

Compiladores

Martínez Coronel Brayan Yosafat

3CM7

Jerarquía
de Chomsky



Jerarquía de Chomsky

Un lenguaje formal en el sentido de la teoría de Chomsky es un conjunto de secuencias o cadenas sobre algún vocabulario *sigma* (Σ). Estos lenguajes pueden ser definidos por significados finitos mediante una gramática: se compone de un conjunto finito de símbolos del vocabulario (llamados terminales), un conjunto finito de símbolos no terminales, un símbolo inicial (que pertenece a los no terminales) y un conjunto finito de reglas. Ahora, en la teoría formal de lenguajes, son clasificados en 4 categorías:

Lenguajes computacionalmente numerables (o recursivamente numerables)

La clase de todos los lenguajes que pueden ser definidos por alguna gramática formal se llama computacionalmente numerable o recursivamente numerable, en otras palabras, cualquier lenguaje definible por una máquina de Turing (o equivalente) entra en esta categoría. *Eso quiere decir que, cualquier programa en las computadoras, con un número finito de pasos, es un ejemplo de esta clase.*

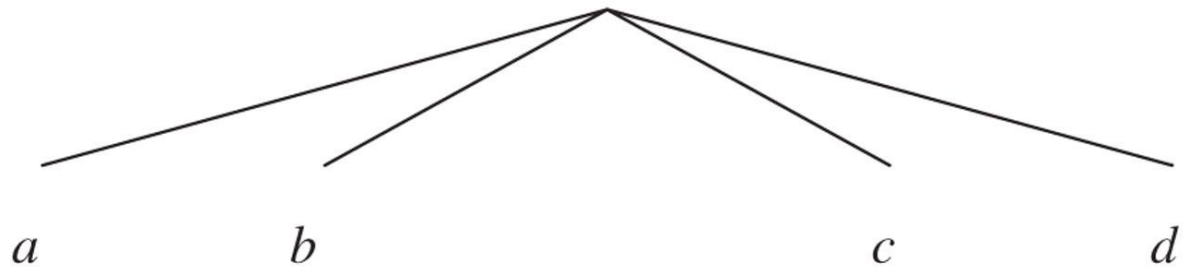
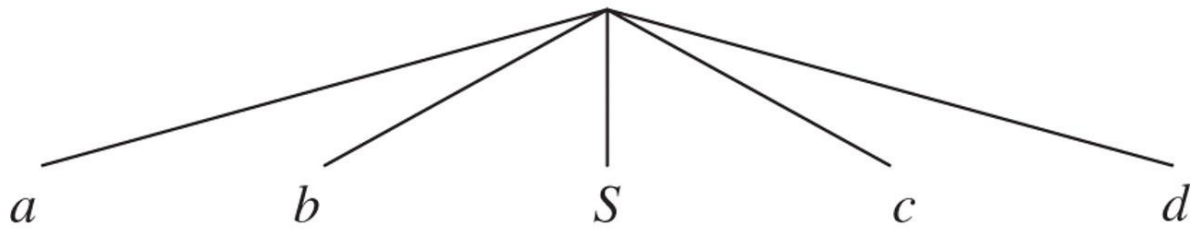
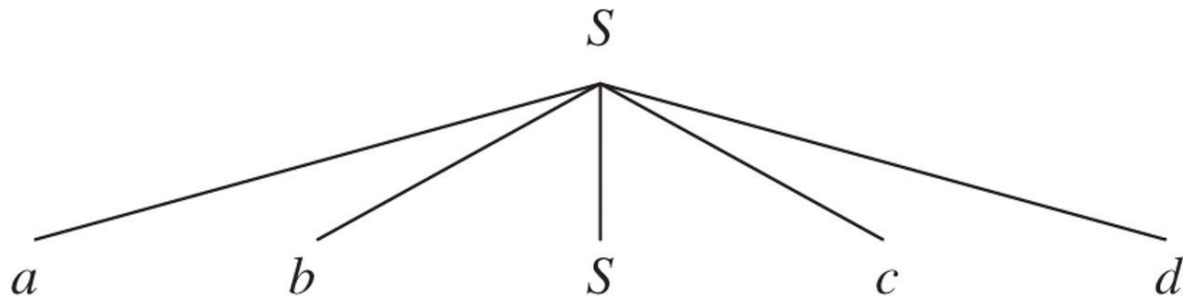
Lenguajes sensibles al contexto

Las gramáticas sensibles al contexto son aquellas donde la derecha de cada regla (llamada producción) no es más larga que el lado izquierdo. Esta definición hace que un procedimiento de decisión (si dada una cadena c determinar si pertenece al lenguaje producido por la gramática) sea posible, aunque, para fines prácticos como puede llegar a ser muy complejo, no se cuenta con ello. *Ejemplos de este tipo son los siguientes, donde x^n denota una cadena de símbolos que contiene n repeticiones de x :*

- *El conjunto de todos los números al cuadrado*
- $a^n b^n c^n$

Lenguajes libres de contexto

En una gramática libre de contexto, todas las reglas o producciones son de la forma $A \rightarrow B$, de forma que A es un no terminal y B es una cadena de símbolos. Por ejemplo, el lenguaje $(ab)^n(cd)^n$, que tiene las reglas: $S \rightarrow abScd$; $S \rightarrow abcd$. Dada la cadena $ababababcdcdcd$, podemos representarlo con un árbol de estructura de frase:



Lenguajes Regulares

En estas gramáticas, las reglas pueden tomar dos formas: $A \rightarrow a$; $A \rightarrow aB$; donde las mayúsculas denotan símbolos no terminales y, las minúsculas, símbolos terminales. De acuerdo con una interpretación natural, los no terminales significan nombres de estados de un autómata. De hecho, por eso, también se les puede llamar como lenguajes de estado finito. Por ejemplo: $a^n b^m$.

La investigación fue realizada con el Journal del Royal Society Publishing:

<https://royalsocietypublishing.org/doi/full/10.1098/rstb.2012.0077#RSTB20120077F3>