Ejercicios de Teoría de Lenguajes Formales

Juan Alejandro Salgado Arcila

10 de agosto de 2025

I. Alfabetos, Palabras y Operaciones Básicas

Enunciado: Dado el alfabeto $\Sigma = \{a, b, c\}$ y las palabras $w = \texttt{abac}$ y $v = \texttt{ca}$.

(1) **Determina cuáles de las siguientes palabras son válidas sobre el alfabeto Σ:** $\texttt{baca}$, $\texttt{cada}$, ε, $\texttt{abcde}$.

Respuesta:

$\texttt{baca}$ es válida, porque todas sus letras $(b, a, c, a) \in \Sigma$.

$\texttt{cada}$ no es válida, porque contiene la letra $d \notin \Sigma$.

ε (palabra vacía) es válida sobre cualquier alfabeto, así que pertenece a Σ^*.

$\texttt{abcde}$ no es válida, porque contiene $d, e \notin \Sigma$.

(2) **Calcula la longitud de w y v. Es decir, encuentra $|w|$ y $|v|$.**

Respuesta:

$$w = \texttt{abac} \implies |w| = 4, \qquad v = \texttt{ca} \implies |v| = 2.$$

(3) **Realiza las siguientes concatenaciones:** wv, vw.

Respuesta:

$$wv = \texttt{abac} \cdot \texttt{ca} = \texttt{abacca},$$
$$vw = \texttt{ca} \cdot \texttt{abac} = \texttt{caabac}.$$

(4) **Calcula la potencia cúbica de v. Es decir, encuentra v^3.**

Respuesta:

$$v^3 = v \cdot v \cdot v = \texttt{ca}\,\texttt{ca}\,\texttt{ca} = \texttt{cacaca}.$$

(5) **Calcula la inversa (reversa) de w y v. Es decir, encuentra w^R y v^R.**

Respuesta:

$$w = \texttt{abac} \implies w^R = \texttt{caba}, \qquad v = \texttt{ca} \implies v^R = \texttt{ac}.$$

(6) **¿Cuál es el resultado de concatenar w con la palabra vacía ε? ($w\varepsilon$).**

Respuesta:

$$w\varepsilon = w = \texttt{abac} \qquad (\varepsilon \text{ es elemento neutro para la concatenación}).$$

(7) **Dados los alfabetos $\Sigma_1 = \{0,1,2\}$ y $\Sigma_2 = \{1,3,5\}$, calcula:** $\Sigma_1 \cup \Sigma_2,\ \Sigma_1 \cap \Sigma_2,\ \Sigma_1 - \Sigma_2$.

Respuesta:

$$\Sigma_1 \cup \Sigma_2 = \{0,1,2,3,5\}, \qquad \Sigma_1 \cap \Sigma_2 = \{1\}, \qquad \Sigma_1 - \Sigma_2 = \{0,2\}.$$

(8) **Encuentra una palabra u de longitud 5 sobre $\Sigma = \{a,b\}$ que sea un palíndromo.**

Respuesta: Un ejemplo es

$$u = \texttt{ababa},$$

que verifica $u^R = \texttt{ababa} = u$ y $|u| = 5$.

(9) **Si $x = \texttt{hola}$ y $y = \texttt{mundo}$, ¿es $|xy| = |x| + |y|$? Justifica tu respuesta.**

Respuesta: Sí. Calculamos

$$|x| = |\texttt{hola}| = 4, \qquad |y| = |\texttt{mundo}| = 5.$$

La concatenación $xy = \texttt{hola} \cdot \texttt{mundo} = \texttt{holamundo}$ tiene longitud

$$|xy| = |\texttt{holamundo}| = 9.$$

Por tanto $|xy| = 9 = 4 + 5 = |x| + |y|$.
Justificación general: al concatenar dos palabras se unen secuencialmente sus símbolos sin solapamiento, por lo que el número total de símbolos es la suma de las longitudes.

(10) **Dado el alfabeto $\Sigma = \{x, y\}$, describe el conjunto de todas las palabras de longitud 2 (el conjunto Σ^2).**

Respuesta:

$$\Sigma^2 = \{\, \texttt{xx},\ \texttt{xy},\ \texttt{yx},\ \texttt{yy} \,\}.$$

(Todas las palabras formadas por dos símbolos donde cada símbolo es x o y.)

II. Lenguajes y Operaciones con Lenguajes

Enunciado: Dados el alfabeto $\Sigma = \{a, b\}$ y los lenguajes $L_1 = \{a,\, ab,\, ba\}$ y $L_2 = \{\varepsilon,\, b,\, bb\}$.

(1) **Describe con palabras el siguiente lenguaje infinito:** $L = \{a,\, aa,\, aaa,\, aaaa, \ldots\}$.

Respuesta: Es el lenguaje formado por todas las palabras no vacías compuestas únicamente por la letra a. Formalmente,

$$L = \{a^n \mid n \geq 1\}.$$

(2) **Calcula la unión de L_1 y L_2.** $(L_1 \cup L_2)$

Respuesta:

$$L_1 \cup L_2 = \{\, \varepsilon,\ a,\ b,\ ab,\ ba,\ bb \,\}.$$

(3) **Calcula la intersección de L_1 y L_2.** $(L_1 \cap L_2)$

Respuesta: No hay palabras en común entre L_1 y L_2, por lo que

$$L_1 \cap L_2 = \varnothing.$$

(4) **Calcula la diferencia $L_1 - L_2$.**

Respuesta: Como $L_1 \cap L_2 = \varnothing$, la diferencia es todo L_1:

$$L_1 - L_2 = L_1 = \{\, a,\ ab,\ ba \,\}.$$

(5) **Calcula la concatenación de los lenguajes L_1 y L_2.** $(L_1 L_2)$

Respuesta: Concatenando cada elemento de L_1 con cada elemento de L_2:

$$\begin{aligned} &a \cdot \varepsilon = a, \quad a \cdot b = ab, \quad a \cdot bb = abb, \\ &ab \cdot \varepsilon = ab, \quad ab \cdot b = abb, \quad ab \cdot bb = abbb, \\ &ba \cdot \varepsilon = ba, \quad ba \cdot b = bab, \quad ba \cdot bb = babb. \end{aligned}$$

Eliminando repeticiones:

$$L_1 L_2 = \{\, a,\ ab,\ ba,\ abb,\ bab,\ abbb,\ babb \,\}.$$

(6) **Calcula la potencia al cuadrado del lenguaje L_1.** $(L_1^2 = L_1 L_1)$

Respuesta: Concatenando cada par de elementos de L_1:

$$\begin{aligned} &a \cdot a = aa, \quad a \cdot ab = aab, \quad a \cdot ba = aba, \\ &ab \cdot a = aba, \quad ab \cdot ab = abab, \quad ab \cdot ba = abba, \\ &ba \cdot a = baa, \quad ba \cdot ab = baab, \quad ba \cdot ba = baba. \end{aligned}$$

Tomando los resultados únicos:

$$L_1^2 = \{\, aa,\ aab,\ aba,\ abab,\ abba,\ baa,\ baab,\ baba \,\}.$$

(7) **Sea L el lenguaje de todas las palabras sobre $\{0,1\}$ que tienen un número par de ceros. Indica si las siguientes palabras pertenecen a L:**

$$1010, \quad 0011, \quad 10001, \quad 111.$$

Respuesta:

$$\begin{aligned}
&1010: \text{ tiene dos ceros} \Rightarrow \text{pertenece a } L.\\
&0011: \text{ tiene dos ceros} \Rightarrow \text{pertenece a } L.\\
&10001: \text{ tiene tres ceros} \Rightarrow \text{no pertenece a } L.\\
&111: \text{ tiene cero ceros (0 es par)} \Rightarrow \text{pertenece a } L.
\end{aligned}$$

(8) **Describe formalmente (usando notación de conjuntos) el lenguaje sobre $\Sigma = \{a,b\}$ que contiene todas las palabras que comienzan con a y terminan con b.**

Respuesta: Si $\Sigma = \{a,b\}$, entonces el conjunto es

$$\{\, awb \mid w \in \Sigma^* \,\} \quad \text{o, equivalentemente,} \quad a\Sigma^*b.$$

(9) **¿Cuál es la diferencia entre el lenguaje vacío $\varnothing$ y el lenguaje que solo contiene la palabra vacía $\{\varepsilon\}$?**

Respuesta: El lenguaje vacío $\varnothing$ no contiene ninguna palabra (cardinalidad 0). El lenguaje $\{\varepsilon\}$ contiene exactamente la palabra vacía ε (cardinalidad 1). Son distintos: $\varnothing \neq \{\varepsilon\}$. Además, $\varnothing^* = \{\varepsilon\}$ mientras que $\{\varepsilon\}^* = \{\varepsilon\}$.

(10) **Si $L = \{a,b\}$, ¿cuál es el resultado de concatenar L con el lenguaje $\{\varepsilon\}$? (Es decir, $L\{\varepsilon\}$).**

Respuesta: Como ε es elemento neutro para la concatenación,

$$L\{\varepsilon\} = \{a\varepsilon,\ b\varepsilon\} = \{a,\ b\} = L.$$

III. Cerraduras (Kleene y Positiva) y Lenguaje Universal

Enunciado: Responde las preguntas sobre cerraduras y lenguajes.

(1) **Dado el lenguaje $L = \{ab\}$ sobre $\Sigma = \{a, b\}$, escribe los primeros 5 elementos (ordenados por longitud) del lenguaje L^* (Cerradura de Kleene).**

Respuesta: Recordando que L^* contiene todas las concatenaciones finitas de elementos de L y además la palabra vacía ε, los primeros cinco elementos ordenados por longitud son

$$\varepsilon, \quad ab, \quad abab, \quad ababab, \quad abababab.$$

(En general $L^* = \{ (ab)^n \mid n \geq 0 \}$.)

(2) **Usando el mismo lenguaje $L = \{ab\}$, escribe los primeros 5 elementos de L^+ (Cerradura Positiva).**

Respuesta: L^+ es igual a L^* pero sin ε, es decir todas las potencias positivas de ab. Los primeros cinco elementos son

$$ab, \quad abab, \quad ababab, \quad abababab, \quad ababababab.$$

(En general $L^+ = \{ (ab)^n \mid n \geq 1 \}$.)

(3) **Describe con palabras el lenguaje universal Σ^* para el alfabeto $\Sigma = \{a\}$.**

Respuesta: Para $\Sigma = \{a\}$, el lenguaje universal Σ^* es el conjunto de todas las palabras (incluida la vacía) formadas exclusivamente por repeticiones de la letra a. Formalmente,

$$\Sigma^* = \{ a^n \mid n \geq 0 \} = \{\varepsilon, a, aa, aaa, \ldots\}.$$

(4) **Describe el lenguaje representado por la expresión a^*b sobre $\Sigma = \{a, b\}$. Da 3 ejemplos de palabras en este lenguaje.**

Respuesta: La expresión a^*b denota el conjunto de todas las palabras que consisten en cero o más a seguidas de una única b. Formalmente

$$a^*b = \{\, a^n b \mid n \geq 0 \,\}.$$

Tres ejemplos son:

$$b\ (n = 0), \qquad ab\ (n = 1), \qquad aaab\ (n = 3).$$

(5) **Sea $L = \{aa,\, b\}$. Demuestra cómo se puede generar la palabra "baaaa" a partir de L^*.**

Respuesta: Recordemos que L^* contiene concatenaciones finitas de elementos de L. Observamos que

$$\texttt{baaaa} = \texttt{b} \cdot \texttt{aa} \cdot \texttt{aa},$$

y puesto que $\texttt{b} \in L$ y $\texttt{aa} \in L$, la concatenación $\texttt{b} \cdot \texttt{aa} \cdot \texttt{aa}$ es una concatenación de tres elementos de L. Por tanto

$$\texttt{baaaa} \in L^*.$$

IV. Gramáticas y Jerarquía de Chomsky (Conceptual)

Enunciado: Preguntas conceptuales sobre gramáticas y la jerarquía de Chomsky.

(1) **Nombra los cuatro componentes que definen una gramática formal.**

Respuesta: Una gramática formal se define por la 4-tupla

$$G = (N, \Sigma, P, S)$$

donde

- N es el conjunto finito de símbolos no terminales (variables),
- Σ es el conjunto finito de terminales (alfabeto) tal que $N \cap \Sigma = \varnothing$,
- P es el conjunto finito de reglas de producción (producciones),
- $S \in N$ es el símbolo inicial.

(2) **Asocia cada tipo de lenguaje de la Jerarquía de Chomsky con su autómata reconocedor:**

Lenguaje Regular $\rightarrow$?
Lenguaje Independiente del Contexto $\rightarrow$?
Lenguaje Recursivamente Enumerable $\rightarrow$?

Respuesta:

Lenguaje Regular (Tipo 3) $\rightarrow$ Autómata finito (AF).
Lenguaje Independiente del Contexto (Tipo 2) $\rightarrow$ Autómata de pila (AP).
Lenguaje Recursivamente Enumerable (Tipo 0) $\rightarrow$ Máquina de Turing (MT).

(3) **¿Qué tipo de gramática (según la jerarquía de Chomsky) es la más utilizada para definir la sintaxis de los lenguajes de programación como Java o Python?**

Respuesta: La sintaxis de la mayoría de lenguajes de programación se define mediante *gramáticas libres de contexto* (Context-Free Grammars, Tipo 2). En la práctica se usan gramáticas libres de contexto combinadas con comprobaciones semánticas adicionales (que a menudo requieren análisis más allá de la CFG, es decir chequeos contextuales o acciones semánticas implementadas en el compilador).

(4) **Dada la gramática con reglas $S \rightarrow aS \mid b$ (donde S es el símbolo inicial y a, b son terminales), genera 3 palabras que pertenezcan al lenguaje definido por esta gramática.**

Respuesta: La gramática genera cadenas que son una (o más) repeticiones de a terminadas en b. Ejemplos:

$$b, \quad ab, \quad aab.$$

(Más explícitamente: $S \Rightarrow b$ produce b. $S \Rightarrow aS \Rightarrow ab$ produce ab. $S \Rightarrow aS \Rightarrow aaS \Rightarrow aab$ produce aab.)

(5) **El lenguaje de todos los palíndromos sobre $\{a, b\}$ (ej. "aba", "bbabb") no es regular. ¿A qué categoría superior de la Jerarquía de Chomsky pertenece y por qué se considera más complejo que un lenguaje regular?**

Respuesta: El conjunto de palíndromos sobre $\{a, b\}$ pertenece a la clase de *lenguajes independientes del contexto* (Tipo 2). Se puede dar una gramática libre de contexto que los genera, por ejemplo

$$S \rightarrow aSa \mid bSb \mid a \mid b \mid \varepsilon,$$

la cual genera exactamente todos los palíndromos.

Se considera más complejo que un lenguaje regular porque reconocer si una palabra es palíndroma requiere, en general, comparar símbolos en posiciones "simétricas"desde los extremos hacia el centro. Un autómata finito (AF) no dispone de memoria no acotada para recordar la mitad de la palabra y hacer esa comprobación; en cambio un autómata con pila (AP) puede *apilar* la primera mitad y luego comparar durante la segunda mitad, por eso los palíndromos son reconocibles por un AP pero no por un AF. Esta necesidad de memoria (pila) es la razón estructural de la mayor complejidad frente a los lenguajes regulares.