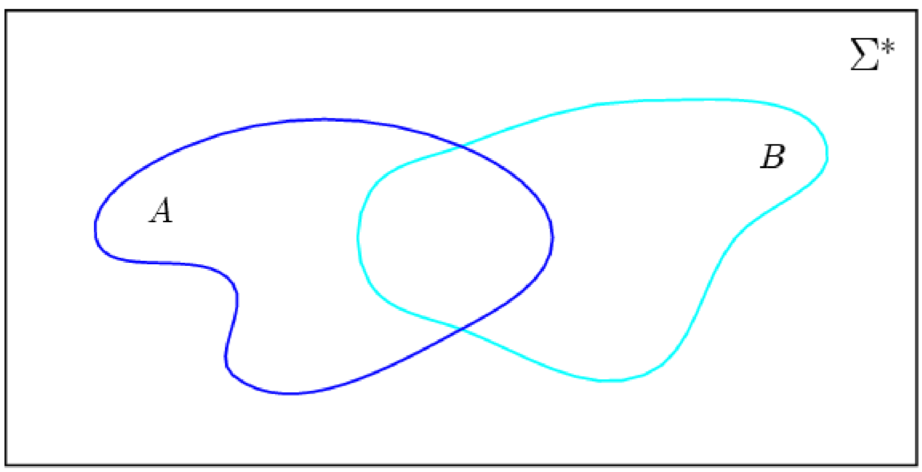
1.7. Lenguajes

Un **lenguaje** L sobre un alfabeto Σ es un subconjunto de Σ^*, es decir $L \subseteq \Sigma^*$.

Casos extremos:

$L = \varnothing$,	lenguaje vacío.
$L = \Sigma^*$,	lenguaje de todas las cadenas sobre Σ.

Todo lenguaje L satisface $\varnothing \subseteq L \subseteq \Sigma^*$, y puede ser finito o infinito. Los lenguajes se denotan con letras mayúsculas $A, B, C, \ldots, L, M, N, \ldots$. En la siguiente gráfica se visualizan dos lenguajes A y B sobre Σ.



Ejemplo Los siguientes son ejemplos de lenguajes sobre los alfabetos especificados.

- $\Sigma = \{a, b, c\}$. $L = \{a, aba, aca\}$.
- $\Sigma = \{a, b, c\}$. $L = \{a, aa, aaa, \ldots\} = \{a^n : n \geq 1\}$.
- $\Sigma = \{a, b, c\}$. $L = \{\lambda, aa, aba, ab^2a, ab^3a, \ldots\} = \{ab^na : n \geq 0\} \cup \{\lambda\}$.
- $\Sigma = \{a, b, c, \ldots, x, y, z\}$. $L = \{u \in \Sigma^* : u$ aparece en el diccionario español DRA$\}$. L es un lenguaje finito.
- $\Sigma = \{a, b, c\}$. $L = \{u \in \Sigma^* : u$ no contiene el símbolo $c\}$. Por ejemplo, $abbaab \in L$ pero $abbcaa \notin L$.

- $\Sigma = \{0, 1\}$. $L =$ conjunto de todas las secuencias binarias que contienen un número impar de unos.

- $\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. El conjunto $\mathbb{N}$ de los números naturales se puede definir como un lenguaje sobre Σ, en la siguiente forma:

$$\mathbb{N} = \{u \in \Sigma^* : u = 0 \text{ ó } 0 \text{ no es un prefijo de } u\}.$$

Ejercicio Definir el conjunto de los enteros $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ como un lenguaje sobre un alfabeto adecuado.

✎ El concepto abstracto de "lenguaje", tal como se ha definido, no es exactamente la misma noción utilizada en la expresión "lenguaje de programación". Para precisar la relación entre estos conceptos, consideremos el alfabeto Σ de los caracteres ASCII. Un programa en C o en Pascal, por ejemplo, es simplemente una cadena de símbolos de Σ y, por lo tanto, un conjunto de programas es un lenguaje (en el sentido formal definido en esta sección).