} \stackrel{\text{def}}{=} LL^n.$$

$$(\cdot)^+ : 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L^+ \stackrel{\text{def}}{=} LL^* = \bigcup_{n \geq 1} L^n.$$

- (vi) **La clausura de Kleene** de un lenguaje es la unión de la potencias finitas del lenguaje:

$$(\cdot)^* : 2^{\Sigma^*} \rightarrow 2^{\Sigma^*}$$

$$L^* \stackrel{\text{def}}{=} \bigcup_{n \geq 0} L^n.$$

Operaciones sobre lenguajes

Algunas propiedades de las operaciones entre lenguajes

- (i) Desde que los lenguajes son conjuntos, las operaciones de unión e intersección sobre lenguajes son asociativas, conmutativas y cada una distribuye sobre la otra. Además, satisfacen las leyes de De Morgan y el lenguaje $\emptyset$ es la identidad para la unión.

Operaciones sobre lenguajes

Algunas propiedades de las operaciones entre lenguajes

- (i) Desde que los lenguajes son conjuntos, las operaciones de unión e intersección sobre lenguajes son asociativas, conmutativas y cada una distribuye sobre la otra. Además, satisfacen las leyes de De Morgan y el lenguaje $\emptyset$ es la identidad para la unión.
- (ii) El lenguaje $\{\epsilon\}$ es la identidad para la concatenación, es decir,

$$L\{\epsilon\} = \{\epsilon\}L = L.$$

Operaciones sobre lenguajes

Algunas propiedades de las operaciones entre lenguajes

- (i) Desde que los lenguajes son conjuntos, las operaciones de unión e intersección sobre lenguajes son asociativas, conmutativas y cada una distribuye sobre la otra. Además, satisfacen las leyes de De Morgan y el lenguaje $\emptyset$ es la identidad para la unión.
- (ii) El lenguaje $\{\epsilon\}$ es la identidad para la concatenación, es decir,

$$L\{\epsilon\} = \{\epsilon\}L = L.$$

- (iii) El lenguaje $\emptyset$ es el **aniquilador** para la concatenación, es decir,

$$L\emptyset = \emptyset L = \emptyset.$$

(continua en la próxima diapositiva)

Operaciones sobre lenguajes

Algunas propiedades de las operaciones entre lenguajes (continuación)

(iv) La clausura de Kleene satisface las siguiente propiedades:

$$L^*L^* = L^*,$$

$$L^{**} = L^*,$$

$$L^* = L^+ \cup \{\epsilon\},$$

$$\emptyset^* = \{\epsilon\}.$$

Referencias



Dexter C. Kozen [1997] (2012). *Automata and Computability*. Third printing. Undergraduate Texts in Computer Science. Springer. DOI: [10.1007/978-1-4612-1844-9](https://doi.org/10.1007/978-1-4612-1844-9) (vid. págs. 2, 10-12).