

ENTREGABLE 1: Gramática Formal del Mini-Lenguaje

Proyecto: Analizador de Lenguaje Natural Simple

Autores: Ricardo Méndez, Emiliano Ledesma, Diego Jiménez, Abraham Velázquez

Fecha: Noviembre 2024

Definición Formal de la Gramática

La gramática del mini-lenguaje está definida como una Gramática Independiente del Contexto (GIC), expresada formalmente como una 4-tupla:

$$G = (V, T, P, S)$$

Donde:

- V:** Conjunto de símbolos no terminales
- T:** Conjunto de símbolos terminales (alfabeto)
- P:** Conjunto de producciones (reglas gramaticales)
- S:** Símbolo inicial

Componentes de la Gramática

1. Símbolos No Terminales (V)

$$V = \{S, SN, SV, SP\}$$

Descripción de cada símbolo:

Símbolo	Nombre	Descripción
S	Oración	Símbolo inicial, representa una oración completa
SN	Sintagma Nominal	Representa el sujeto de la oración
SV	Sintagma Verbal	Representa el predicado (verbo + complementos)
SP	Sintagma Preposicional	Representa complementos circunstanciales

2. Símbolos Terminales (T)

El alfabeto terminal está compuesto por categorías gramaticales y sus instancias específicas:

$T = \{DET, N, V, PRON, PREP, ADV\} \cup \text{Vocabulario}$

2.1 Categorías Gramaticales

Categoría	Código	Descripción
Determinante	DET	Artículos y demostrativos
Nombre/Sustantivo	N	Sustantivos comunes y propios
Verbo	V	Verbos de acción o estado
Pronombre	PRON	Pronombres personales
Preposición	PREP	Preposiciones
Adverbio	ADV	Modificadores adverbiales

2.2 Vocabulario del Mini-Lenguaje

Determinantes (8 palabras)

$DET = \{\text{el, la, un, una, los, las, mi, tu}\}$

Sustantivos (14 palabras)

$N = \{\text{gato, perro, niño, niña, libro, parque, pescado, jardín, casa, María, Juan, hermano, matemáticas, niños}\}$

Verbos (9 palabras)

$V = \{\text{come, corre, estudia, lee, camino, juega, juegan, escribe, canta}\}$

Pronombres (5 palabras)

$PRON = \{\text{yo, tú, él, ella, nosotros}\}$

Preposiciones (5 palabras)

PREP = {por, en, de, con, a}

Adverbios (4 palabras)

ADV = {rápidamente, bien, mal, rápido}

Total de palabras en el alfabeto: 45 palabras

3. Producciones (P)

El conjunto de reglas de producción define cómo se pueden generar oraciones válidas:

```
P = {  
  R1: S → SN SV  
  R2: SN → DET N  
  R3: SN → PRON  
  R4: SN → N  
  R5: SV → V SN  
  R6: SV → V SP  
  R7: SV → V ADV  
  R8: SV → V  
  R9: SP → PREP SN  
}
```

3.1 Explicación Detallada de Cada Regla

Regla 1: S → SN SV

- **Descripción:** Una oración se compone de un Sintagma Nominal (sujeto) seguido de un Sintagma Verbal (predicado)
- **Ejemplo:** "el gato" (SN) + "come pescado" (SV)
- **Justificación:** Estructura básica del español: Sujeto + Predicado

Regla 2: SN → DET N

- **Descripción:** Un Sintagma Nominal puede formarse con un determinante y un sustantivo
- **Ejemplo:** "el gato", "la casa", "un libro"
- **Uso:** Sujetos con artículo determinado o indeterminado

Regla 3: SN → PRON

- **Descripción:** Un Sintagma Nominal puede ser simplemente un pronombre
- **Ejemplo:** "yo", "tú", "ella"
- **Uso:** Sujetos pronominales

Regla 4: SN → N

- **Descripción:** Un Sintagma Nominal puede ser solo un sustantivo (nombres propios)
- **Ejemplo:** "María", "Juan"
- **Uso:** Nombres propios sin determinante

Regla 5: SV → V SN

- **Descripción:** Un Sintagma Verbal con verbo transitivo y objeto directo
- **Ejemplo:** "come" + "pescado" → "come pescado"
- **Uso:** Verbos que requieren objeto directo

Regla 6: SV → V SP

- **Descripción:** Un Sintagma Verbal con verbo y complemento circunstancial
- **Ejemplo:** "camino" + "por el parque" → "camino por el parque"
- **Uso:** Verbos con complementos de lugar, tiempo, modo

Regla 7: SV → V ADV

- **Descripción:** Un Sintagma Verbal con verbo y modificador adverbial
- **Ejemplo:** "corre" + "rápidamente" → "corre rápidamente"
- **Uso:** Verbos con adverbios de modo

Regla 8: SV → V

- **Descripción:** Un Sintagma Verbal puede ser solo un verbo intransitivo
- **Ejemplo:** "corre", "juega"
- **Uso:** Verbos que no requieren complemento

Regla 9: SP → PREP SN

- **Descripción:** Un Sintagma Preposicional se forma con preposición y Sintagma Nominal
- **Ejemplo:** "por" + "el parque" → "por el parque"
- **Uso:** Complementos circunstanciales

4. Símbolo Inicial (S)

S = S

El símbolo S (Oración) es el punto de partida para generar cualquier cadena válida del lenguaje.

Propiedades de la Gramática

4.1 Tipo de Gramática

Clasificación según Jerarquía de Chomsky: Tipo 2 (Gramática Independiente del Contexto)

Justificación:

- Todas las producciones tienen la forma $A \rightarrow \alpha$, donde:
 - $A \in V$ (símbolo no terminal único en el lado izquierdo)
 - $\alpha \in (V \cup T)^*$ (cadena de terminales y/o no terminales)
- No hay producciones sensibles al contexto (no hay $\alpha A \beta \rightarrow \alpha \gamma \beta$)

4.2 Ambigüedad

Estado: Gramática NO ambigua

Explicación: Para cualquier oración válida del lenguaje, existe un único árbol de derivación. No hay múltiples formas de aplicar las producciones para generar la misma cadena.

Ejemplo: "el gato come pescado" tiene una única derivación:

$S \rightarrow SN \quad SV \rightarrow DET \ N \quad SV \rightarrow DET \ N \ V \quad SN \rightarrow DET \ N \ V \ N$

4.3 Recursividad

Estado: Gramática NO recursiva

Explicación: No hay producciones de la forma $A \rightarrow \alpha A \beta$ donde A aparece en su propia expansión. Esto limita la gramática a oraciones simples pero simplifica el análisis.

4.4 Poder Expresivo

Oraciones que puede generar:

- Oraciones simples con estructura SVO (Sujeto-Verbo-Objeto)
- Oraciones con sujeto pronominal
- Oraciones con complementos circunstanciales
- Oraciones con verbos intransitivos

Oraciones que NO puede generar:

- Oraciones compuestas (coordinadas o subordinadas)
- Oraciones con adjetivos
- Oraciones con múltiples complementos
- Oraciones pasivas
- Oraciones interrogativas o imperativas

Ejemplos de Derivación

Ejemplo 1: "el gato come pescado"

Derivación paso a paso (leftmost):

- | | |
|---------------------------|--------------------|
| 1. S | |
| 2. S ⇒ SN SV | [Aplicar R1] |
| 3. ⇒ DET N SV | [Aplicar R2 en SN] |
| 4. ⇒ el N SV | [Sustituir DET] |
| 5. ⇒ el gato SV | [Sustituir N] |
| 6. ⇒ el gato V SN | [Aplicar R5 en SV] |
| 7. ⇒ el gato come SN | [Sustituir V] |
| 8. ⇒ el gato come N | [Aplicar R4 en SN] |
| 9. ⇒ el gato come pescado | [Sustituir N] |

Reglas aplicadas: R1, R2, R5, R4

Ejemplo 2: "yo camino por el parque"

Derivación paso a paso:

- | | |
|-------------------------------|--------------------|
| 1. S | |
| 2. S ⇒ SN SV | [Aplicar R1] |
| 3. ⇒ PRON SV | [Aplicar R3 en SN] |
| 4. ⇒ yo SV | [Sustituir PRON] |
| 5. ⇒ yo V SP | [Aplicar R6 en SV] |
| 6. ⇒ yo camino SP | [Sustituir V] |
| 7. ⇒ yo camino PREP SN | [Aplicar R9 en SP] |
| 8. ⇒ yo camino por SN | [Sustituir PREP] |
| 9. ⇒ yo camino por DET N | [Aplicar R2 en SN] |
| 10. ⇒ yo camino por el N | [Sustituir DET] |
| 11. ⇒ yo camino por el parque | [Sustituir N] |

Reglas aplicadas: R1, R3, R6, R9, R2

Ejemplo 3: "María estudia matemáticas"

Derivación paso a paso:

- | | |
|--------------------------------|--------------------|
| 1. S | |
| 2. S ⇒ SN SV | [Aplicar R1] |
| 3. ⇒ N SV | [Aplicar R4 en SN] |
| 4. ⇒ María SV | [Sustituir N] |
| 5. ⇒ María V SN | [Aplicar R5 en SV] |
| 6. ⇒ María estudia SN | [Sustituir V] |
| 7. ⇒ María estudia N | [Aplicar R4 en SN] |
| 8. ⇒ María estudia matemáticas | [Sustituir N] |

Reglas aplicadas: R1, R4, R5, R4

Lenguaje Generado

El lenguaje $L(G)$ generado por esta gramática es:

$$L(G) = \{w \in T^* \mid S \Rightarrow^* w\}$$

Es decir, todas las cadenas de símbolos terminales que pueden ser derivadas desde el símbolo inicial S .

Cardinalidad del Lenguaje

El lenguaje es **finito** debido a:

1. Vocabulario limitado (45 palabras)
2. Ausencia de recursividad
3. Producciones no recursivas

Estimación: Aproximadamente 10,000 - 50,000 oraciones válidas distintas (considerando todas las combinaciones posibles de las reglas con el vocabulario disponible).

Extensiones Futuras Posibles

Para ampliar el poder expresivo de la gramática, se podrían añadir:

1. Adjetivos:

```
SN → DET ADJ N
ADJ → {grande, pequeño, bonito, ...}
```

2. Conjunciones coordinantes:

```
S → S CONJ S
CONJ → {y, o, pero}
```

3. Subordinadas:

```
S → S CONJ_SUB S
CONJ_SUB → {que, porque, cuando}
```

4. Adverbios de tiempo/lugar:

```
SV → ADV_T V
ADV_T → {hoy, ayer, mañana}
```

Validación de la Gramática

Verificación de Propiedades

Propiedad 1: Todas las producciones son independientes del contexto

- Estado: CUMPLE
- Todas las reglas tienen forma $A \rightarrow \alpha$

Propiedad 2: El símbolo inicial S no aparece en el lado derecho

- Estado: CUMPLE
- S solo aparece en lado izquierdo de R1

Propiedad 3: No hay símbolos inútiles

- Estado: CUMPLE
- Todos los no terminales {S, SN, SV, SP} son alcanzables y generan terminales

Propiedad 4: No hay producciones vacías (ϵ -producciones)

- Estado: CUMPLE
- Ninguna producción genera la cadena vacía

Conclusiones

La gramática definida cumple con los requisitos del proyecto:

1. Es una Gramática Independiente del Contexto válida
2. Define un alfabeto limitado de 45 palabras
3. Genera oraciones simples en español con estructura SVO
4. Es no ambigua y no recursiva
5. Permite análisis sintáctico determinista
6. Es adecuada para implementación con Autómatas Finitos (en casos simples)

La gramática sirve como base formal para el analizador de oraciones implementado en el sistema.

Fin del Documento: Gramática Formal del Mini-Lenguaje