Universidad Autónoma del Estado de México

Centro Universitario UAEM Zumpango

Ingeniería en Computación

Autómatas y Lenguajes Formales

Apuntes Complementarios del Curso

Unidad de Competencia 1

Representar información a través del uso de grafos

M. en C. Rafael Rojas Hernández

Zumpango, México Agosto, 2015

Índice general

Índice de tablas

Índice de figuras

Presentación

Estos apuntes complementarios a la Unidad de Aprendizaje "Autómatas y Lenguajes Formales", tienen como propósito brindar a los alumnos, además del contenido, conceptos principales y ejercicios ilustrativos de cada uno de los temas; con la finalidad de que amplíen y reafirmen sus conocimientos.

Se pretende que el alumno aprenda los conocimientos de la Unidad de Competencias 1 "Representar información a través del uso de grafos" del programa de estudios permitiéndole acceder a las siguientes unidades y cursos posteriores a esta materia.

En este documento se incluyen conceptos, definiciones, ejemplos y ejercicios, se estudian temas como: terminología básica (introducción a los grafos, conceptos y representación de los grafos) y la aplicación a máquinas de estado finito de los autómatas (historia y usos). También se incluye un glosario de términos y la bibliografía básica.

Introducción

Dentro de las ciencias básicas las matemáticas son una herramienta clave para hallar una exllicación a diversos fenómenos, ya que esta escrita en un lenguaje que el mundo entiende.

Uno de los hechos que se destaca en la informática es que las áreas genéricas del conocimiento humano como es la lógica y el álgebra, han tenido que especializarse, o particularizarse para ser utilizados en ésta, de aquí surge el uso de la lógica matemática, lógica de conjuntos, teoría de grafos, entre otros, para su aplicación en las ciencia de las computación; extendiéndose en tantas direcciones como la teoría del lenguaje o el no determinismo así como las expresiones regulares y las gramáticas libres de contexto.

Todo ello con el fin de llegar a entender, desde su partes más básicas, la teoría del funcionamiento de las computadoras. Los modelos de como deben ser programadas o como es que funcionan los programas.

Uan parte importante es conocer el término de autómata como uno de los principales objetos dentro del mundo real y como es visto dentro del área de la computación, además de sus analogías con los lenguajes, procesos, caminos, programas, entre otros.

Unidad 1

Representar información a través del uso de grafos

1.1. Terminología básica

1.1.1. Introducción a los grafos

Gráficamente un grafo es un conjunto de puntos (llamados vértices o nodos), unidos por una o más líneas (aristas, lazos o caminos). El uso de grafos permite el estudio de las interrelaciones entre las unidades que se encuentran en interacción, con ejemplos se tienen los de la figura 1.1.

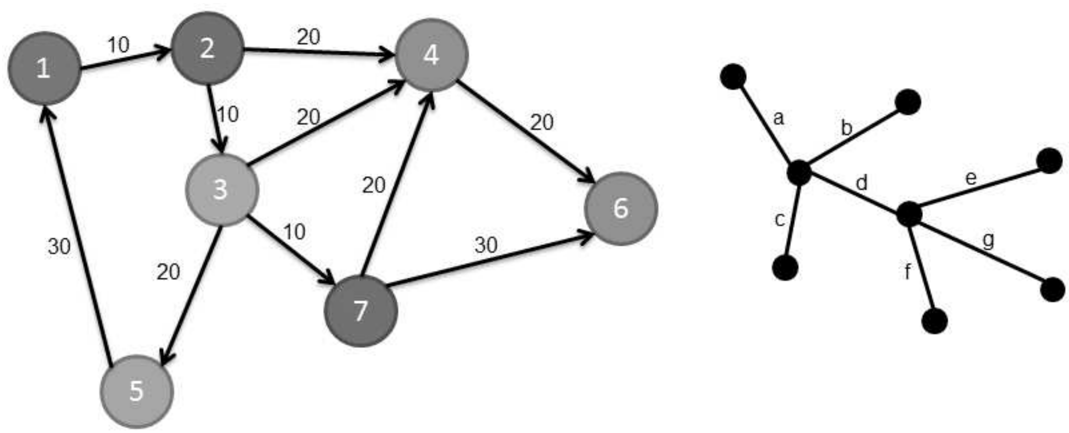


Figura 1.1: Ejemplos de grafos.

Definición 1 *Un **grafo simple** $G(V, E)$ consiste de un conjunto no vacío V de vértices o nodos, y un conjunto E de pares no ordenados de distintos elementos de V llamados aristas.*

Los elementos comunes dentro de un grafo puede observarse en la figura 1.2.

Definición 2 *Un **multigrafo** $G(V, E)$ consiste de un conjunto V de vértices, un conjunto E de aristas, y una función f de E en $\{u, v | u, v \in V, u \neq v\}$.*
Las aristas e_1 y e_2 son llamadas aristas múltiples o paralelas si $f(e_1) = f(e_2)$.

La figura 1.3 es un ejemplo de un multigrafo con aristas paralelas.

Definición 3 *Un **pseudografo** $G(E, V)$ consiste de un conjunto V de vértices, un conjunto E de aristas, y una función f de E en $\{\{u, v\} | u, v \in V\}$.*

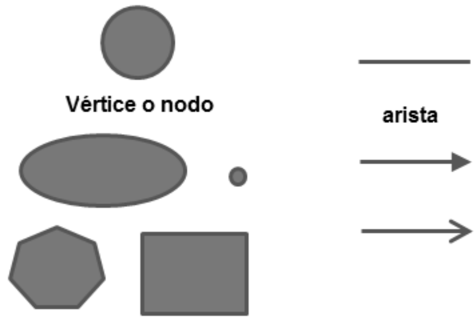


Figura 1.2: Elementos de un grafo simple.

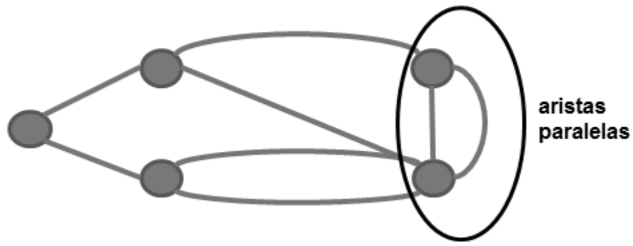


Figura 1.3: Elementos de un grafo simple.

Definición 4 *Una arista es un lazo o bucle si $f(e) = \{u, u\} = \{u\}$ para algún $u \in V$.*

En la figura 1.4 se observa el ejemplo de los pseudografos y un lazo.

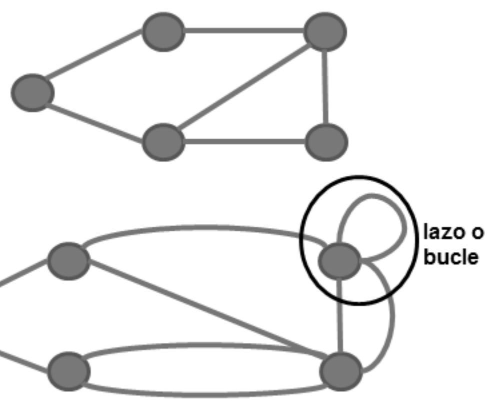


Figura 1.4: Pseudografo y lazo.

Definición 5 *Un* ***grafo dirigido*** *o* ***digrafo*** *$G(E, V)$ consiste de un conjunto de vértices V, y un conjunto de aristas E, que son pares ordenados de los elementos de V.*

Para representar un grafo dirigido es necesario que las aristas sean representadas con flechas, como se ve en la figura 1.5.

1.1.2. Conceptos

Definición 6 *Dos vértices u y v en un grafo no dirigo G son llamados* ***vértices adyacentes*** *(o vecinos) en G si $\{u, v\}$ están en una arista de G. Si $e = \{u, v\}$, la arista e es llamada incidente con el vértice u y v.*

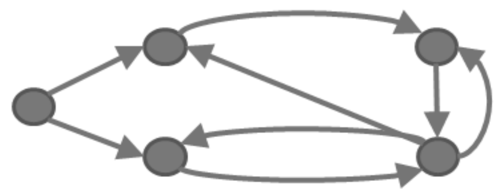


Figura 1.5: Grafo dirigido.

La arista e es también llamada como conectada con u y v. Los vértices u y v son llamados puntos finales de las aristas $\{u, v\}$.

Los vértices adyancentes, como ejemplo se puede ver en la figura 1.6.

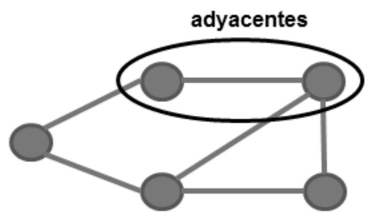


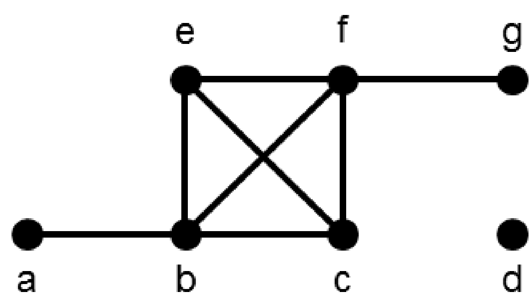
Figura 1.6: Pseudografo y lazo.

Definición 7 *El* ***grado*** *de un vértice es el número de aristas que inciden en este, y es denotado por* $gr(v)$ *(degree o* $deg(v)$*).*

Un lazo de un vértice contribuye dos veces el grado del mismo, figura 1.7.

Figura 1.7: Grado de un vértice.

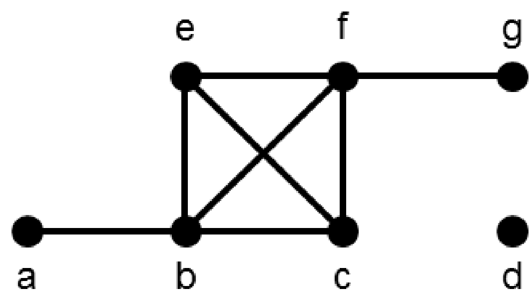
Ejemplo 1 *Grado de un vértice.*



$$gr(a) = 1$$
$$gr(b) = 3$$
$$gr(c) = 3$$
$$gr(d) = 0$$

$$gr(e) = 3$$
$$gr(f) = 4$$
$$gr(g) = 1$$

Ejemplo 2 *Grado de un vértice.*



$$gr(a) = 6$$
$$gr(b) = 5$$
$$gr(c) = 4$$
$$gr(d) = 4$$
$$gr(e) = 3$$

Definición 8 *Cuando (u, v) es una arista del grafo G con aristas dirigidas, u se dice que es adyacente a v y v se dice que es adyacente de u. El vértice u es llamado el **vértice inicial** de (u, v), y v es llamado el **vértice final** o **terminal** de (u, v). El vértice inicial y final en un lazo es el mismo.*

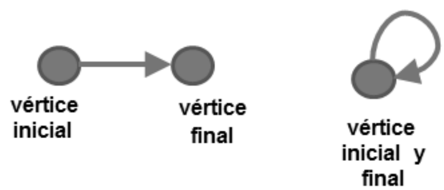
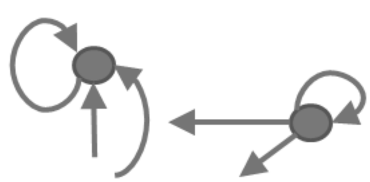


Figura 1.8: Grados.

Definición 9 *En un grafo con aristas dirigidas el **grado entrante** (indegree) de un vértice v, denotado por $gr^-(v)$ (o $deg^-(v)$), es el número de aristas con v en su vértice final. El **grado de salida** (outdegree) de v, denotado por $gr^+(v)$ ($deg^+(v)$), es el número de aristas con v como su vértice inicial.*

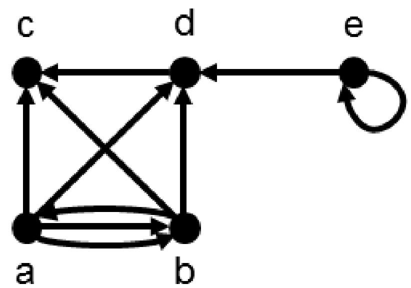
Para identificar la diferencia entre un grado de entrada y uno de salida puede observarse la figura 1.9.

Ejemplo 3 *Grado de un vértice.*



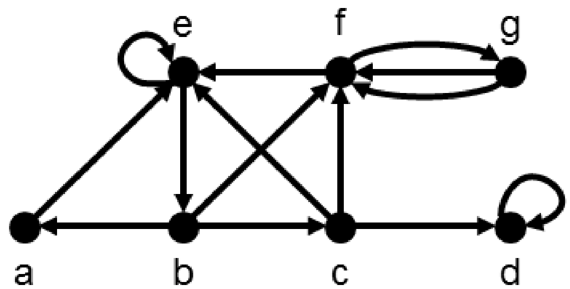
(a) Entrada (b) Salida

Figura 1.9: Grados.



$gr^{-}(a) = 1$ $gr^{+}(a) = 4$
$gr^{-}(b) = 2$ $gr^{+}(b) = 3$
$gr^{-}(c) = 3$ $gr^{+}(c) = 0$
$gr^{-}(d) = 3$ $gr^{+}(d) = 1$
$gr^{-}(e) = 1$ $gr^{+}(e) = 2$

Ejemplo 4 *Grado de un vértice.*



$gr^{-}(a) = 1$ $gr^{+}(a) = 1$
$gr^{-}(b) = 1$ $gr^{+}(b) = 3$
$gr^{-}(c) = 1$ $gr^{+}(c) = 3$
$gr^{-}(d) = 2$ $gr^{+}(d) = 1$
$gr^{-}(e) = 4$ $gr^{+}(e) = 2$
$gr^{-}(f) = 4$ $gr^{+}(f) = 2$
$gr^{-}(g) = 1$ $gr^{+}(g) = 2$

Definición 10 *Un* ***subgrafo*** *de un grafo $G(V, E)$ es un grafo $H(W, f)$ donde $W \in V$ y $F \in E$.*

La forma de un subgrafo a partir de un grafo se observa en la figura 1.10

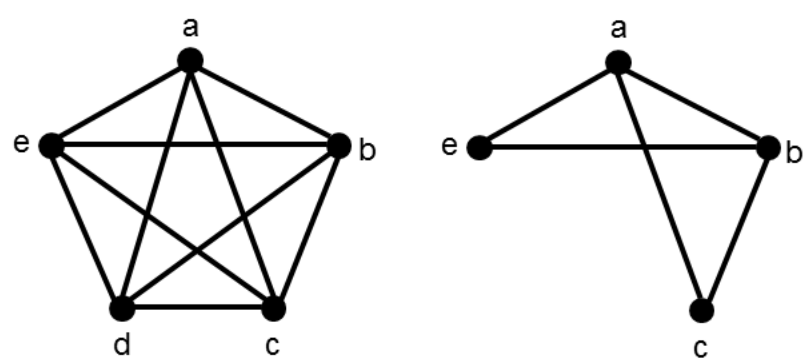


Figura 1.10: Subgrafo.

Definición 11 *La* ***unión*** *de dos grafos $G_1(V_1, E_1)$ y $G_2(V_2, E_2)$ es un grafo con el conjunto de vértices $V_1 \cup V_2$ y el conjunto de aristas $E_1 \cup E_2$. La unión de G_1 y G_2 es denotada por $G_1 \cup G_2$.*

En la figura 1.11 na unión de dos grafos puede comprobarse.

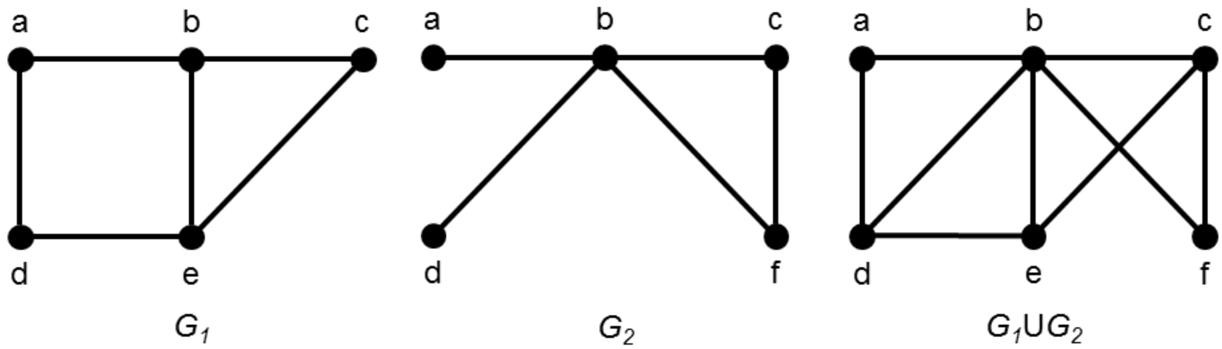


Figura 1.11: Union de 2 grafos.

Definición 12 *Los grafos simples $G_1 = (V_1, E_1)$ y $G_2 = (V_2, E_2)$ son isomorfos si existe una función uno a uno f de V_1 a V_2 con la propiedad de que a y b son adyacentes en G_1 si y sólo si $f(a)$ y $f(b)$ son adyacentes en G_2, para todas las a y b en V_1. La función f es llamada isomorfismo.*

El ejemplo de isomorfismo esta presente en la figura 1.12.

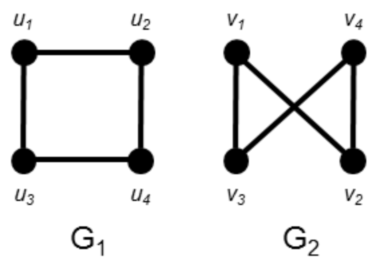


Figura 1.12: Isomorfismo entre 2 grafos.

Definición 13 *Un camino de longitud n de u a v, con n entero positivo, en un grafo no dirigido es una secuencia de aristas $e_1, \ldots, e_n$ en el grafo, tales que $f(e_1) = \{x_0, x_1\}, f(e_2) = \{x_1, x_2\}, \ldots, f(e_n) = \{x_{n-1}, x_n\}$, donde $x_0 = u$ y $x_n = v$..*
Un camino es un circuito si este comienza y termina en el mismo vértice.
Un camino de longitud n de u a v, con n entero positivo, en un multigrafo dirigido es una secuencia de aristas $e_1, \ldots, e_n$ en el grafo, tales que

$$f(e_1) = \{x_0, x_1\}, f(e_2) = \{x_1, x_2\}, \ldots, f(e_n) = \{x_{n-1}, x_n\}$$

donde $x_0 = u$ y $x_n = v$.

Definición 14 *Un grafo no dirigido es llamado conectado si existe un camino entre cada par de vértices distintos del grafo.*

En la figura 1.13 se ejemplifica el ejemplo de la definición de grafos conectado (H) y no conectado (G).

Definición 15 *Un grafo completo de n vértices, denotado por K_n, es un grafo simple que contiene exactamente una arista entre cada par de vértices distintos.*

Los grafos conectados para 1 hasta 5 vértices se muestra en la figura 1.14

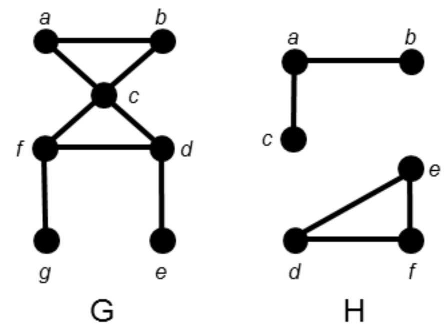


Figura 1.13: Isomorfismo entre 2 grafos.

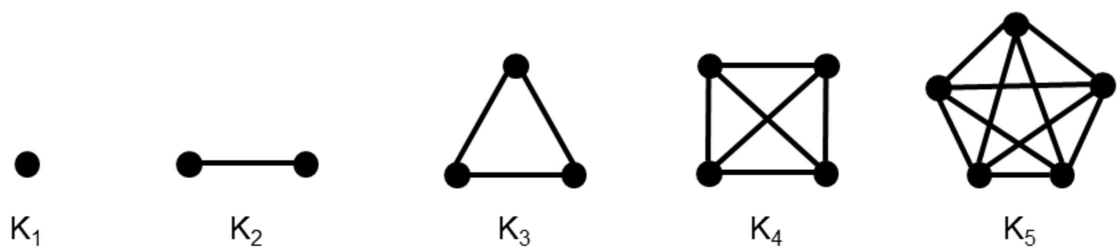


Figura 1.14: Grafos completos.

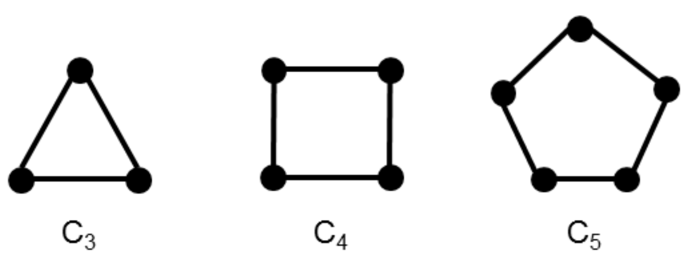


Figura 1.15: Ciclos en un grafo.

Definición 16 *El ciclo C_n, $n \geq 3$, consiste de n vértices $v_1, v_2, \ldots, v_n$ y aristas $\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{n-1}, v_n\}$, y v_n, v_1.*

Para los primeros tres ciclos se tiene la figura 1.15.

Definición 17 *Para obtener la rueda W_n cuando se agrega un vértice a un ciclo C_n, para $n \geq 3$, este se conecta a cada uno de los n vértices en C_n.*

Las tres primeras ruedas en un grafo se observan en la figura 1.16.

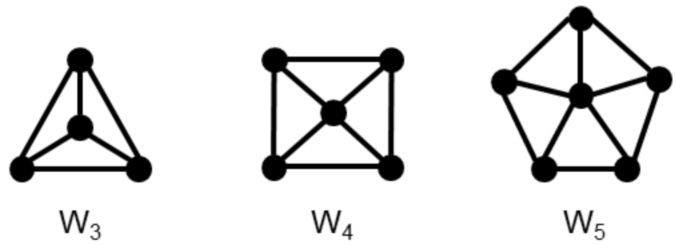


Figura 1.16: Ruedas en un grafo.

Definición 18 *El n-Cubo, denotado por Q_n, es el grafo que tiene a sus vérticer representando los 2^n bits en cadenas de longitud n.*

los grafos n-Cubos para $n = 1, 2, 3$ se muestran en la figura 1.18.

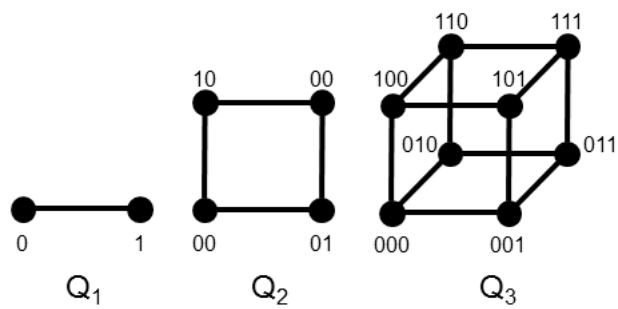


Figura 1.17: Grafos n-Cubos.

1.1.3. Representación de los grafos

Definición 19 *Sea $G(V, E)$ un grafo no dirigido, con vértices $v_1, v_2, \ldots, v_a$ y aristas $e_1, e_2, \ldots, e_b$, la **matriz de incidencia** de G es la matriz $M(G) = [m_{ij}]$ de dimensiones $a \times b$ dada por:*

$$m_{ij} = \begin{cases} 1 & \textit{cuando la arista } e_j \textit{ es incidente en } v_i \\ 0 & \textit{en otro caso} \end{cases}$$

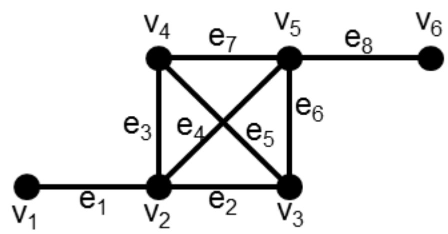
La manera en como se asignan columnas y filas se observa en la tabla 1.1.

Ejemplo 5 *Matriz de incidencia $M(G)$*

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
v_1	1	0	0	0	0	0	0	0
v_2	1	1	1	1	0	0	0	0
v_3	0	1	0	0	1	1	0	0
v_4	0	0	1	0	1	0	1	0
v_5	0	0	0	1	0	1	1	1
v_6	0	0	0	0	0	0	0	1

Tabla 1.1: Tabla de incidencias

	e_1	e_2	$\vdots$	e_b
v_1				
v_2				
$\dots$				
v_a				



Definición 20 *Sea $G(V, E)$ un grafo no dirigido, con vértices $v_1, v_2, \dots, v_a$, la* ***matriz de adyacencia*** *de G es la matriz $A(G) = [a_{ij}]$ de dimensiones $a \times a$ dada por:*

$$a_{ij} = \begin{cases} 1 & \textit{if } \{v_i, v_j\} \textit{ es una arista de } G \\ 0 & \textit{en otro caso} \end{cases}$$

La manera en como se asignan columnas y filas se observa en la tabla 1.2.

Tabla 1.2: Tabla de adyacencia

	v_1	v_2	$\dots$	v_a
v_1				
v_2				
$\vdots$				
v_a				

Ejemplo 6 *Matriz de adyancia $A(G)$*

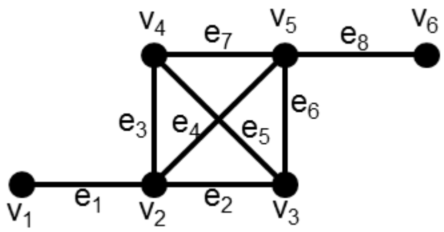
	v_1	v_2	v_3	v_4	v_5	v_6
v_1	*0*	*1*	*0*	*0*	*0*	*0*
v_2	*1*	*0*	*1*	*1*	*1*	*0*
v_3	*0*	*1*	*0*	*1*	*1*	*0*
v_4	*0*	*1*	*1*	*0*	*1*	*0*
v_5	*0*	*1*	*1*	*1*	*0*	*1*
v_6	*0*	*0*	*0*	*0*	*1*	*0*

1.2. Aplicación a máquinas de estado finito

Para entender el origen de las cosas es necesario, en primer lugar, como es que se llego al punto en el cual estamos. Por lo que en esta sección hablaremos de como hemos se relaciona la téoria de autómatas con los lenguajes formales, el mundo de la programación y las máquinas de estado finito.

1.2.1. Historia

Desde el inicio de los tiempos el ser humano a desarrollado teorías o modelos de como es el funcionamiento de las cosas, ya sean procesos naturales o herramientas, máquinasy objetos creados por él. Muchas de las



disciplinas de las ciencias se basan en principios que provienen de las partes más comunes de la naturaleza como la física, química, botánica, matemáticas o incluso la lingüística.

Gracias a la interrelacón entre las ciencias se comenzaron a desarrollar artefactos y equipos que apoyaron al hombreen sus tareas diarias, sin embargo no podian ser creadas de manera que sustituyeran cien por ciento al hombre. Con el paso de los siglos, los inventos (o máquinas) se fueron perfeccionando hasta llegar a un punto que podían realizar las actividades de varios hombres al mismo tiempo, más rápido y de manera más eficiente. Sin embargo eran equipos mecánicos que necesitaban ser impulsados por animales o el hombre para su funcionamiento.

Con la llegada de las máquinas de vapor se fueron realizando acciones más complejas, pero con equipos a gran escala, sin embargo esto cambio al inicio del siglo XX, con el auge de la electricidad. Ya que con ella fue posible disminuir el tamaño de lo objetos.

Al poder crear máquinas eléctricas estás cada fueron más complejas y comenzaron a ser utilizadas para funciones poco convencionales o como apoyo a las ciencias, por ejemplo las cálculadoras. Sin embargo existía campos en los cuales no era posibles utilizarlas, y simplemente se desarrollaban toerías para su uso en la resoluci´on de problemas.

Uno de ellos era la manera en como comunicarse con las máquinas para que una sóla entendiera diferentes instrucciones, con ello se pretendía utilizar sólo un dispositivo para realizar diversas tareas sin cambiar su configuración física.

Ya que la manera que utiliza el hombre para dar instrucciones a otros o incluso a los animales que utiliza es a través de palabras, las cuales son agrupadas como un lenguaje que sólo es entendido por quién lo conoce, se consideró que esta sería una forma adecuada de hacerlo, sin embargo el problema no era tan sencillo.

Para lograr dicho objetivo investigadores comenzaron a entrelazar diversas áres como la ingeniería, las matemáticas y la lngüística. Generando nuevas disciplinas como: autómatas, computabilidad y lenguajes y gramáticas, respectivamente; todo con el fin de desarrollar una máquina que reciba instrucciones o que sea pensante.

La teoría de autómatas surge dentro del campo de la ingeniería eléctrica, utilizada para aplicaciones de los circuitos eléctricos combinatorios y secuenciales, en matemáticas en la resolución de problemas, teoría de grafos; y para lingüística los fundamentos para la generación de lenguajes y gramáticas, figura **??**.

Dentro del área de las matemáticas se tienen bases en el término *computabilidad* que se refiere a las operaciones o cálculos que puede realizar una máquina. Existen una gran cantidad de trabajos elaborados al respecto, pero el de mayor impacto fue desarrollado por Alan Mathison Turing (1912-1954) que mostró la teoría para identificar que problemas podrían ser computables, además de que invento la **Máquina de Turing**, modelo básico de las computadoras actuales. La Máquina de Turing constaba de una unidad de procesamiento y en una memoria.

Otro trabajos científicos fueron también de suma importancia, por citar algunos, la demostración de problemas indecidibles y las capacidades algortmicas de Alonzo Church (1903-1995); la teoría de funciones recursivas y el desarrollo de espresiones regulares de Stephen Kleene (1909-1994); Giuseppe Peano (1858-1932) con sus fundamentos de lógica; David Hilbert (1862-1943) y los Principios de lógica teórica; Kurt G´odel(1906-1978) proponiendo el Teorema de Imcompletud;

Dentro de la ingeniería eléctrica se tiene el desarrollo de la lógica de primer orden u operadores lógicos por Frederich L. Gottlon Frege (1848-1925); la Ley de Moore y las investigaciones en transistores de Gordon E. Moore(1929-); redes nueronales artificiales de Warren McCulloch(1898-1969) y Walter Pitts(1923-1969); y la Teoría de la Información de Clause E. Shannon(1916-2001).

Noam Chomsky(1928-) y el origen de la estructura l"ogica de la teoría ling´uística computacional; las

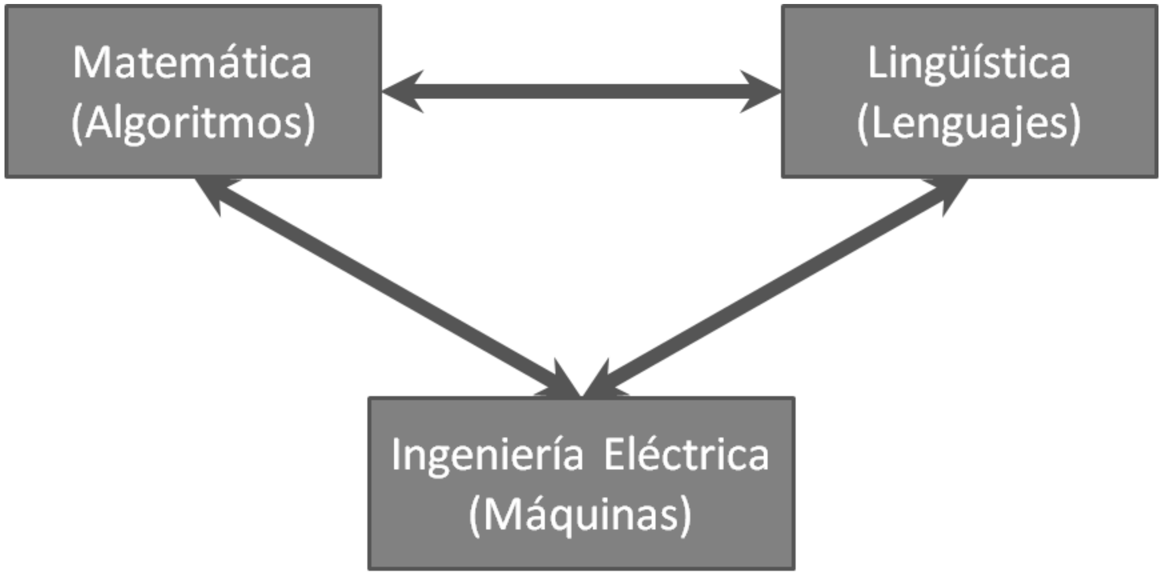


Figura 1.18: Relación Matemáticas-Lin´uística-Ingeniería

formas normales para gramáticas de John Backus(1924-2007); dentro del campo de la ling´uística.

Al unirse las tres ciencias mencionadas anteriormente surge la actual Teoría de Autómatas, que si bien explica el funcionamiento de las máquinas para realizar cálculos, a sido pilar para el desarrollo de una infinidad de métodos y algoritmos utilizados por otras ciencias.

1.2.2. Usos

Si bien la teoría de autómatas ha encontrado aplicación en muy diversos campos, no es dificil saber el porque. Si consideramos que todos los problemas pueden ser tratados como máquinas secuenciales que son capaces de procesar información, un autómata es la manera más natural de considerarlo.

Esto es similar a cualquier sistema existente en la naturaleza, que recibe señales de su entorno, reacciona ante ellas y emite nuevas senñales al ambiente que lo rodea.

Algunos de los campos másrepresentativos donde se encuentra el uso de la teoría de autómatas son:

- Teoría de la Comunicación. Involucra todo lo relevanta los procesos y medios de comunicación de información.
- Teoría de Control. Modelos donde se involucra la toma de desiciones y control de máquinas eléctricas.
- Lógica de circuitos secuenciales. Para el desarrollo de la electrónica digital.
- Arquitectura de computadoras. Diseño y desarrollo de componentes para computadoras personales y dispositivos electrónicos.
- Redes de computadoras. Equipos y protocolos para la interconexión de equipos de cómputo.
- Reconocimiento de patrones. Ayuda en el reconocimiento de obtetos a través de sus rasgos característicos.
- Redes nueronales. Desarrollo de algoritmos de redes nueronales artificiales.
- Reconocimiento y procesamiento de lenguaje. Métodos para entender el habla por una máquina computadora.
- Traducción de lenguajes. Automatización de las equivalencias de palabras entre lenguajes.

- Teoría lógica de los sistemas evolutivos. Simulaciones en el área de las ciencias m"edicas para concer el comportamiento de sistemas reproductivos.

Glosario

- **Alfabeto**. Un alfabeto es un conjunto finito no vacío de símbolos.
- **Autómata Finito Determinista (AFD)** Es aquel en donde es posible determinar claramente cuál es el estado siguiente..
- **Arista**. Son las líneas que unen un vértice con otro y se les asigna una letra, un número o una combinación de ambos.
- **Autómata finito no determinista (AFN)**. Es aquel en donde la función de estado siguiente, no conduce a un estado único determinado.
- **Cadena**. Una cadena es una secuencia finita de símbolos partenecientes a un alfabeto.
- **(Cadena vacía**. Cadena que contiene cero símbolos.
- **Composiciones**. Conjunto de reglas que se deben usar para la estructuración de las palabras válidas en el lenguaje.
- **Expresiones Regulares**. Forma de expresar los lenguajes regulares con la finalidad de facilitar la manipulación y simplificación de los mismos.
- **Grafo**. Diagrama que consta de un conjunto de vértices y un conjunto de aristas.
- **Gramática**. Están integradas por el alfabeto y las composiciones para la estructuración correcta de las palabras validas en un lenguaje.
- **Lenguaje**. s un conjunto de símbolos (o palabras) y métodos para estructurar y combinar dichos símbolos.
- **Máquina de estado finito**. Es una forma especial de representar los autómatas finitos, en donde no existen estados aceptados y donde los símbolos de salida se colocan juntamente con los símbolos de entrada en cada una de las aristas de la máquina.
- **Máquina de Turing (MT)**. Consiste de una cinta que se extiende de manera infinita en donde se escribe o se lee información por medio de una cabeza de lectura-escritura.
- **Vértice**. También llamados nodos son círculos para representar en forma gráfica los elementos de un conjunto en un grafo.

Bibliografía

[1] Hopcrof John E., Ullman Jeffrey D; "Introducción a la teoría de autómatas, lenguajes y computación", CECSA, México 2000.

[2] Ayres Frank Jr., "Álgebra Moderna", Mc Graw Hill, México 1991.

[3] Barco Gómez C., Barco Gómez G., Aristazábal Botero W., "Matemática Digital", Mc Graw Hill, Colombia. 1998.

[4] Grossman W.- Jerrold, "Discrete Mathematics an Introduction to Concepts, Methods and Applications", Mac Millan Publishing Co., USA, 1992.

[5] Jonhnsonbaugh R., "Matemáticas Discretas", Grupo Editorial Iberoamérica, México, 1988.

[6] Kelley Dean; "Teoría de Autómatas y Lenguajes Formales", Prentice Hall, España, 1995.

[7] Kolman Bernard, Busby Robert C., "Estructuras de Matemáticas Discretas para la Computación", Prentice Hall, México, 1997.

[8] Manno Morris, "Lógica Digital", Prentice Hall.

[9] Ross Kenneth A., et.al., "Matemáticas Discretas", Prentice Hall, México, 1990.

[10] Harrison, "Introduction to witching and automata theory"; Mc Graw Hill,

[11] Kosen, D., "Automata and Computability", Springer Verlag, 1997.

[12] Brookshear, "Theory of Computation; Formal Languages, Automata and Complexity", Cummings, 1989.