

Teoría de la Computación 2025

Proyecto 2

10.septiembre.2025

Este proyecto consiste en implementar un algoritmo de *parsing* para una gramática CFG.

Gramáticas

Investigar la implementación del algoritmo para convertir una gramática CFG a su forma normal de Chomsky (CNF).

Investigar sobre el algoritmo CYK (Cocke-Younger-Kasami) para realizar el *parsing* de una gramática CFG.

Esta parte del proyecto consiste en implementar el algoritmo CKY para determinar si una frase simple en el idioma inglés es parte de un lenguaje generado por una gramática.

La gramática que usaremos como ejemplo es la siguiente:

```
S  →  NP VP
VP →  VP PP
VP →  V NP
VP →  cooks | drinks | eats | cuts
PP →  P NP
NP →  Det N
NP →  he | she
V  →  cooks | drinks | eats | cuts
P  →  in | with
N  →  cat | dog
N  →  beer | cake | juice | meat | soup
N  →  fork | knife | oven | spoon
Det →  a | the
```

Objetivos

- Convertir una gramática CFG a su CNF.
- Implementar del algoritmo CYK.

Especificaciones

a) Entrada:

- La gramática CFG: lista de variables, lista de símbolos terminales, lista de reglas.
- Una expresión sentencia (frase) en el idioma inglés w . Por ejemplo, se ingresa la expresión $w = \text{She eats a cake with a fork}$; $w = \text{The cat drinks the beer}$.
- Indique en su reporte ejemplos de frases en el lenguaje, y ejemplos que no estén en el lenguaje.

b) Salida:

- un SÍ, si la expresión w pertenece al lenguaje descrito por la gramática, o un NO en caso contrario.

- También se debe indicar el tiempo que se tarda el algoritmo en realizar dicha validación.
- Construir el *parse tree* de la expresión w . (Para ello, basta extender el algoritmo CYK, e ir guardando los nodos de la tabla dinámica, en lugar sólo del valor booleano).

Nota: El algoritmo CYK requiere que la gramática a usar esté en la Forma Normal de Chomsky.

El algoritmo requiere el uso de *programación dinámica*. Deberán investigar en qué consiste y cómo se implementa.

Las frases aceptadas son sintácticamente correctas (nadie garantiza que sean semánticamente correctas).

Ponderación

El proyecto en total tiene un valor de 20 puntos, distribuidos de la siguiente forma:

Característica	Ponderación
Implementación de Conversión a la forma normal de Chomsky	6
Implementación de programación dinámica	2
Implementación del algoritmo CYK (algoritmo)	5
Construcción del Parse Tree (algoritmo)	4
Informe técnico	3
Total	20 puntos

Debe entregarse documentación para tener derecho a nota. La documentación debe de abarcar:

- Diseño de la aplicación (el método de modelado a discreción del desarrollador o desarrolladora).
- Discusión (obstáculos encontrados, recomendaciones, etc.).
- Ejemplos y pruebas realizadas.
 - Dar 2 ejemplos de cadenas aceptadas que sean semánticamente correctas.
 - Dar 2 ejemplos de cadenas aceptadas que no sean semánticamente correctas.
 - Dar 2 ejemplos de cadenas no aceptadas por la gramática.

Fecha Importantes

Entrega y revisión del proyecto: **lunes 13 de octubre.**

Entrega del informe técnico y código: **viernes 17 de octubre.**

Referencias

- <https://devimam.github.io/cfgtocnf/>
- https://en.wikipedia.org/wiki/CYK_algorithm
- https://en.wikipedia.org/wiki/Phrase_structure_rules
- <https://www.geeksforgeeks.org/cyk-algorithm-for-context-free-grammar/>
- <https://www.geeksforgeeks.org/cocke-younger-kasami-cyk-algorithm/>
- <https://www.xarg.org/tools/cyk-algorithm/>
- <https://web.cs.ucdavis.edu/~rogaway/classes/120/winter12/CYK.pdf>