



Universidad
Nacional Autónoma
de México



INSTITUTO DE
INVESTIGACIONES
EN MATEMÁTICAS
APLICADAS Y
EN SISTEMAS

Licenciatura en Ciencia de Datos

PLN

Tema 5

Actividad Sumativa 1

Integrante:

- Villalón Pineda Luis Enrique

¿Qué estudia la semántica?

La semántica es la parte de la lingüística que estudia el significado de las palabras y expresiones. Es uno de los aspectos más complejos del lenguaje, especialmente cuando trabajamos con Procesamiento del Lenguaje Natural (PLN), porque analizar significados no es tarea sencilla.

Las dos ramas principales:

1. Semántica Léxica - Se enfoca en palabras individuales

- Estudia los "sememas" o rasgos básicos que forman el significado de cada palabra
- Analiza cómo se relacionan las palabras entre sí: sinónimos, antónimos, palabras polisémicas, términos generales y específicos, etc.

2. Semántica Composicional - Se enfoca en frases y oraciones completas

- Estudia cómo se construye el significado cuando combinamos palabras
- Sigue el Principio de Composicionalidad: el significado de una frase depende de las palabras que la forman y cómo están organizadas

¿Que es el significado?

Según Saussure, el significado es el concepto o idea mental que asociamos a un sonido o palabra (lo que él llamaba "significante"). Es algo que existe en nuestra mente, no es algo físico. Básicamente cuando analizamos palabras, vemos que su significado está formado por pequeños bloques llamados sememas o rasgos semánticos. Estos son las piezas más básicas e indivisibles del significado.

Es decir , el significado es justamente lo que estudia la semántica.

¿Qué son las expresiones complejas?

Son combinaciones de palabras que estudia la semántica composicional. Básicamente, cualquier cosa que sea más que una palabra suelta:

- Oraciones (frases completas)
- Sintagmas (grupos de palabras que funcionan juntos)

Características importantes:

- Pueden ser tan cortas o largas como sea necesario
- Se construyen uniendo palabras simples
- Deben seguir las reglas gramaticales del idioma

El significado de estas expresiones no es aleatorio. Depende de dos cosas: qué palabras usas y cómo las organizas. Esto es lo que llamamos el Principio de Composicionalidad: el todo se entiende por sus partes y cómo se conectan. Un ejemplo claro sería "El gato negro"

significa algo diferente a "El negro gato" o "Gato el negro", aunque usemos las mismas palabras.

Menciona las relaciones léxicas y da un ejemplo de cada una.

Estas relaciones son fundamentales para bases de datos léxicas como WordNet y para extraer información de textos:

1. Sinonimia - Palabras intercambiables

- Significan prácticamente lo mismo
- Ejemplo: computadora / ordenador

2. Antonimia - Palabras opuestas

- Sus significados se contradicen
- Ejemplo: bueno / malo

3. Polisemia - Una palabra, varios significados relacionados

- Misma palabra, múltiples sentidos con el mismo origen
- Ejemplo: "cometa" (el astro o el juguete)

4. Homonimia - Palabras que suenan/escriben igual pero no tienen relación

- Parecen idénticas pero sus significados no están conectados
- Ejemplo: "cobre" (el metal) vs "cobrar" (del verbo cobrar)

5. Hiperonimia/Hiponimia - Relación general-específico

- Hiperónimo: término general que abarca otros
- Hipónimos: términos específicos dentro de esa categoría
- Ejemplo: "colores" (general) → rojo, verde, naranja (específicos)

6. Meronimia - Relación parte-todo

- El merónimo es una parte del holónimo (el todo)
- Ejemplo: "camisa" (el todo) → manga, cuello, botones (las partes)

Estas relaciones son clave cuando queremos que las computadoras entiendan y procesen el lenguaje.

Elabora el campo semántico de las flores anotando sus semas compartidos.

El campo semántico de las flores

Palabras del grupo:

- Rosa
- Tulipán
- Clavel
- Margarita
- Lirio

¿Qué tienen en común?

Todas estas palabras comparten estos rasgos de significado:

- [+ Vegetal] → son seres vivos del reino vegetal
- [+ Parte de una planta] → no son la planta completa, sino una parte específica
- [+ Reproductor] → tienen función reproductiva
- [+ Ornamental] → se usan para decorar

Aunque cada flor es diferente, todas comparten estos elementos básicos de significado. Eso es lo que las agrupa en el mismo "campo semántico".

Explica el principio de recursividad.

La recursividad es usar las mismas reglas una y otra vez para crear estructuras cada vez más complejas. Es como si tuvieramos con piezas limitadas puedes construir infinitas cosas.

Ejemplo:

Puedes decir: "El gato" Luego: "El gato negro" Seguir: "El gato negro que vive en mi casa"
Continuar: "El gato negro que vive en mi casa que está en la esquina" ...y así al infinito.

El resultado:

Con herramientas finitas (palabras + reglas) podemos crear y entender mensajes infinitos. Por eso el lenguaje funciona como un algoritmo: reglas simples que se repiten generan complejidad ilimitada.

Explica como funciona WordNet.

1. Como una red semántica No es un diccionario lineal, sino una red donde las palabras se conectan entre sí según su significado.

2. Divide las palabras en 5 grupos:

- Nombres (sustantivos)
- Verbos
- Adjetivos
- Adverbios
- Elementos funcionales

3. El corazón: los "synsets" Un synset es un conjunto de sinónimos que representan el mismo concepto. Es la unidad básica que permite identificar y organizar cada palabra dentro de la red.

4. Las conexiones que usa: WordNet vincula palabras mediante estas relaciones:

- Sinonimia (palabras equivalentes)
- Antonimia (palabras opuestas)
- Hiperonimia (términos generales)
- Hiponimia (términos específicos)
- Meronimia (relaciones parte-todo)

¿Qué es la similitud semántica?

Es una forma de medir qué tan conectadas o relacionadas están dos palabras en cuanto a su significado.

¿Cómo se calcula?

Con distintos métodos

- Longitud de camino (qué tan "lejos" están dos palabras en la red)
- Algoritmo de Resnik
- Algoritmo de Dekang Lin

La similitud semántica es la base de cómo funcionan herramientas como WordNet. Les permite a las computadoras determinar, por ejemplo, que "auto" y "coche" están muy relacionados, mientras que "auto" y "manzana" no lo están.

Es darle números a algo subjetivo: ¿qué tan parecidas son dos palabras en significado?

Reflexiona y escribe cuál es la diferencia entre ambigüedad sintáctica y ambigüedad semántica.

Ambigüedad Semántica vs Ambigüedad Sintáctica:

Ambas crean confusión en el lenguaje, pero por razones diferentes:

Ambigüedad Semántica - El problema está en las palabras

¿Por qué pasa?

- Una palabra tiene significado vago o informal
- Polisemia: una misma palabra tiene varios significados

Ejemplo: "Banco" puede significar:

- Institución financiera
- Asiento para sentarse
- Conjunto de peces

Solución: La Desambiguación del Sentido de las Palabras (WSD) ayuda a las computadoras a elegir el significado correcto según el contexto.

Ambigüedad Sintáctica - El problema está en cómo armamos la frase

¿Por qué pasa?

- La estructura u organización de la oración permite varias interpretaciones
- Las palabras individuales pueden estar claras, pero su relación no

Ejemplo: "Juan le compró un coche a Pedro"

Puede significar:

1. Pedro le vendió el coche a Juan
2. Juan compró el coche para regalárselo a Pedro

Mismas palabras, pero la estructura deja dudas sobre quién hace qué.

Para resumir:

- Semántica: "No sé qué significa esta palabra"
- Sintáctica: "No sé cómo se relacionan estas palabras entre sí"

Escribe tres oraciones de tu autoría en las que se pueda presentar ambigüedad semántica y explícalas.

1. Oración: "La banda tocó un excelente concierto antes de ir a prisión."

¿Qué pasó aquí? La palabra "banda" puede significar:

- Grupo musical → Tocarón bien y luego... ¿fueron arrestados?
- Grupo criminal → Ya eran delincuentes que tocaron música antes de ser encarcelados

Una sola palabra cambia completamente la historia. Aquí es donde entra la desambiguación para que las computadoras elijan el significado correcto según el contexto.

2. Oración: "El guardia persiguió al ladrón con la linterna."

¿Quién tiene la linterna?

- El guardia usó su linterna para perseguir al ladrón

- El ladrón tenía la linterna (y por eso lo persiguieron)

Las palabras están claras, pero su organización deja dudas sobre quién hace qué con qué.

3. Oración: "Necesitamos asegurar la red antes de que se desconecte."

¿De qué red hablamos?

- Una red informática (internet, WiFi) → seguridad digital
- Una red física (de pesca, de voleibol) → amarrarla o fijarla físicamente

Misma palabra, mundos completamente diferentes. Solo el contexto nos dice si necesitamos un experto en ciberseguridad o una cuerda más fuerte.