

UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS

Inteligencia Artificial

Análisis y Generación del Lenguaje Natural

- Hugo David Calderón
- Willy Ugarte Rojas
- Jorge Valverde Rebaza

Niveles de análisis

- **Análisis morfológico**
- **Análisis sintáctico.**
- **Análisis semántico.**
- **Análisis pragmático.**

Análisis morfológico

La morfología estudia la estructura de las palabras y su relación con las categorías del lenguaje.

El objetivo del análisis morfológico automático es llevar a cabo una clasificación morfológica de las palabras.

Ejemplo 1: El análisis de la palabra *gatos* resulta en
gato+Noun+Masc

Nos indica que se trata de un sustantivo plural con género masculine

Ejemplo 2:

| Lexema | Gramema |
|--------|---------|
| cant | o |
| | es |
| | a |
| | emos |
| | en |

Análisis morfológico

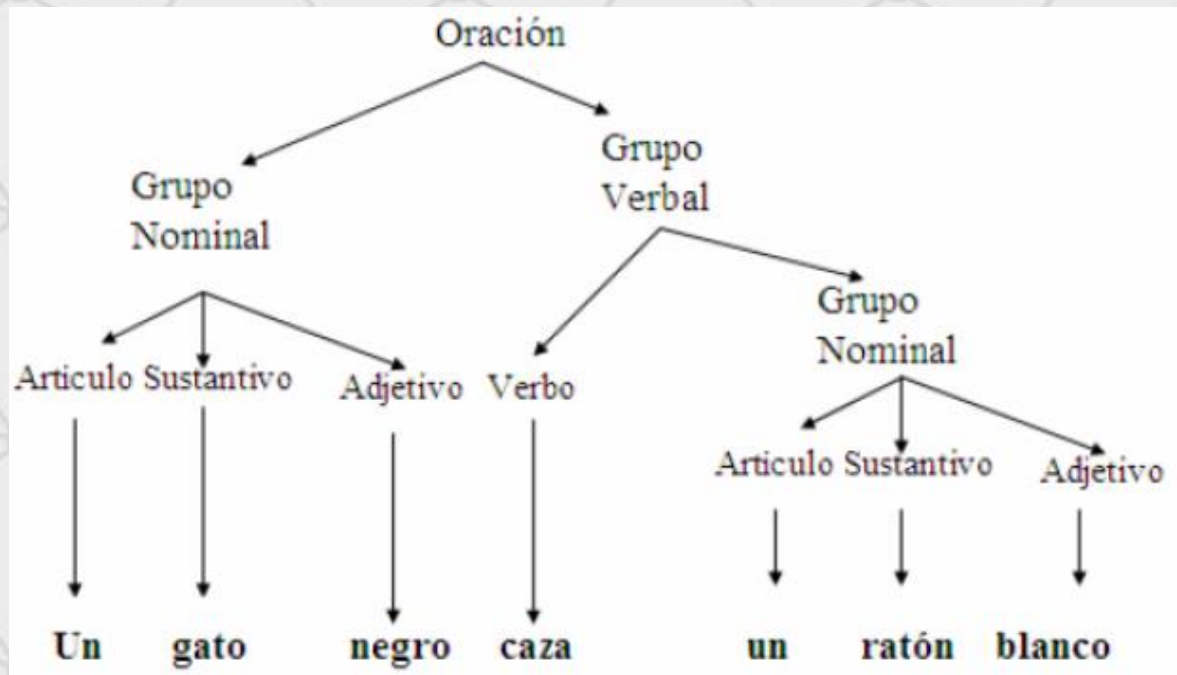
En los analizadores morfológicos se usan:

- Diccionarios de morfemas: de raíces (lexemas), de sufijos, de prefijos, de infijos
- Morfotáctica: reglas de combinación de morfemas

| | |
|---------------|---|
| un | Artículo, singular, masculino |
| gato | Sustantivo, común, masculino, singular |
| negro | Adjetivo, singular, masculino |
| caza | Verbo <i>cazar</i> . Principal, indicativo, presente, tercera persona, singular |
| un | Artículo, singular, masculino |
| ratón | Sustantivo, común, masculino, singular |
| blanco | Adjetivo, singular, masculino |

Análisis Sintáctico

Se estudia cómo combinar las palabras para formar oraciones. Su función es asignar etiquetas a cada uno de los componentes que aparecen en la oración a analizar, de manera que se sepa cómo se combinan las palabras y forman estructuras gramaticales correctas.



Análisis semántico

Se estudia el significado de las oraciones, representándolo de manera formal.

Existen varias formas de representación formal semántica de las oraciones, tales como las siguientes:

- Lógica de primer orden
- Redes semánticas
- Grafos conceptuales

Análisis pragmático

- Se estudia cómo el contexto afecta a la interpretación de las oraciones.
- Se interesa por el modo en que el contexto influye en la interpretación del significado

Técnicas de análisis del lenguaje

- **Técnicas lingüísticas formales:** Se basan en el desarrollo de reglas estructurales que se aplican en las fases de análisis del lenguaje.
- **Técnicas probabilísticas:** Se basan en el estudio en base a un conjunto de textos de referencia (corpus) de características de tipo probabilístico asociadas a las distintas fases de análisis del lenguaje

Modelos para el procesamiento del lenguaje natural

- Lógicos (gramáticas) basado en reglas
- Probabilísticos (basados en corpus)

Gramáticas Independientes de contexto

Una gramática formal es objeto o modelo matemático que permite especificar un lenguaje o lengua, es decir, es el conjunto de reglas capaces de generar todas las posibilidades combinatorias de ese lenguaje, ya sea éste un lenguaje formal o un lenguaje natural.

Definición

Una gramática es un objeto $G = (\Sigma_T, \Sigma_N, S, P)$, donde

Σ_T es un alfabeto cuyos símbolos llamamos **símbolos terminales** de la gramática.

Σ_N es un alfabeto tal que $\Sigma_T \cap \Sigma_N = \emptyset$, cuyos símbolos llamamos **símbolos no terminales** de G .

S es un símbolo no terminal señalado al que llamamos **axioma** de la gramática.

P es un conjunto de **reglas** definidas sobre el alfabeto $\Sigma = \Sigma_T \cup \Sigma_N$, tales que en la parte izquierda de todas las reglas aparece al menos un símbolo no terminal.

Ejemplo

Una gramática es un objeto $G = (\Sigma_T, \Sigma_N, S, P)$, donde

Σ_T { juan, pedro, salgado, antonio, maría, pepa, ama, lava, peina, adora, paula, sultán, corre, salta, camina }

Σ_N {=<frase>, <sujeto>, <predicado>, <verbo transitivo>, <verbo intransitivo>, <objeto directo>}

S <frase> es el axioma

P es el conjunto formado por las diecisiete reglas:

1. <frase> ::= <sujeto><predicado>
2. <sujeto> ::= juan | pedro | maría | salgado
3. <predicado> ::= <verbo transitivo><objeto directo>
4. <predicado> ::= <verbo intransitivo>
5. <verbo transitivo> ::= ama | lava | peina | adora
6. <objeto directo> ::= paula | antonio | sultán
7. <verbo intransitivo> ::= corre | salta | camina

Ejemplo

Una gramática es un objeto $G = (\Sigma_T, \Sigma_N, S, P)$, donde

Si partimos de <frase> y seguimos las reglas se puede obtener:

1. <frase> \rightarrow <sujeto><predicado>
2. \rightarrow maría<predicado>
3. \rightarrow maría <verbo transitivo><objeto directo>
5. \rightarrow maría ama <objeto directo>
6. \rightarrow *maría ama antonio*

Ejercicio: Usando la gramática del ejemplo anterior genere toda las frases permitidas.

Ejercicio

Derivar la frase “el hombre toma café”

- La regla se aplica sustituyendo el símbolo que aparece en la parte izquierda por los que aparecen en su parte derecha.
- Una derivación es la aplicación sucesiva de reglas hasta obtener una expresión que sólo contiene símbolos terminales.

```
S  => NP VP
NP => DT N
   |N
   |NP PP
VP => V NP
   |V
   |VP PP
PP => P NP
   |P
DT => el | los
N  => hombre | amigos | café | leche
V  => toma | toman
P  => con | solo
```

Qué otras frases se pueden derivar del lenguaje definido?



Fin