MD-01

# Secuencias finitas y palabras. Conceptos fundamentales

***Alberto A. Herrera Becerra*, Francisco J. Rodríguez Ramírez ***


## RESUMEN

Se presenta una introducción al tema de Secuencias Finitas y Palabras; estas entidades matemáticas se caracterizan por tres propiedades peculiares: son entidades compuestas, formadas a partir de elementos primitivos; estos elementos pueden aparecer repetidas veces; y todas estas apariciones están ordenadas. Estas construcciones son fundamentales en las áreas de la Teoría de los Lenguajes Formales, donde son los bloques de construcción de las entidades llamadas lenguajes, y de la Ciencia de la Computación Teórica, donde son los elementos clave para construir modelos de computadoras como procesadores lingüísticos o sistemas dinámicos discretos. A pesar de su importancia, el tema de las secuencias finitas se presenta de una manera muy informal e intuitiva en muchos de los libros sobre la Teoría de los Lenguajes Formales y Autómatas. Nosotros estamos convencidos de que este enfoque no permite que los estudiantes aprecien todas las propiedades y características de las secuencias finitas. Debido a esto, por varios años hemos desarrollado materiales educativos que ayuden y motiven al estudiante a profundizar en el estudio de estas fascinantes entidades matemáticas. Así pues, en este reporte damos a conocer nuestro enfoque al tema, presentando las definiciones fundamentales y resaltando la naturaleza combinatoria de las secuencias finitas y palabras.

## ABSTRACT

This paper is an introduction to Finite Sequences and Words, mathematical entities with very peculiar properties: they are composed entities, that is, they are formed from primitive elements; these primitive elements could appear several times in the composed one; and the primitive elements appeared ordered within the composed one. These constructions are of a fundamental importance in the fields of Formal Language Theory, where they are the building blocks of the entities known as languages, and Theoretical Computer Science, where they are the key elements to build computer models as linguistic processors or discrete dynamical systems. Despite their importance, the theme of finite sequences is presented in a very informal or intuitive way in most of the textbooks on Formal Languages and Automata. In our view, this approach does not allow students to fully appreciate all the properties and features of these mathematical objects; certainly, they are much more complex than the intuitive approach suggest. For several years, we have been developing learning aids to help and motivate students to immerse themselves in a very exciting theme of discrete math. So, in this paper we present our approach to the theme considering the fundamental definitions and the combinatoric nature of words.


**Palabras claves**: Matemáticas Discretas, Indexación, Secuencias Finitas, Combinatoria.

## INTRODUCCIÓN

Es muy común encontrar en los libros de nivel introductorio sobre Lenguaje Formales afirmaciones similares a la siguiente. “Una secuencia finita de símbolos de un determinado alfabeto se conoce como una palabra sobre dicho alfabeto Kelley (1995)”. Es cierto que, desde un punto de vista informal, afirmaciones como la anterior son apropiadas, pero ciertamente carecen de rigor matemático. El defecto principal, no se aclara qué se debe entender por el concepto de secuencia finita como una entidad matemática. La situación no mejora si se recurre a definiciones de diccionario. Por ejemplo, en el Diccionario de la Lengua Española de la Real Academia Española se encuentran aseveraciones como las siguientes. “Serie o sucesión de cosas que guardan entre sí cierta relación” o “conjunto de cantidades u operaciones ordenadas de tal modo que cada una está determinada por las anteriores RAE (2001)”. De hecho, en este diccionario se señala que la segunda aseveración anterior se puede entender en un sentido matemático. De inmediato queda la duda si la aseveración anterior realmente es una definición de carácter matemático; por decir lo menos, parece una aseveración ambigua. El problema con todo esto, es que la idea intuitiva se usa como base para desarrollar toda una teoría sobre los llamados Lenguajes Formales. Es bien sabido en las Matemáticas que es muy peligroso basar toda una teoría en ideas intuitivas: más temprano que tarde se pueden encontrar inconsistencias en los razonamientos desarrollados y, tal vez peor, contradicciones.

Debido a lo anterior, durante varios años nos hemos dedicado a revisar las definiciones matemáticas del concepto de secuencia o sucesión en lo general, y las características particulares que tienen las secuencias finitas. Hemos encontrado diversos enfoques, pero todos basados en los conceptos fundamentales de las matemáticas discretas: Conjuntos, relaciones y funciones. Algunos de los enfoques son introductorios, algunos otros son más sofisticados, aunque se debe señalar que todos ellos parten de una definición precisa, no sólo de una idea intuitiva. Nosotros creemos que un enfoque apropiado para cursos introductorios a la teoría de los

* Instituto de Ciencias Aplicadas y Tecnología, UNAM, Departamento de Instrumentación Científica e Industrial, alberto.herrera@icat.unam.mx, M. en I.B.B. Facultad de Ingeniería, UNAM, División de Ingeniería Eléctrica, frajorora@yahoo.co.uk, Ing.

lenguajes formales y sus aplicaciones a la teoría de las computadoras se debe basar en el concepto de función, tal y como se maneja en las matemáticas discretas. Con base en este enfoque, se pueden describir las características y propiedades esenciales de las sucesiones en lo general, las características particulares de las sucesiones finitas y, de manera muy especial, las principales operaciones que es conveniente definir para su manipulación.

En este documento presentamos una breve descripción del enfoque basado en funciones para describir las características y propiedades esenciales de las secuencias finitas, en el contexto de una realización particular: las entidades llamadas palabras en la teoría de los lenguajes formales. En especial, mostramos que una palabra (una secuencia finita) es una entidad combinatoria, esto es, existe, en lo general, una gran cantidad de realizaciones diferentes de una palabra particular que satisfacen las mismas características esenciales. Este aspecto de las palabras (secuencias finitas) es básico para entender la naturaleza de los métodos de procesamiento de palabras que se desarrollan en la teoría de los lenguajes formales y que, eventualmente, permiten establecer diferentes familias de lenguajes, cada una de ellas con sus propias propiedades distintivas.

## 1. INDEXACIÓN

Como ya se mencionó muy brevemente en la Introducción, la herramienta básica para construir las secuencias son las funciones, entendidas de manera abstracta. Desde este punto de vista, una función no es más que un conjunto de pares ordenados (esto es, una relación binaria) que satisface algunas condiciones adicionales. Así, una función (o mapeo) $f$ de un conjunto $A$ a un conjunto $B$, denotada $f: A \to B$, es cualquier relación binaria $R$ de $A$ a $B$ tal que para toda $a \in A$ existe exactamente una $b \in B$ tal que $(a, b) \in R$. Con esto, las funciones mantienen su característica básica identificada en sus muchas aplicaciones: las funciones se pueden entender como una regla que permite asociar un objeto (en muchos casos, un valor) con otro objeto (otro valor). Adicionalmente, una función se puede usar para "ordenar" los elementos de un conjunto. Es esta noción de ordenamiento la que se va a usar en la definición del concepto de secuencia. Esta definición se introduce en dos etapas. Primero se introduce la idea de una indexación; luego se introduce la idea de secuencia como una indexación particular.

Sea $I$ un conjunto. Una *indexación* del conjunto $D$ por $I$ es una función $s: I \to D$; la función $s$ también se denomina *familia de elementos indexados* por $I$. Así, $s$ es un mapeo de la forma $i \mapsto s(i) = d$, donde $i \in I$ y $d \in D$. Se dice entonces que el elemento $d$ ha sido señalado o seleccionado por medio del elemento $i$; debido a esto, el elemento $i$ se denomina un *índice* para el elemento $s(i) = d$ y, en consecuencia, el conjunto $I$ se denomina el *conjunto de índices*. Una indexación $s$ es una regla para señalar o seleccionar elementos del conjunto $D$.

Un elemento $s(i)$ de una indexación $s$ usualmente se denota $s_i$. Con esto, la familia de elementos indexados $s$ es el conjunto $s = \{(i, s_i) : i \in I\}$. El conjunto anterior puede ser denotado de manera compacta $(s_i)_{i \in I}$, o simplemente $(s_i)$ si el conjunto de índices es conocido. A manera de ejemplo, sea el conjunto de índices $I = \{a, b, c, d, e, f\}$ y sea la familia de conjuntos $M = \{A, B, C, D, E, F, G, H\}$. Con los dos conjuntos $I$ y $M$ se puede construir la familia indexada $s$ siguiente:

$$s = \{(a, C), (b, H), (c, B), (d, E), (e, D), (f, A)\}. \qquad (1)$$

La familia anterior se representa de manera compacta como sigue:

$$s = (s_a, s_b, s_c, s_d, s_e, s_f) = (C, H, B, E, D, A). \qquad (2)$$

Se debe destacar que el orden en el que se listan los elementos de $s$ no es relevante. Debido a esto, se tiene lo siguiente:

$$s = (s_f, s_d, s_e, s_b, s_a, s_c) = (A, E, D, H, C, B). \qquad (3)$$

El rango de $s$ se denota $\{s_i : i \in I\}$; esto permite hablar de una familia indexada en términos de conjuntos, precisamente el conjunto de imágenes definidas por la función $s$. De esto se puede construir una imagen mental muy sugerente sobre las familias indizadas. Una familia puede ser vista como un conjunto de "contenedores" tales que cada "contenedor" está identificado con uno y sólo un índice, y cada "contenedor" contiene uno y sólo un objeto del conjunto blanco. Esta imagen va a ser usada más adelante en el documento.

Es muy importante destacar que un elemento $d \in D$ puede tener varios índices bajo $s$ debido a que esta función no tiene porque ser inyectiva. Esto es, puede ocurrir que $s_i = s_j = d$ para $i, j \in I$ con $i \neq j$.

A continuación, se describe una clase muy especial de familias indexadas, la clase fundamental para el estudio de los lenguajes formales.

## 2. SECUENCIAS FINITAS

Un caso especial de indexación surge cuando el conjunto de índices es un conjunto de números naturales de la forma general $I_n = \{k \in \mathbb{N} : k < n\}$ para algún $n \in \mathbb{N}$ dado. Para estos conjuntos es más conveniente adoptar la notación $\{0, 1, \ldots, n-1\}$, obviamente más sugerente. Nótese que cuando $n = 0$, la especificación anterior denota al conjunto vacío $\emptyset$.

Se define a una *secuencia finita* $s$ como una indexación cuyo conjunto de índices es precisamente $I_n = \{0, \ldots, n-1\}$ para algún número natural $n$ dado y cuyo rango es un conjunto $D$. Esto es, se tiene que $s: \{0, \ldots, n-1\} \to D$. El número natural $n$ referido en la definición se denomina la *longitud* de $s$ y, en consecuencia, se dice que $s$ es una secuencia (finita) de longitud $n$; la longitud de la secuencia se denota $|s|$, o $long(s)$. Los valores de $s$ son denominados las *componentes* de $s$. Una secuencia finita $s$, de longitud $n$, es denotada $s = (s_0, s_1, \ldots, s_{n-1})$, con $s_i = s(i)$ para $0 \leq i \leq n-1$. Aquí es necesario establecer una regla de escritura: las componentes $s_i$ se deben listar siguiendo el orden natural de sus índices.

Una secuencia finita también se puede definir eligiendo el conjunto de índices $I_n = \{m, m+1, m+2, \ldots, m+n-1\}$, para $m, n \in \mathbb{N}$ dados, con $m < n$. Nótese que también se tienen $n$ números naturales consecutivos como índices. Sin embargo, el uso de este

AÑO 8. No. 8. ISSN-2448-7236. SEPTIEMBRE 2023 – AGOSTO 2024. MD-01, pág.: 1 a la 10.

conjunto de índices añade complicaciones innecesarias para los fines que se buscan. También se puede introducir y formalizar la noción de una secuencia infinita usando como conjunto de índices a todo el conjunto $\mathbb{N}$. Sin embargo, este concepto no va a ser usado en el presente documento y, en consecuencia, no se discutirá más.

Se define ahora a la secuencia (finita) de longitud cero como la función de indexación cuyo dominio es el conjunto vacío y cuyo rango es $D$. Ya se sabe que existe exactamente una función como tal, a saber, la función vacía, por lo que sólo hay una secuencia de longitud cero. Esta secuencia será denominada la *secuencia nula* o la *secuencia vacía*, y será denotada $\varepsilon$. Así, $\varepsilon: \emptyset \to D$ y $|\varepsilon| = 0$. (Cabe señalar que otros autores utilizan notaciones alternativas, como $\epsilon$ o como $\lambda$, para denotar a la secuencia vacía.)

Una propiedad fundamental de las secuencias finitas es la siguiente. Se dice que dos secuencias finitas $s$ y $t$ sobre el mismo conjunto $D$ son *iguales*, lo cual se denota $s = t$, si y sólo si todas sus componentes coinciden. Esto es, se requiere que $(s_0, \dots, s_{n-1}) = (t_0, \dots, t_{m-1})$, pero esta igualdad sólo puede suceder si $n = m$, las secuencias tienen las mismas longitudes, y $s_i = t_i$ para $0 \le i \le n-1$, sus valores son iguales. Así pues, resulta que dos secuencias de longitudes diferentes no pueden ser iguales, y dos secuencias con la misma longitud todavía pueden ser diferentes si por lo menos una componente no coincide.

La notación $(d, d')$ se usa para denotar a una secuencia de longitud dos. Sin embargo, en la teoría de conjuntos se suele usar la expresión $(d, d')$ como el par ordenado $\{\{d\}, \{d, d'\}\}$, el cual es un objeto matemático cuya característica distintiva es que $(d, d') = (d_1, d_1')$ si y sólo si $d = d_1$ y $d' = d_1'$. En el contexto de las secuencias finitas, resulta que la última igualdad también es válida, por lo que la ambigüedad no es muy dañina si se le maneja con precaución.

Por su parte, una secuencia $t$ de longitud uno, una secuencia unitaria, está completamente determinada por su valor $(t_0)$. Muchas veces, sin embargo, se encontrará que es conveniente considerar que un objeto $d$ puede ser usado como la secuencia unitaria $(d)$, cuya componente es precisamente $d$. Este uso permite simplificar muchas descripciones que, de manera más formal, se vuelven un tanto engorrosas. Nuevamente, el uso señalado no es muy dañino si se hace con precaución.

Se introduce una clase especial de secuencias finitas en la sección siguiente.

## 3. PALABRAS

Se presenta una clase especial de secuencias finitas definidas sobre símbolos. Un *símbolo* es una marca, signo o palabra que indica, significa o se entiende que representa a una idea, objeto o relación. Los símbolos son un medio para la comunicación compleja que a menudo puede tener múltiples niveles de significación. Los símbolos son la base de toda la comprensión humana y sirven como vehículos de concepción para todo el conocimiento humano. Los símbolos facilitan la comprensión del mundo en el que vivimos, sirviendo en consecuencia como un cimiento sobre el que realizamos nuestros juicios. En este sentido, las personas no sólo usan los símbolos para hacer sentido del mundo a su alrededor, sino también para identificar y cooperar en sociedad a través de una retórica constitutiva (creación de una identidad colectiva).

Para empezar, un *alfabeto* es un conjunto finito no vacío de símbolos. Algunos autores consideran que si un conjunto $A$ es un alfabeto, entonces en $A$ no se deben incluir símbolos que puedan formarse por yuxtaposición de otros símbolos de $A$. Esta restricción impide confusiones con el concepto de palabra, que se dará más adelante.

Puesto que un alfabeto es simplemente un conjunto finito no vacío, dados los alfabetos $A_1$ y $A_2$, se tiene que $A_1 \cup A_2$ también es un alfabeto. Es más, si los conjuntos $A_1 \cap A_2$, $A_1 - A_2$ y $A_2 - A_1$ son no vacíos, entonces también son alfabetos.

Una *palabra* de longitud $n$ sobre un alfabeto $A$ es una secuencia finita de longitud $n$ de símbolos de $A$. Una palabra $w$ de longitud $n$ se denota $w = (w_0, w_1, \dots, w_{n-1})$ donde $w_i \in A$ para $0 \le i \le n-1$. Se debe recordar que los valores $w_i$ no necesariamente son todos diferentes entre sí. Debido a esto, la *longitud* de la palabra $w$, denotada por $|w|$ o $long(w)$, es simplemente el número total de símbolos que la forman, incluyendo todas las posibles repeticiones de símbolos.

Ahora bien, debido a que se está considerando que cualquier símbolo que forme parte de un alfabeto no resulta de la combinación de otros símbolos, es más conveniente denotar a la palabra $w = (w_0, w_1, \dots, w_{n-1})$ simplemente como $w = w_0 w_1 \dots w_{n-1}$; esto es, se escriben los símbolos que forman a la palabra uno detrás de otro, sin marcadores que los separen. Debido a esto, en lo que sigue a menudo se usa el término *cadena* como sinónimo de palabra.

Es conveniente señalar que las palabras de longitud unitaria se escriben ahora simplemente $a_0$, en lugar de $(a_0)$. Nótese que esto se presta a una mayor confusión con el símbolo $a_0$ del alfabeto en cuestión. Deberá ser el contexto el que ayude a resolver la posible ambigüedad; lo mejor es señalar siempre de manera explícita la diferencia.

Para un alfabeto $A$ dado se define una única palabra de longitud cero, o *palabra vacía* o palabra *nula*, que se denota $\varepsilon$. Se debe cuidar que $\varepsilon$ no sea un símbolo de ningún alfabeto, ya que $\varepsilon$ denota a una palabra. Dados dos alfabetos diferentes, digamos $A_1$ y $A_2$, para cada uno de ellos se define una palabra vacía, pero estas palabras vacías no son iguales, ya que una está definida sobre $A_1$, mientras que la otra está definida sobre $A_2$.

Sea $w$ una palabra sobre un alfabeto $A$ y sea $a$ un símbolo de $A$. Se define la *longitud* $a$ de $w$, denotada $|w|_a$ o $long_a(w)$, como el número de veces que aparece $a$ en $w$. Es claro que $|\varepsilon|_a = 0$ para todo $a \in A$ y que para toda palabra $w$ de cualquier alfabeto $A$ se satisface la igualdad:

$$|w| = \sum_{a \in A} |w|_a. \qquad (4)$$

Como una consecuencia más de que las palabras sean casos particulares de las secuencias finitas, se tiene que dos palabras $u$, $v$ sobre un alfabeto $A$ común son *iguales*, lo cual se denota $u = v$, si ambas palabras tienen la misma longitud y los mismos valores.

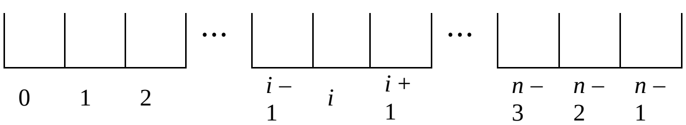
AÑO 8. No. 8. ISSN-2448-7236. SEPTIEMBRE 2023 – AGOSTO 2024. MD-01, pág.: 1 a la 10.

Nótese que esto significa que las palabras tienen los mismos símbolos para los mismos índices o, dicho de manera coloquial, tienen los mismos símbolos en las mismas posiciones. Dos palabras que no tienen la misma longitud no pueden ser iguales, y aunque dos palabras tengan la misma longitud puede que no sean iguales si por lo menos uno de sus valores no coincide.

Ahora bien, dado que cada componente $w_i$ de una palabra $w = w_0 w_1 \dots w_{n-1}$ sobre un alfabeto $A$ conlleva un significado individual intrínseco, se puede esperar entonces que la composición ordenada $w$ conlleve un significado compuesto. Sin embargo, en la presentación que se va a realizar no se tomarán en cuenta ni los significados individuales de los símbolos disponibles, ni el posible significado compuesto de las palabras formadas con esos símbolos. Desde un punto de vista lingüístico, se va a realizar un estudio sintáctico de las palabras, esto es, sólo va a ser relevante el orden secuencial de los símbolos que forman a la palabra. En la sección que sigue se presentan las particularidades más relevantes de los estudios sintácticos de las palabras.

### 3.1 Las palabras como cadenas de contenedores

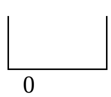
Para guiar el estudio sintáctico de las palabras se va a hacer uso de la idea de "contenedores" presentada en la sección de Indexación. Así, una palabra $w$ de longitud $n$ se puede visualizar como una colección finita y ordenada de $n$ "contenedores", cada uno de ellos etiquetado con un número natural. El ordenamiento de los "contenedores" se basa en el orden natural de los índices numéricos empleados.



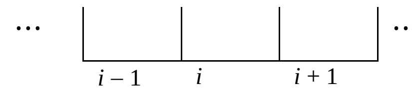

Así pues, las palabras se visualizan como cadenas de "contenedores" consecutivos, uno de detrás de otro. En cada "contenedor" de la cadena se debe colocar uno y sólo un símbolo de un alfabeto $A$ previamente establecido. No pueden existir "contenedores" vacíos y no pueden existir "contenedores" con dos o más símbolos.

Con esta imagen, la palabra vacía se puede entender como una cadena vacía de "contenedores". Dicho de otra manera, en la palabra vacía no hay "contenedor" alguno en el cual colocar símbolos.

Una característica distintiva fundamental de toda palabra no vacía de longitud $n$ es que siempre deben existir un *primer* "contenedor" -el cual tiene la etiqueta $0$-, y un *último* "contenedor" -con la etiqueta $n-1$. Un caso especial es la palabra de longitud unitaria, ya que su único "contenedor" es tanto el primero como el último.

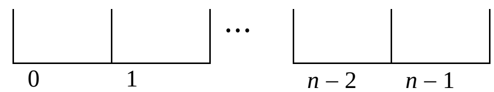



Otra característica distintiva es que todo "contenedor" debe poseer un "contenedor" antecesor inmediato y un "contenedor" sucesor inmediato, como se ilustra a continuación.



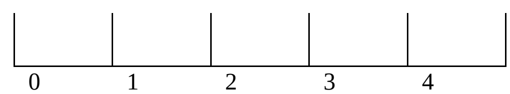

En este ejemplo, el "contenedor" con etiqueta $i-1$ es el antecesor inmediato del "contenedor" $i$, mientras que el "contenedor" con etiqueta $i+1$ es su sucesor inmediato.

Existen dos excepciones a lo anterior. La primera excepción corresponde al primer "contendor", el cual no tiene un antecesor inmediato. La segunda excepción corresponde al último "contenedor", el cual no tiene un sucesor inmediato.




Esto es, el primer "contenedor", con etiqueta $0$, tiene como sucesor inmediato al "contenedor" con etiqueta $1$ (el segundo "contenedor"), pero no tiene un "contenedor" que sea un antecesor inmediato. Por su parte, el último "contenedor", con etiqueta $n-1$, tiene al "contenedor" con etiqueta $n-2$ (el penúltimo) como su antecesor inmediato, pero no tiene un "contenedor" que sea su sucesor inmediato. Nuevamente, la palabra de longitud unitaria representa un caso especial, ya que es una palabra no vacía cuyo único "contenedor" no posee ni antecesor ni sucesor inmediatos.

En toda palabra de longitud mayor o igual a $3$ se pueden considerar "contenedores" antecesores o sucesores a un "contenedor" dado. Por ejemplo, considérese la palabra de longitud $5$ ilustrada a continuación.




Si se toma como referencia al "contenedor" $1$, se observa que tiene al "contenedor" $0$ como su antecesor inmediato y único antecesor; que tiene al "contenedor" $2$ como su sucesor inmediato; y que tiene a los "contenedores" $2$, $3$ y $4$ como sus sucesores. Por su parte, el "contenedor" $2$ tiene a los contenedores $0$ y $1$ como sus antecesores, siendo $1$ su antecesor inmediato, y tiene a los contenedores $3$ y $4$ como sus sucesores, siendo $3$ su sucesor inmediato.

### 3.2 Las palabras como cadenas de símbolos

Considérese un alfabeto $A$ y considérese que $a_j$ es un símbolo arbitrario de $A$. Si se considera ahora una cadena de $n$ "contenedores", se puede formar la palabra $w$ de longitud $n$ siguiente:

Memorias del Congreso Científico Tecnológico de las carreras de
Ingeniería Mecánica Eléctrica, Industrial y Telecomunicaciones, sistemas y electrónica

AÑO 8. No. 8. ISSN-2448-7236. SEPTIEMBRE 2023 – AGOSTO 2024. MD-01, pág.: 1 a la 10.

| $a_j$ | $a_j$ | $\cdots$ | $a_j$ | $\cdots$ | $a_j$ | $a_j$ |
|---|---|---|---|---|---|---|
| 0 | 1 | | $i$ | | $n-2$ | $n-1$ |

Se debe tener cuidado al interpretar la figura anterior. No se debe entender que en cada "contenedor" de la cadena se coloca un mismo símbolo, $a_j$ en la figura. Lo que se quiere decir es que en cada "contenedor" se puede colocar un símbolo arbitrario $a_j$ del alfabeto $A$. La elección es totalmente arbitraria, se puede escoger <u>cualquier</u> símbolo de $A$.

Con base en la imagen anterior, la palabra $w$ se entiende que es una función de la forma $w: I \to A$, con $I = \{0,1,\dots,n-1\}$, y que está dada por:

$$\begin{array}{rcccl} w(0) & = & w_0 & = & a_{j_0} \\ w(1) & = & w_1 & = & a_{j_1} \\ & \vdots & & & \\ w(i) & = & w_i & = & a_{j_i} \\ & \vdots & & & \\ w(n-1) & = & w_{n-1} & = & a_{j_{n-1}} \end{array} \qquad (5)$$

Se debe entender que dos símbolos $a_{j_k}$ y $a_{j_l}$, con $k \neq l$, pueden ser, o no ser, iguales. Dicho de otra manera, puede suceder que $w(k) = w(l)$, o puede suceder que $w(k) \neq w(l)$, para $k \neq l$.

La representación funcional anterior también se puede entender como el conjunto de pares ordenados siguiente:

$$w = \{(0, a_{j_0}), (1, a_{j_1}), \dots, (i, a_{j_i}), \dots, (n-2, a_{j_{n-2}}), (n-1, a_{j_{n-1}})\}, \qquad (6)$$

En este conjunto, los valores $a_{j_i}$ se deben interpretar como el símbolo arbitrario $a_j$ del alfabeto $A$ con índice $i$, donde $0 \leq i \leq n-1$. En estas descripciones, se ha hecho uso de la noción de secuencia finita como un conjunto indizado por el conjunto de índices numéricos $I$.

En las aplicaciones, suele ser útil comenzar con la representación de conjunto de pares ordenados y luego construir la representación funcional.

A manera de ejemplo, considérese el alfabeto $A = \{a, b, c, d, e\}$. En este caso, los caracteres $a$, $b$, $c$, $d$ y $e$ no necesariamente se deben interpretar como letras de algún vocabulario; simplemente indican un alfabeto formado por cinco símbolos diferentes y distinguibles. Considérese ahora la cadena de símbolos de longitud $5$ siguiente:

| $b$ | $c$ | $a$ | $b$ | $d$ |
|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 |

Esta figura se debe interpretar como sigue. En el "contenedor" $0$ se colocó el símbolo $b$; en el "contenedor" $1$ se colocó el símbolo $c$; en el "contenedor" $3$ se colocó el símbolo $a$; y así sucesivamente. La cadena anterior también se puede representar de manera más compacta como sigue:

| $b_0$ | $c_1$ | $a_2$ | $b_3$ | $d_4$ |
|---|---|---|---|---|

Esta última figura también se puede interpretar de manera funcional. Así, la figura se puede interpretar como el conjunto de pares ordenados siguiente:

$$w = \{(0,b), (1,c), (2,a), (3,b), (4,d)\}. \qquad (7)$$

A partir de esta representación, ya es fácil establecer la representación funcional siguiente:

$$\begin{array}{rcccl} w(0) & = & w_0 & = & b \\ w(1) & = & w_1 & = & c \\ w(2) & = & w_2 & = & a \\ w(3) & = & w_3 & = & b \\ w(4) & = & w_4 & = & d \end{array} \qquad (8)$$

En particular, nótese que $w(0) = w(3) = b$. Esto es, el símbolo $b$ tiene dos índices diferentes, $0$ y $3$. Dicho de otra manera, en la cadena $w$ hay dos "contenedores" diferentes, $0$ y $3$ precisamente, que contienen copias del mismo símbolo, $b$ en este caso.

De todo lo anterior, debería resultar claro que una palabra definida sobre un alfabeto $A$ dado, es una entidad compuesta (de símbolos), cuyos componentes están ordenados (por una enumeración), y donde algunos componentes pueden aparecer repetidas veces. Se debería añadir que no es necesario usar todos los símbolos del alfabeto para formar palabras.

### 3.4 Notaciones alternativas

Considérese que la palabra $w$ se representa inicialmente en la forma de conjunto de pares ordenados:

$$w = \{(0, a_{j_0}), (1, a_{j_1}), \dots, (i, a_{j_i}), \dots, (n-2, a_{j_{n-2}}), (n-1, a_{j_{n-1}})\}, \qquad (9)$$

Esta palabra $w$ también se puede expresar de manera simplificada como sigue:

$$w = (a_{j_0}, a_{j_1}, \dots, a_{j_{n-2}}, a_{j_{n-1}}). \qquad (10)$$

Si además se satisfacen las recomendaciones establecidas para los símbolos de un alfabeto, se puede emplear la notación simplificada:

$$w = a_{j_0} a_{j_1} \cdots a_{j_{n-2}} a_{j_{n-1}}. \qquad (11)$$

Retomando el ejemplo de la sección anterior, la palabra particular $w$ que se analizaba puede ser representada de manera muy compacta como sigue:

$$w = b_0 c_1 a_2 b_3 d_4. \qquad (12)$$

Se pueden omitir los índices, llegando a la expresión siguiente:

$$w = bcabd. \qquad (13)$$

En esta última forma de representación, el orden en el que se escriben los símbolos es fundamental. El primer símbolo de la palabra es el primero que se escribe; el segundo es el segundo que se escribe; y así sucesivamente hasta llegar al último símbolo.

### 3.5 Algunos conjuntos de palabras

Los conjuntos de palabras sobre un alfabeto dado se denominan *lenguajes*. El estudio de las características distintivas y propiedades de los lenguajes se presentará en un reporte futuro.
Dado un alfabeto $A$, el conjunto de todas las palabras sobre $A$ se denota $A^*$, el cual se lee "$A$ estrella". Más adelante se discuten la cardinalidad y ciertas propiedades de algunos subconjuntos interesantes de $A^*$. Estos subconjuntos son los siguientes.
El conjunto de palabras $A^i$, el cual se lee "$A$ a la $i$", donde $i$ es un número natural tal que $i \geq 0$, es el conjunto de todas las palabras de longitud $i$ sobre $A$. Esto es, se tiene que $A^i = \{w \in A^* : |w| = i\}$, donde $i \geq 0$ es un número dado.

## 4. CONSTRUCCIÓN DE PALABRAS

En la sección previa se presentaron los conjuntos de palabras $A^i$, señalando qué palabras los forman, pero no se dijo nada acerca de cómo construir tales palabras. Este problema de construcción es discutido en las secciones siguientes.
La construcción de palabras específicas se puede realizar de la manera siguiente.

- Primero, es indispensable fijar el alfabeto $A$ de referencia. Los símbolos de este alfabeto serán los únicos que se podrán emplear en la construcción.
- Segundo, se deberán establecer las longitudes de las palabras que se desean construir. De esta manera se deja establecida la cantidad total de símbolos que se van a utilizar.
- Tercero, se deberán establecer posibles restricciones al uso de los símbolos. Por decir, se debe establecer cuáles símbolos del alfabeto realmente van a ser utilizados, y cuáles no van a ser considerados. Se debe establecer cuáles de los símbolos realmente utilizados van a aparecer múltiples veces y señalar la cantidad de repeticiones. Se debe indicar si aparecen algunas combinaciones de símbolos (o patrones) y se debe fijar la cantidad de veces que aparece el patrón. Finalmente, se deben señalar las posiciones en las que deben aparecer los símbolos o patrones distinguidos.

De lo anterior, debe resultar claro que el proceso de construcción de palabras particulares no es un proceso simple. Se debe tomar en cuenta que, en lo general, va a existir una cantidad muy grande de palabras diferentes que satisfacen las condiciones establecidas en el proceso de construcción. De entre todas estas palabras se deben seleccionar las palabras específicas que se desean.
Lo anterior señala claramente que todo problema de construcción de palabras involucra un *cálculo combinatorio*, y es bien conocido que los problemas combinatorios no son triviales y generalmente son muy engorrosos.

### 4.1 Primeros cálculos combinatorios

Todo conjunto $A^i$ es finito, aunque posiblemente sea muy grande. Esta afirmación se establece a partir del razonamiento siguiente. Para un conjunto $A^i$ dado, cada una de sus palabras tiene la forma general siguiente:

$x_{j_0} x_{j_1} \dots x_{j_{i-1}}$, donde cada $x_j$ es alguno de los símbolos de $A$.

La forma anterior se puede representar como una cadena de $i$ "contenedores", cada uno de los cuales contiene uno, y sólo uno, de los símbolos del alfabeto, y no puede haber contenedores vacíos. Ahora bien, en la primera posición (o primer contenedor) de la palabra se puede colocar cualquiera de los símbolos del alfabeto; las elecciones posibles dependen de la cantidad de símbolos distinguibles, o de la cardinalidad del alfabeto $A$. Colocado el primer símbolo, en la segunda posición (o segundo contenedor) se puede colocar nuevamente a cualquiera de los símbolos del alfabeto. Y así sucesivamente. Debe resultar claro que la elección de un símbolo para una posición (o contenedor) de la palabra no depende de las elecciones realizadas previamente, y no influye sobre las elecciones que se van a realizar después. Todas las elecciones son independientes entre sí. En consecuencia, la cantidad de palabras que forman al conjunto $A^i$ es $\#\left(A^i\right) = (\#A)^i$; este resultado se obtiene usando el llamado principio de la multiplicación. Ya que se ha establecido una cardinalidad para el conjunto $A^i$, se concluye que el conjunto es finito y su tamaño depende del tamaño del alfabeto $A$.
Con base en el resultado general anterior se tienen dos resultados particulares inmediatos. El primero corresponde al caso en que $i = 0$. Se tiene que $\#(A^0) = (\#A)^0 = 1$, ya que, según la definición anterior, $A^0 = \{\varepsilon\}$, donde $\varepsilon$ es la palabra vacía sobre $A$ (es la única palabra sobre $A$ que tiene longitud cero). El siguiente caso corresponde a $i = 1$, donde $\#(A^1) = \#A$. En efecto, si el alfabeto $A$ está formado por $n$ símbolos distinguibles, esto es, $A = \{a_1, a_2, \dots, a_n\}$, entonces el conjunto $A^1$ está formado por $n$ palabras diferentes, esto es, $A^1 = \{a_1, a_2, \dots, a_n\}$, todas las palabras unitarias sobre $A$. De manera rigurosa es importante no confundir $A$ con $A^1$, ya que el primero es un conjunto de símbolos, mientras que el segundo es un conjunto de palabras.
Con base en los resultados anteriores, los casos que requieren una investigación más detallada son los que corresponden a $i > 1$. Estos casos también dependen de la cantidad de símbolos distinguibles que se puedan emplear. En la siguiente sección se determina la cardinalidad de algunos conjuntos $A^i$ particulares, considerando tres alfabetos con cardinalidades diferentes.

### 4.2 Más cálculos combinatorios

Alfabetos con un solo símbolo
La construcción de palabras con un alfabeto de un único símbolo distinguible resulta muy trivial. Por ejemplo, sea el alfabeto $A = \{0\}$ (se debe recordar que se ignora el significado de los símbolos considerados) y sean las palabras $w$ de longitud $n$, donde

$n$ es un número arbitrario. En lo particular, estas palabras tienen la forma general:

$$0_0 0_1 \cdots 0_{n-1}, \text{ donde } 0 \in A. \qquad (14)$$

Esta es una cadena de $n$ "contenedores" en la que en cada contenedor sólo se puede colocar al único símbolo disponible. Así, cada conjunto $A^i$, con $i \geq 0$, está formado por una sola palabra; esto es, se tiene que:

$$A^i = \{0_0 0_1 \cdots 0_{i-2} 0_{i-1}\}, \text{ con } i \geq 0. \qquad (15)$$

El conjunto $A^i$ está constituido por la única palabra formada por $i$ ceros.
Así pues, los casos más interesantes de analizar se forman con alfabetos que tengan dos o más símbolos distinguibles. Algunos de estos casos se analizan a continuación.

Alfabetos con dos símbolos distinguibles
Se considera el caso de un alfabeto formado por dos símbolos distinguibles, por ejemplo $A = \{0, 1\}$. A continuación, se presenta un método para construir las palabras que forman parte de algunos conjuntos $A^i$ sobre el alfabeto anterior.

**1.a.** El primer caso no trivial corresponde a $A^2$, el conjunto de todas las palabras de longitud dos. En lo particular se tiene que $\#(A^2) = (\#A)^2 = 2^2 = 4$; esto es, se tienen cuatro palabras diferentes de longitud dos. Nótese que se ha establecido la cantidad de palabras diferentes que constituyen a $A^2$, pero, por el momento, no se sabe qué palabras son. En este caso, sin embargo, es muy fácil especificar directamente tales palabras:

$$A^2 = \{00, 01, 10, 11\}. \qquad (16)$$

Claramente, dos palabras se forman repitiendo dos veces uno de los símbolos disponibles, mientras que las otras dos palabras se forman con los dos símbolos disponibles, en los dos órdenes posibles. Esto se puede decir de otra manera, las palabras de longitud dos se forman considerando que uno de los símbolos disponibles se repite una vez (con lo cual, el otro símbolo también se debe repetir una vez), o que uno de los símbolos se repite dos veces (con lo que el otro símbolo se repite cero veces).
El resultado anterior es fácil de explicar. Todas las palabras de longitud dos tienen la forma general $x_{i_0} x_{i_1}$, donde $x_i = 0, 1$. Estas son las cadenas de dos "contenedores". Falta considerar qué los símbolos se pueden encontrar en cada contenedor. Están disponibles dos símbolos distinguibles, por lo que se puede presentar el caso en que cada símbolo se encuentre en un contenedor diferente. Y hay dos formas de acomodar los símbolos, o primero se considera al cero y luego al uno, o viceversa. Queda la posibilidad de que ambos contenedores tengan al mismo símbolo; y nuevamente hay dos posibilidades, sólo se usa el cero, o sólo se usa el uno. Es claro que, en total, hay cuatro palabras diferentes, como ya se había establecido.
Con base en la descripción anterior, el conjunto $A^2$ se puede particionar como:

$$A^2 = \{01, 10\} \cup \{00, 11\}. \qquad (17)$$

Cada una de las particiones anteriores se ajusta a una *estructura* particular para las palabras de longitud dos. Por una estructura para una palabra se debe entender una descripción compacta de los símbolos usados, las veces que esos símbolos se repiten y las posiciones que ocupan en una cierta palabra. Las estructuras para palabras van a ser especificadas en dos partes: (a) los símbolos usados y el número de veces que se repiten; y (b) las posiciones (o contenedores) en las que se colocan. Los símbolos usados y sus repeticiones se dice que constituyen el *esqueleto* de la estructura, mientras que las posiciones en las que se colocan constituyen la *realización* de la estructura.
En el caso que nos ocupa, los esqueletos de las estructuras se pueden representar usando la notación $(|w|_0, |w|_1)$, donde $w$ es la palabra bajo análisis; esto es, primero se señalan las repeticiones del símbolo $0$, y luego las repeticiones del símbolo $1$. En lo anterior se debe garantizar que $|w|_0 + |w|_1 = |w|$; en el caso que se analiza se tiene que $|w| = 2$. Debe ser claro que la notación permite señalar tanto a los símbolos que se usan, como la cantidad de veces que se repiten.
Con base en lo anterior, la cantidad de esqueletos diferentes es igual a las soluciones de la ecuación:

$$|w|_0 + |w|_1 = |w|. \qquad (18)$$

El primer esqueleto está caracterizado por las palabras en las que uno, y sólo uno, de los símbolos se repite dos veces (obviamente, el símbolo restante no se usa), mientras que el segundo esqueleto se caracteriza por las palabras en las que cada símbolo aparece una sola vez. Así, los esqueletos de las estructuras para las palabras de longitud dos son:

$$\begin{aligned} &\text{a)} \quad (2,0), (0,2), \\ &\text{b)} \quad (1,1). \end{aligned} \qquad (19)$$

Ahora bien, un mismo esqueleto se puede realizar de varias formas diferentes. Cada uno de los dos primeros esqueletos –apartado (a)– se puede realizar de una sola forma: las dos repeticiones ocupan los dos contenedores. El tercer esqueleto –apartado (b)– se puede construir de dos formas diferentes: uno de los símbolos se coloca en el primer contenedor, y el símbolo restante se coloca en el segundo contenedor. En total, se obtienen cuatro palabras diferentes, como ya se había establecido previamente.

**1.b.** El siguiente caso corresponde al conjunto de todas las palabras de longitud tres, esto es, $A^3$. Ahora se tiene que $\#(A^3) = (\#A)^3 = 2^3 = 8$. Nuevamente, se conoce la cantidad total de palabras, pero se desconoce qué palabras son. Para resolver esto, se procede a determinar los posibles esqueletos de la estructura de las palabras de longitud tres. Según lo que se estableció en la Sección 1.a, la cantidad de esqueletos diferentes es igual a las soluciones de la ecuación:

$$|w|_0 + |w|_1 = |w| = 3. \qquad (20)$$

Esto es, hay que determinar el número de formas en las que 3 se puede expresar como la suma de 2 números. No es difícil establecer que el problema anterior tiene las soluciones siguientes:

a) $(3,0),(0,3)$,
b) $(2,1),(1,2)$. (21)

Así pues, hay cuatro esqueletos diferentes para formar a las estructuras de las palabras de longitud tres. Nótese que para las palabras de longitud dos sólo hay tres esqueletos diferentes. Con esto se puede establecer un resultado parcial: para un mismo alfabeto, al aumentar la longitud de una palabra en una unidad (en este caso de dos a tres), también aumenta en una unidad la cantidad de esqueletos de las estructuras para las palabras (en este caso, de tres a cuatro).

Los esqueletos de la estructura general tienen las características siguientes. Uno de los símbolos se repite cero veces y, por consiguiente, el otro símbolo se repite tres veces; o uno de los símbolos se repite una sola vez y el otro se repite dos veces. Debe ser claro que las palabras de longitud tres tienen un símbolo que se repite más de una vez.

La última afirmación también se puede validar razonando como sigue. Dado que sólo hay dos símbolos distinguibles, únicamente dos de los tres contenedores que forman a la palabra pueden tener símbolos diferentes; el tercer contenedor necesariamente va a tener una repetición de algunos de los símbolos ya usados.

Los esqueletos recién establecidos se pueden realizar como sigue. Se debe recordar que se están considerando palabras de longitud tres o cadenas de tres "contenedores"; estas cadenas tienen la forma general $x_{i_0}x_{i_1}x_{i_2}$, donde $x_i = 0,1$.

- Cada uno de los esqueletos en los que un símbolo que se repite tres veces –apartado (a)–, se puede realizar de una sola forma: en cada uno de los tres contenedores se encuentre el mismo símbolo. Así, se tienen dos palabras diferentes en las que uno de los símbolos se repite tres veces.
- Por su parte, cada uno de los esqueletos en los que un símbolo que se repite dos veces –apartado (b)–, se puede realizar de tres formas diferentes: se deben escoger los dos contenedores en los que se colocan las dos repeticiones; hay tres selecciones posibles. Así, se tienen seis palabras diferentes de longitud tres que tienen dos símbolos repetidos.

Esto explica las ocho palabras diferentes. Con los resultados anteriores, el conjunto $A^3$ se puede particionar de la manera siguiente:

$$A^3 = \{000\} \cup \{111\} \cup \{100,010,001\} \cup \{110,101,011\}. \qquad (22)$$

Ahora es fácil observar que por lo menos uno de los símbolos disponibles se repite dos o más veces en cada una de las palabras.

Alfabeto con tres símbolos distinguibles

Se considera ahora el caso de un alfabeto formado por tres símbolos distinguibles, en particular $A = \{0,1,2\}$. Como en el ejemplo previo, se ilustra el procedimiento para construir las palabras de algunos conjuntos $A^i$ sobre este alfabeto.

**2.a.** Se investiga primero el caso no trivial más simple, esto es, el conjunto $A^2$, el conjunto de todas las palabras de longitud dos sobre $A$. La cardinalidad de este conjunto es $\#(A^2) = (\#A)^2 = 3^2 = 9$. Las palabras de $A^2$ tienen la estructura general $x_{i_0}x_{i_1}$, donde $x_i = 0,1,2$. Claramente, tanto el primer contenedor como el segundo deben tener uno de tres símbolos posibles, lo cual explica la cardinalidad de $A^2$ que se había calculado. La cantidad de esqueletos de esta estructura se encuentra resolviendo la ecuación:

$$|w|_0 + |w|_1 + |w|_2 = |w| = 2. \qquad (23)$$

Esto es, 2 se debe expresar en términos de la suma de tres sumandos. Debe resultar claro que por lo menos uno de los sumandos debe ser nulo. También debe ser claro que ninguna de las palabras de $A^2$ puede estar formada por los tres símbolos disponibles.

Las soluciones de la ecuación anterior, y los esqueletos de la estructura general para las palabras de longitud dos, son:

a) $(2,0,0),(0,2,0),(0,0,2)$,
b) $(1,1,0),(1,0,1),(0,1,1)$. (24)

Nótese que se ha generalizado la notación introducida en el Ejemplo 1. Debe resultar claro que por lo menos uno de los símbolos disponibles no se usa.

En este punto de la presentación se puede señalar otro resultado parcial. Considerando palabras de una misma longitud, en este caso longitud dos, al aumentar en uno la cantidad de símbolos disponibles, también aumenta la cantidad de esqueletos de las estructuras para las palabras: en este caso se pasó de tres esqueletos para dos símbolos a seis esqueletos para tres símbolos; la cantidad se duplicó.

Por lo que respecta a la realización de los esqueletos anteriores se establece lo siguiente.

- Cada uno de los esqueletos del apartado (a) de la lista se puede realizar de una sola manera: en cada contenedor se coloca una de las repeticiones del único símbolo que se utiliza. Se obtienen tres palabras diferentes.
- Por su parte, cada uno de los esqueletos del apartado (b) se puede realizar de dos formas diferentes: en uno de los contenedores se coloca uno de los símbolos, en el contenedor restante se coloca otro símbolo; uno de los símbolos disponibles no se usa. Se obtienen seis palabras diferentes.

En total, se obtienen 9 palabras diferentes. El conjunto $A^2$ se puede especificar directamente como sigue:

$$A^2 = \{00\} \cup \{11\} \cup \{22\} \cup \{01,10\} \cup \{02,20\} \cup \{12,21\}. \qquad (25)$$

Debe resultar claro que las palabras se forman usando un solo símbolo dos veces, o usando dos símbolos diferentes, una vez cada uno.

**2.b.** Se analiza ahora el conjunto $A^3$, esto es, el conjunto de todas las palabras de longitud tres. La cardinalidad de este conjunto es $\#(A^3) = (\#A)^3 = 3^3 = 27$. En este caso también es fácil explicar esta cardinalidad. Todas las palabras de longitud tres se ajustan a la estructura $x_{i_0}x_{i_1}x_{i_2}$, donde $x_i = 0, 1, 2$, y en cada uno de los contenedores se puede colocar uno, y sólo uno, de los tres símbolos disponibles. Por el principio de la multiplicación se obtienen nueve palabras diferentes.
La cantidad de esqueletos para la estructura de las palabras se obtiene resolviendo la ecuación $|w|_0 + |w|_1 + |w|_2 = |w| = 3$. No es difícil establecer que las soluciones de la ecuación dan origen a los esqueletos siguientes:

$$\begin{array}{ll} \text{a)} & (3,0,0), (0,3,0), (0,0,3), \\ \text{b)} & (2,1,0), (2,0,1), (1,2,0), (1,0,2), (0,2,1), (0,1,2), \\ \text{c)} & (1,1,1). \end{array} \qquad (26)$$

La obtención del último esqueleto permite establecer que en $A^3$ se encuentran palabras en las que se usan los tres símbolos disponibles, una sola vez cada uno de ellos. Una situación similar no se había observado en $A^2$. Las restantes palabras de $A^3$ se caracterizan por un símbolo que se repite dos o tres veces.
En este punto se pueden mencionar otros resultados parciales. Primero, para un mismo alfabeto, en este caso de tres símbolos distinguibles, al aumentar la longitud de las palabras en una unidad, en este caso de dos a tres, la cantidad de esqueletos de las estructuras para las palabras también aumenta, pasando de seis a diez. El segundo resultado parcial se adiciona al resultado mencionado en la sección 2.a. Para palabras de una misma longitud, al aumentar la cantidad de símbolos disponibles en uno, la cantidad de esqueletos también aumenta, pero en este caso, las palabras de longitud tres, se pasa de 4 esqueletos con dos símbolos a 10 esqueletos con tres símbolos. En la sección 2.a se había observado que la cantidad se duplicaba, ahora ya no. Debería resultar claro que todavía no se puede hablar de una generalización, y que más casos deben ser investigados con más detalle.
Por lo que respecta a la realización de los esqueletos se encuentra lo siguiente.

- Los tres esqueletos del apartado (a) de la lista se pueden realizar de una sola forma cada uno de ellos. Se obtienen entonces 3 palabras diferentes.
- Por su parte, cada uno de los seis esqueletos del apartado (b) se puede construir de tres formas diferentes. Esto es fácil de explicar, se deben escoger dos contenedores de tres posibles para acomodar el símbolo que se repite dos veces; hay tres formas diferentes de elegir esos dos contenedores. Así, se obtienen 18 palabras diferentes.
- El último esqueleto –apartado (c)– se puede construir de seis formas diferentes, con lo cual se obtienen 6 palabras diferentes.

Tal vez sea conveniente explicar la razón de la última realización. En el primer contenedor se puede colocar uno de los tres símbolos diferentes, pero en el segundo ya sólo se puede colocar uno de los dos símbolos restantes y, en consecuencia, en el tercer contener ya sólo se puede colocar el último símbolo disponible. En este caso, la elección hecha para el primer contener afecta a la elección que se puede hacer para el segundo contenedor, y las dos elecciones realizadas afectan la elección para el tercer contenedor. Claramente, las elecciones ya no son independientes entre sí.
En total, se obtienen 27 palabras diferentes. A pesar de que el conjunto $A^3$ todavía se podría especificar directamente, ya no se presenta.

## COMENTARIOS FINALES

Debe resultar claro que las secuencias finitas y las palabras son construcciones matemáticas no triviales, con propiedades nada simples. Un enfoque exclusivamente intuitivo al tema no permite resaltar tales características. En el documento hemos resaltado que, usando conceptos fundamentales de las matemáticas discretas como conjuntos, números naturales y funciones, se puede estructurar una presentación rigurosa, aunque sencilla, del tema. La rigurosidad empleada permite mostrar las principales propiedades de las secuencias finitas, propiedades que son fundamentales para comprender sus muy diversas aplicaciones y usos en las teorías de los Lenguajes Formales, Autómatas y Computación Teórica. En especial, se destaca la naturaleza combinatoria de las secuencias finitas y palabras y, con ello, se establecen las bases para emprender investigaciones más detalladas sobre métodos de construcción, caracterización y selección de palabras particulares en contextos muy diversos. Estas aplicaciones y usos de las secuencias finitas (y de sus parientes cercanos, las secuencias infinitas) cada vez son más solicitadas en diversas áreas de las matemáticas aplicadas, por ejemplo, criptología y biomatemáticas.

# Referencias Bibliográficas

Cohen, D. I. A. (1986) **Introduction to Computer Theory.** New York, John Wiley & Sons, Inc.
Epp, S. S. (2012) **Matemáticas Discretas con Aplicaciones.** Cuarta edición, México, Cengage Learning.
Fejer, P. A. and Simovici, D. A. (1990) **Mathematical Foundations of Computer Science. Volume 1: Sets, Relations, and Induction.** Texts and Monographs in Computer Science, Volume X, New York, Springer-Verlag.
Herrera Becerra, Alberto A. (2021). **Palabras y Operaciones para Palabras. Teoría y Ejemplos.** Notas de Curso, Instituto de Ciencias Aplicadas y Tecnología, Universidad Nacional Autónoma de México, NC-INME-2021-572.
Kelley, D. (1995) **Teoría de Autómatas y Lenguajes Formales.** Madrid, Prentice Hall.
Mateescu, A. and Salomaa, A. (1997) Formal Languages: an Introduction and a Synopsis. En: Rozenberg, G. and Salomaa, A., eds. **Handbook of Formal Languages, Volume 1: Word, Language, Grammar.** Berlin-Heidelberg, Springer-Verlag, p. 1-39.
Real Academia Española (2001). **Diccionario de la Lengua Española**, vigésima segunda edición. Madrid, España, Editorial Espalsa Calpe, S. A.

AÑO 8. No. 8. ISSN-2448-7236. SEPTIEMBRE 2023 – AGOSTO 2024. MD-01, pág.: 1 a la 10.

Tremblay, J. P. and Manohar R. (1987) **Discrete Mathematical Structures with Applications to Computer Science.** New York, McGraw-Hill Book Company.

## INFORMACIÓN ACADÉMICA

**Alberto Arturo Herrera Becerra:** Químico egresado de la Facultad de Química de la UNAM, Maestro en Investigación Biomédica Básica egresado de la UACPyP del CCH de la UNAM y candidato a Doctor en Química por parte de la UAM-Iztapalapa. Profesor de Asignatura en el Departamento de Computación de la División de Ingeniería Eléctrica de la Facultad de Ingeniería de la UNAM.

**Francisco José Rodríguez Ramírez:** Ingeniero Mecánico Electricista de la Facultad de Ingeniería de la UNAM. Diploma en Microprocessors Engineering and Digital Electronics, egresado del University of Manchester Institute of Science and Technology (UMIST), Manchester, Inglaterra. Profesor de Tiempo Completo en la Facultad de Ingeniería de la UNAM.