# Índice.

Índice ………………………………...…………………………………………………………………. 1

Introducción…………………………………………………………………………………………...…2

Desarrollo………………………………………………………………………………………………...3

Lenguajes formales……………………………………………………………………………………3

Gramática formal……………………………………………………………………………………...4

Jerarquía de Chomsky…………………………………………………………………………………4

Conclusiones……………………………………………………………………………………………..6

Referencias……………………………………………………………………………………………….7

# Introducción.

En el año 1931 Kurt Gödel publicó su famoso articulo, que tal vez sea el descubrimiento más importante del siglo XX.

En pocas palabras, el teorema de Gödel dice lo siguiente:

Toda formulación axiomática consistente de la teoría de números contiene proposiciones indecidibles.  
- Teorema de la incompletud, Kurt Gödel 1931.

Esto hizo que cualquier sistema completo y consistente que fuese capaz de demostrar cualquier teorema terminara.

En 1937 Alan Mathison Turing publicó un articulo sobre los números calculables, que desarrolló el teorema de Gödel y se considera como el origen de la informática teórica. En el artículo introdujo la *maquina de Turing* (una entidad matemática abstracta que formalizó el concepto de algoritmo). El teorema que Turing publicó demostraba que existen problemas irresolubles (o sea, que ninguna máquina de Turing es capaz de obtener su solución).

En 1938 el artículo de Claude Elwood Shannon vino a establecer las bases para la aplicación de la lógica matemática a los circuitos combinatorios y secuenciales. A lo largo de las décadas, las ideas de Shannon se convirtieron en la teoría de las máquinas secuenciales y de los autómatas finitos.

En sí los autómatas son sistemas que transmiten información. Es decir, todo sistema que acepta las señales de su entorno, y como resultado cambia de estado y transmite una señal al medio. Se puede decir que un autómata es cualquier máquina, una central telefónica, una computadora o los mismos seres humanos son en sí autómatas. Aunque este concepto es muy general para el estudio teórico, por lo que se le tiene que meter limitaciones a esta definición.

Las aplicaciones de los autómatas tienen campos de acción muy variados, pero tienen en común que esos campos manejan los conceptos de control, acción y memoria.

Estos son algunos ejemplos donde se ha encontrado una aplicación de la Teórica de los Autómatas.

* Teoría de la comunicación.
* Teoría del control.
* Lógica de los circuitos secuenciales.
* Computadoras.
* Redes de computadoras y codificadoras.
* Reconocimiento de patrones.
* Fisiología del sistema nervioso.
* Estructura y análisis de los lenguajes de programación para computadoras.
* Traducción automática de lenguajes.
* Teoría algebraica de lenguajes.

En la década de los 50, el lingüista norteamericano Avram Noam Chomsky revolucionó su campo de actividad con la teoría de las gramáticas transformacionales, que estableció las bases de la lingüística matemática y proporcionó una herramienta que facilitó considerablemente el estudio y la formalización de los lenguajes de computadora.

El estudio de los lenguajes se divide en:

* **Gramática**: Análisis de la estructura de las frases.
* **Semántica**: Su significado.

A su vez, la gramática puede analizar las formas que toman las palabras (**morfología**), su combinación para formar frases correctas (**sintaxis**) y las propiedades del lenguaje hablado (**fonética**).

La teoría de lenguajes formales resultó tener una relación sorprendente con la teoría de máquinas abstractas. Los mismos fenómenos aparecen independientemente en ambas disciplinas y es posible establecer correspondencias entre ellas.

Chomsky clasificó las gramáticas y lenguajes formales de acuerdo con una jerarquía de cuatro grados los cuales se verán a continuación.

# Desarrollo.

## Lenguajes formales.

En esta parte se verán los temas relacionados con los lenguajes formales, sus definiciones y operaciones.

Un lenguaje formal es un conjunto (finito o infinito) de cadenas finitas de símbolos primitivos. Por ejemplo, el lenguaje "número" es el conjunto infinito de cadenas finitas formadas por los digitos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.

### Definiciones básicas.

* **Alfabeto**: Σ; conjunto no vacío finito de símbolos.  
   Ejemplos: el alfabeto español, el inglés, el alfabeto de los números, el alfabeto formado por todos los símbolos del teclado de un ordenador, el alfabeto formado por los símbolos {a, e, i, o, u}.
* **Palabra o cadena**: Secuencia finita de símbolos de un alfabeto.  
   Ejemplos: "palabra", "autómatas", "lenguajes", "sistemas".
* **Longitud de una palabra**: *|x|*; dada una palabra formada por x formada por símbolos de un alfabeto Σ, su longitud es el número de símbolos del alfabeto que la forman.  
   Ejemplos: |palabra|=7 en el caso que los símbolos sean del alfabeto español.

### Operaciones con palabras.

* **Concatenación**: si *x* y *y* son palabras, la concatenación de *xy* es una palabra formada por los símbolos de *x* seguidos de los símbolos de *y*.  
   Ejemplo: En Σ1, si *x=pa* y *y=labra*, *xy=palabra*.
* **Potencia**: La potencia i-ésima de una palabra *x, x^i*, se forma por la concatenación *i* veces de *x*.  
   Ejemplo: si *x=pala*, *x²=palapala*.
* **Reflexión**: Si la palabra *x* está formada por los símbolos A1, A2, ..., An, entonces la palabra inversa *x^-1 = An, ..., A2, A1.*  
   *Ejemplo: En Σ1, si* x=pala\_, *x^-1 = lapa*.

### Operaciones con lenguajes.

* **Unión**: Si L1 y L2 son lenguajes, su unión, L1UL2, contendrá las palabras que pertenezcan a cualquiera de los dos lenguajes.
* **Intersección**: Si L1 y L2 son lenguajes, su intersección L1nL2, contendrá todas las palabras que pertenezcan a los lenguajes.
* **Resta**: Si L1 y L2 son lenguajes, la resta de L1 y L2 será L1 - L2. Contendrá todas las palabras que pertenezcan a L1 y no pertenezcan a L2.
* **Concatenación**: Dados dos lenguajes L1 y L2, su concatenación contendrá todas las palabras que se pueden formar.
* **Potencia**: La potencia i-ésima de un lenguaje corresponde a la concatenación i veces del lenguaje con él mismo.

## Gramática formal.

Una gramática formal está dada por cuatro elementos G:= (V,Σ,S,P), donde:

* **V**: Es un conjunto finito llamado alfabeto de símbolos no terminales o, alfabeto de variables.
* **Σ**: Es otro conjunto finito, que verifica V∩Σ =∅ y se suele denominar alfabeto de símbolos terminales.
* **S∈ V**: Es una variable distinguida que se denomina variable inicial.
* **P ⊆(V∪Σ)+×(V∪Σ)∗** es un conjunto finito llamado conjunto de producciones(o, simplemente, sistema de re escritura).

La forma de definir una gramática solo tiene sentido despues de darle una utilidad. Seguidamente veremos la forma de operar sobre palabras usando una gramática, esto es, la dinámica asociada a una gramática. Para poder definir la dinámica asociada a una gramática, necesitamos definir como operar con los símbolos no terminales.

## Jerarquía de Chomsky.

Chomsky pretende la modelización de los lenguajes (formales y naturales) mediante gramáticas en su trabajo.

### Gramáticas regulares o de tipo 3.

Definiremos las gramáticas con producciones lineales del modo siguiente:

* Llamaremos gramática lineal por la izquierda a toda G:= (V,Σ,S,P) gramática tal que todas las producciones de P son de uno de los dos tipos siguientes:
* A->, donde A ∈ V ya ∈ Σ ∪{λ}.
* A→ a B, donde A, B ∈ V ya ∈ Σ ∪ {λ}.
* Llamaremos gramática lineal por la derecha a toda gramática G:= (V,Σ,S,P) tal que todas las producciones de P son de uno de los dos tipos siguientes:
* A→ a, donde A ∈ V ya ∈ Σ ∪ {λ}.
* A→ B a, donde A, B ∈ V ya ∈ Σ ∪ {λ}.
* Llamaremos gramáticas regularesa las gramáticas lineales por la izquierda o lineales por la derecha.

La dualidad (y simetría) entre las gramáticas lineales a izquierda o lineales a derechaes obvia y nos quedaremos solamente con las gramáticas lineales a izquierda.

#### Lenguajes regulares.

Un lenguaje L ⊆ Σ ∗ se denomina un lenguaje regular si existe una gramática regular G:= (V,Σ,S,P) tal que L=L(G). Por definición una producción puede ser una transformación del tipo xAy → w,donde x, y, w ∈ (Σ ∪ V)∗, A ∈ V. A las palabras α y β se las denomina contexto de la producción (o contexto de la variable A en esa producción). Así, una producción libre de contexto es una producción en la que ninguna variable tiene contexto, esto es, de la forma A → w, con A ∈ V.

### Gramáticas libres de contexto o de tipo 2.

Una gramática libre de contexto es una gramática G:= (V,Σ,S,P) tal que todas las producciones de P son del tipo siguiente:

A → w, donde A ∈ V y w∈ (Σ∪V)∗.

Un lenguaje libre de contexto es un lenguaje generado por una gramática libre de contexto. El siguiente peldaño en la jerarquía de Chomsky son los lenguajes donde aparece el contexto. Pero con una gran restricción, las partes derechas de una producción tienen más longitud que la parte izquierda.

### Gramáticas sensibles al contexto o de tipo 1.

Llamaremos gramática sensible al contexto a toda gramática G:= (V,Σ,S,P) tal que todas las producciones de P son del tipo siguiente:

x Ay → xwy,

donde A ∈ V y x,y,w ∈(Σ∪V)∗ pero w=/=λ. Esto implica que en todas las producciones hay al menos una variable en la parte “izquierda” de la producción y la parte derecha tiene más longitud que la parte derecha. Un lenguaje sensible al contexto es un lenguaje generado por una gramática sensible al contexto.

### Gramáticas formales de tipo 0.

Llamaremos gramática formal de tipo 0 a toda gramática G:= (V,Σ,S,P) que admite todo tipo de producciones, esto es, sus producciones son de la forma

x → y,

donde x,y ∈ (Σ∪V)∗, x=/=λ.

En las gramáticas de tipo 0 (las más generales) admitimos que haya producciones sin ninguna variable en el lado izquierdo de la producción.

# Conclusiones.

## Víctor Hugo Vázquez Gómez.

Un lenguaje formal es un lenguaje cuyos símbolos primitivos y reglas están **formalmente especificados**. Se usan como lenguajes naturales, como puede ser el español o un lenguaje artificial como los lenguajes de programación.

Su entendimiento es importante para el desarrollo de un compilador, interpretes o ensambladores.

En la realización de esta investigación nos percatamos de que hay muy poca información a la vista y que para obtener mejores resultados deberíamos cuando menos leer por completo los primeros capítulos de los libros que estamos utilizando. El tiempo ha sido un factor que ha jugado en nuestra contra puesto que no nos ha sido posible leer y **entender** todo en un fin de semana.

## Nicole Rodríguez González.

Los lenguajes formales son el conjunto de cadenas que se crean a partir de un alfabeto creado con símbolos que pueden ser finitos o infinitos.

Las cadenas están formadas por símbolos del alfabeto y se forman con una gramática especificad las cuales contienen reglas para crear estas cadenas.

# *Referencias.*

*[1]* Alfonseca Moreno, M; de la Cruz Echeandía, M; Ortega de la Puente, A; Pulido Cañabate, E. <> PEARSON EDUCACIÓN, S.A., Madrid, 2006, pp. 1-4.

*[2]* Viñuela, P; Fernandez P; Millán, D. <>. ADDISON-WESLEY. pp. 7-11.

*[3]*https://ocw.unican.es/pluginfile.php/1038/course/section/1209/material\_teorico\_del\_curso\_lenguajes\_formales.pdf