

ENFOQUE LINGÜÍSTICO
Fernando Galindo Soria
Instituto Politécnico Nacional
UPIICSA ESCOM

Julio de 1994

RESUMEN

En este trabajo se presenta el *Enfoque Lingüístico* como una nueva forma de ver la realidad, en la cual se considera que, *cualquier cosa se puede ver como una oración de algún lenguaje.*

Mediante este enfoque se pueden representar como oraciones de algún lenguaje las imágenes, las reglas de un sistema experto, las trayectorias de un planeta, el movimiento de la mano, la trayectoria que sigue una pieza de ajedrez al moverse, una huella digital, la señal de un electrocardiograma, etc., con lo que se amplía el concepto de lenguaje que normalmente se restringía a los lenguajes naturales (como el Español, Inglés, Chino o Árabe) y artificiales (como Fortran, Pascal o C), para incluir cualquier cosa.

Como primer punto se comenta que el Enfoque Lingüístico surgió en México a mediados de los 70 y que tiene como antecedentes los trabajos desarrollados por *Noam Chomsky* a mediados de los 50 para representar la estructura de los lenguajes naturales, la aplicación de estos trabajos a la construcción de compiladores, *con lo que se generalizó el concepto de lenguaje para incluir a los lenguajes artificiales de programación* y el surgimiento del *Reconocimiento Sintáctico de Formas* donde se plantea que *la forma de cualquier objeto o el patrón de comportamiento de cualquier proceso se puede ver como una oración de algún lenguaje*, con lo que *se generalizó aún más el concepto de lenguaje para incluir a cualquier patrón o forma*, de donde se llegó finalmente a la idea del *Enfoque Lingüístico*.

Como siguiente punto se *generaliza el concepto de unidad léxica para incluir a cualquier cosa que se pueda percibir o conceptualizar.*

Se presentan ejemplos de múltiples tipos de lenguajes, como: lenguajes naturales restringidos, lenguajes de trayectoria, lenguajes para representar sistemas de diagnóstico y toma de decisiones y *una ecuación lingüística con la cual se tiene la estructura de cualquier elemento de la naturaleza.*

Más adelante se ve que *la Gramática* (o conjunto de reglas que representan a un lenguaje) *es equivalente a un sistema de información.* por lo que *si se tienen un conjunto de oraciones de un lenguaje se puede encontrar su gramática y de ahí obtener el sistema que trabaja con ese lenguaje.*

Finalmente se muestra que *las oraciones tienen una estructura fractal*, ya que, según el enfoque lingüístico *cada una de las unidades léxicas que componen a la oración se pueden ver a su vez como oraciones.*

ENFOQUE LINGÜÍSTICO
Fernando Galindo Soria
Instituto Politécnico Nacional
UPIICSA ESCOM
Julio de 1994

INTRODUCCIÓN

En este trabajo se presenta el *Enfoque Lingüístico* como una nueva forma de ver la realidad, en la cual se considera que, *cualquier cosa que se pueda percibir o conceptualizar se puede representar como una oración de algún lenguaje.*

Mediante este enfoque se pueden ver como oraciones de algún lenguaje las imágenes, las reglas de un sistema experto, las trayectorias de un planeta, el movimiento de la mano, la trayectoria que sigue al moverse una pieza de ajedrez, una huella digital, la señal de un electrocardiograma, etc., y se amplía el concepto de lenguaje que normalmente se restringía a los lenguajes naturales (como el Español, Inglés, Chino o Árabe) y artificiales (como Fortran, Pascal o C), para incluir cualquier cosa.

1. ANTECEDENTES DEL ENFOQUE LINGÜÍSTICO

El Enfoque Lingüístico surgió como resultado de múltiples trabajos realizados desde mediados de los 70 principalmente en El Colegio de México, la Escuela Superior de Física y Matemáticas (ESFM), el Centro Nacional de Cálculo (CENAC) y la UPIICSA del IPN.

Empero sus primeros antecedentes se encuentran en los trabajos desarrollados por *Noam Chomsky* para representar la estructura de los lenguajes naturales y plasmados en su libro *Syntactic Structures* publicado en 1957, en la aplicación de estos trabajos a la construcción de compiladores y en el surgimiento del Reconocimiento Sintáctico de Formas.

A principios de los 60 los trabajos de Chomsky se aplicaron para describir la gramática del Lenguaje Algol propiciando que ya para finales de esa década empezaran a surgir libros donde se mostraba como construir un compilador a partir de la gramática de un lenguaje dado, *con lo cual se generalizó el concepto de lenguaje para incluir a los lenguajes artificiales de programación.*

En paralelo con lo anterior y desde mediados de los 60 se empezó a aplicar la Lingüística Matemática al Reconocimiento de Patrones (o sea al estudio de la forma o patrón de comportamiento asociado con algún objeto o proceso) siendo la base de los métodos de *Reconocimiento Sintáctico de Formas* en los cuales se plantea que *la forma de cualquier objeto o el patrón de comportamiento de cualquier proceso se puede ver como una oración de algún lenguaje*, por ejemplo se ve a una imagen o proceso (voz, movimiento de los planetas, jugada de ajedrez, etc.) como una oración del lenguaje de imágenes o de procesos, *con lo que se generalizó aún más el concepto de lenguaje para incluir a cualquier patrón o forma.*

A finales de los 60 *Rafael C. Gonzalez* y *Michael C. Thomason* escribieron el libro *Syntactic Pattern Recognition* donde ya se presenta en una forma elaborada el concepto de reconocimiento sintáctico de patrones y *aparece claramente la idea de representar patrones como oraciones de algún lenguaje*, con lo que se extiende el concepto de lenguaje para incluir figuras geométricas, dibujos e imágenes en general, además de la forma o patrón de cualquier objeto o proceso, por lo que *podemos considerar que este libro es el precursor del enfoque lingüístico*.

Durante todos los 70 se continuaron realizaron aplicaciones de las herramientas lingüísticas a la construcción de compiladores y al reconocimiento de formas y se extendió su uso a otras áreas de la Informática como la complejidad de algoritmos y la construcción de sistemas operativos, manejadores de bases de datos y editores.

2. RESULTADOS GENERALES

En particular desde mediados de los 70 en el IPN se ha desarrollado el Enfoque Lingüístico y en la actualidad ya se cuenta con una gran cantidad de resultados y aplicaciones, por lo que en los siguientes puntos se presentaran algunos de éstos.

2.1. CONCEPTOS GENERALES

En general se define *un lenguaje como un conjunto*. Mas específicamente cualquier mecanismo que tenga un conjunto de elementos o palabras (*Alfabeto*) sobre los que se pueda aplicar un conjunto de reglas para relacionarlos (*Sintaxis*) y asociarle un significado (*Semántica*) se puede decir que cuenta con un lenguaje.

En particular una oración está formada por palabras (o unidades léxicas) relacionadas entre sí (o sintaxis) y con un cierto significado (o semántica).

2.2. PERCEPCIÓN Y CONCEPTUALIZACIÓN DE UNIDADES LÉXICAS

Generalmente solo se aceptan como unidades léxicas a los componentes de las oraciones de los lenguajes naturales, y desde el surgimiento de los lenguajes de programación se ha ampliado el término para abarcarlos también a ellos, sin embargo aún en la actualidad no se incluyen dentro de este concepto por ejemplo a las partes de un gato o a las ramas de un árbol o a las señales que emite una neurona, y el estudio de estos elementos se incluye como problemas independientes dentro de la Semiótica. Por lo que, lo primero que se tiene que hacer es *generalizar el concepto de unidad léxica*.

En el enfoque lingüístico se *generaliza el concepto de unidad léxica a cualquier cosa que se pueda percibir o conceptualizar*. Por ejemplo *la imagen que percibimos de un gato se puede considerar como la oración que representa al gato, y las diferentes partes del gato equivalen a las palabras o unidades léxicas que componen a la oración*.

2.3. EJEMPLOS DE LENGUAJES

A continuación se mostraran algunos ejemplos de lenguajes no tradicionales con el fin de mostrar la diversidad de aplicaciones de este enfoque y su fuerza.

2.3.1. LENGUAJE NATURAL RESTRINGIDO

Como primer ejemplo se presenta el caso de los lenguajes naturales restringidos donde se incluyen: los Lenguajes Imperativos formados por oraciones en las que se plantea algún requerimiento u orden, los Lenguajes Declarativos que se centran en el planteamiento de hechos o reglas de inferencia, los Lenguajes Interrogativos mediante los que se plantean preguntas sobre el sistema y sus combinaciones.

Como posibles ejemplos de oraciones de este tipo de lenguajes se tienen:

En diagnóstico médico: la conversación del paciente con el médico.

En un sistema de nómina: la descripción de los cambios o modificaciones requeridas.

En pruebas de personalidad: la oración formada por una cadena de signos, donde cada signo representa la respuesta a una pregunta de la prueba (normalmente tienen un máximo de cinco posibles respuestas) y la cadena representa todas las respuestas dadas por una persona, de donde una oración típica se parece a una cadena binaria.

2.3.2. LENGUAJES DE TRAYECTORIA.

Estamos acostumbrados a manejar el concepto de lenguaje como sinónimo de lenguaje natural escrito o hablado y si nos fuerzan mucho podemos aceptar la idea de lenguaje de señas o movimientos.

Sin embargo, atrás de todos estos tipos de lenguajes podemos encontrar un concepto más general que los incluye y generaliza conocido como *Lenguaje de Trayectoria*. en el cual se incluyen todos los lenguajes visuales (Gráficas), de movimiento, (Jugadas en un tablero de Ajedrez, movimiento de un robot, etc.) y cualquier otro lenguaje que se pueda representar por una trayectoria (Sonido, recorrido por una ciudad, etc.).

Un Lenguaje de Trayectoria se puede conceptualizar como un conjunto de oraciones que representan la trayectoria espacial o temporal entre dos puntos para lo cual en lugar de lexemas o fonemas utiliza como unidades básicas símbolos que representan trayectorias.

Prácticamente todos los elementos de los lenguajes tradicionales (Natural escrito y hablado, señas, etc.) se pueden ver como una combinación de trayectorias ya que al final de cuentas lo que se representa mediante una letra es el resultado de la trayectoria que sigue la pluma de un punto a otro, o lo que se escucha como un fonema es la abstracción de la trayectoria que sigue una onda sinusoidal y existen múltiples fenómenos que captamos como resultado de una trayectoria en alguna dimensión entera o fractal (el movimiento de los planetas, el movimiento de un pie, el paso del tiempo, el crecimiento de una célula, etc.).

2.4. DEL LENGUAJE AL SISTEMA

El concepto de lenguaje es muy amplio y es importante saber elegir el que se adecue más a nuestro problema. Ya que se eligió un lenguaje X, el siguiente paso consiste en captar muchas oraciones (Grabando, observándolo, sintiéndolo, etc.).

Ahora bien, si se tuviera que tener la lista de todas las oraciones de un lenguaje no terminaríamos, por lo que, comúnmente *en lugar de la lista de oraciones se utiliza una Gramática o conjunto de reglas que representan al lenguaje.*

La Gramática que representa al lenguaje es equivalente a un sistema de información. por lo que *si se tienen un conjunto de oraciones de un lenguaje se puede encontrar su gramática y de ahí obtener el sistema que trabaja con ese lenguaje.*

2.5. ESTRUCTURA FRACTAL DE LAS ORACIONES

Un resultado que se encontró junto con Juan Martín González Vázquez en 1984, fue que *las oraciones tienen una estructura fractal, ya que, según el enfoque lingüístico cada una de las unidades léxicas que componen a la oración se pueden ver a su vez como oraciones.* Por ejemplo en la siguiente oración:

El perro mordió al gato

Cada una de las palabras puede ser una oración y entonces la palabra

perro

es por sí sola una oración compuesta por sus propias unidades léxicas, que dependiendo de la aplicación podrían ser por ejemplo :

- a) *pe, rro* para un analizador silábico.
- b) *p, e, r, o* para un programa de ordenamiento de palabras.

A su vez *cada letra se puede ver como una oración* de otro lenguaje, por ejemplo la letra

p

sería una oración donde las unidades léxicas podrían ser los diferentes trazos que la forman.

CONCLUSIÓN

En este documento se mostró como el estudio de los lenguajes se ha ampliado desde la idea original del tratamiento de los lenguajes naturales, su ampliación para incluir a los lenguajes artificiales de programación y su generalización para incluir a cualquier patrón o forma, de donde se llegó finalmente a la idea del *Enfoque Lingüístico que plantea que cualquier cosa se puede ver como una oración de algún lenguaje.* Además se mostraron algunas de sus aplicaciones con el fin de que se vea que es un concepto general y que se encuentra en el núcleo de las herramientas que nos permiten estudiar la realidad.

BIBLIOGRAFÍA

1. Przemyslaw *Prosinkiewicz*, Aristid *Lindenmayer* y James *Hanan*. *Developmental Models of Herbaceous Plants for Computer Imagery Purposes*. Computer Graphics, Vol 22(4), Agosto 1988.
2. Sofia *Bueno Peralta* y Antonio *Simancas López*. *Generador de Árboles Fractales*. en Memorias del III Congreso Nacional sobre Informática y Computación, Jalapa, Ver. México, Octubre 1990
3. Heinz-Otto *Peitgen* and Dietmar *Saupe* (Editores). *The Science of Fractal Images*. ED Springer-Verlag, 1988.
4. Fernando *Galindo Soria*. *Sistemas Evolutivos: Nuevo Paradigma de la Informática*. en Memorias XVII Conferencia Latinoamericana de Informática, Caracas Venezuela, julio de 1991.
5. Fernando *Galindo Soria*. *Sistemas Evolutivos*. en Boletín de Política Informática. México, Septiembre de 1986.
6. Fernando *Galindo Soria*. *Aplicaciones de la Lingüística Matemática y los Fractales a la Generación de Imágenes*. en Memorias Simposium Nacional de Computación. México, Nov de 1991.
7. Rafael C. *Gonzalez* y Michael C. *Thomason*. *Syntactic Pattern Recognition*. Ed. Addison-Wesley.
8. Emmon *Bach*. *Teoría Sintáctica*. Ed Anagrama.
9. Arto *Salomaa*. *Formal Languages*. Ed. Academic Press.
10. Herbert. A. *Simon*. *Las Ciencias de lo Artificial*. Ed. ATE
- 11 Noam *Chomsky*. *Estructuras Sintácticas*. Ed. Siglo XXI
12. John E. *Hopcroft*, Jeffrey D. *Ullman*. *Formal Languages and Their Relation to Automata*. Ed. Addison-Wesley.