

La Jerarquía de Chomsky

Apuntes sobre la Complejidad

Universidad de Cantabria

Esquema

1 Motivación

2 Jerarquía de Chomsky

- Gramáticas Regulares
- Gramáticas Libres de Contexto
- Gramáticas Sensibles al Contexto

Ideas y Nociones

Como se ha mencionado anteriormente, los lenguajes son conjuntos de palabras definidos por un alfabeto. No estamos interesados en cualquier conjunto, los lenguajes que nos interesan tienen una estructura inherente.

Ideas y Nociones

Las gramáticas nos marcan las reglas que han construido el lenguaje. Para detectar esas reglas nos podemos preguntar:

- ¿Es nuestro lenguaje una sucesión de símbolos donde el siguiente depende del anterior (o del siguiente)? **en cualquiera de los casos vemos una relación simétrica.**
- ¿Se puede separar cada palabra del lenguaje en partes más pequeñas que no dependan unas de otras?
- ¿Hay elementos que cambien su significado dependiendo del contexto?
- No se ajusta a ninguna de estas premisas pero se reconoce su estructura de alguna manera (mediante algún algoritmo).

Ideas y Nociones

Las gramáticas nos marcan las reglas que han construido el lenguaje. Para detectar esas reglas nos podemos preguntar:

- ¿Es nuestro lenguaje una sucesión de símbolos donde el siguiente depende del anterior (o del siguiente)? **en cualquiera de los casos vemos una relación simétrica.**
- ¿Se puede separar cada palabra del lenguaje en partes más pequeñas que no dependan unas de otras?
- ¿Hay elementos que cambien su significado dependiendo del contexto?
- No se ajusta a ninguna de estas premisas pero se reconoce su estructura de alguna manera (mediante algún algoritmo).

Ideas y Nociones

Las gramáticas nos marcan las reglas que han construido el lenguaje. Para detectar esas reglas nos podemos preguntar:

- ¿Es nuestro lenguaje una sucesión de símbolos donde el siguiente depende del anterior (o del siguiente)? en cualquiera de los casos vemos una relación simétrica.
- ¿Se puede separar cada palabra del lenguaje en partes más pequeñas que no dependan unas de otras?
- ¿Hay elementos que cambien su significado dependiendo del contexto?
- No se ajusta a ninguna de estas premisas pero se reconoce su estructura de alguna manera (mediante algún algoritmo).

Ideas y Nociones

Las gramáticas nos marcan las reglas que han construido el lenguaje. Para detectar esas reglas nos podemos preguntar:

- ¿Es nuestro lenguaje una sucesión de símbolos donde el siguiente depende del anterior (o del siguiente)? en cualquiera de los casos vemos una relación simétrica.
- ¿Se puede separar cada palabra del lenguaje en partes más pequeñas que no dependan unas de otras?
- ¿Hay elementos que cambien su significado dependiendo del contexto?
- No se ajusta a ninguna de estas premisas pero se reconoce su estructura de alguna manera (mediante algún algoritmo).

Ejemplos

- Las sumas expresiones matemáticas que involucren simplemente números sumas y restas. ¿Que significa $3 - 4 + 2$?
- ¿Que pasa si a esa expresión le añadimos paréntesis de este modo $3 - (4 + 2)$?
- En JAVA, ¿que significa $i + j$?, ¿es lo mismo que las variables i, j sean strings que sean enteros ó números en punto flotante?
- Pensemos ahora en alguna lengua humana, como español, ¿es fácil detectar ironía?

Ejemplos

- Las sumas expresiones matemáticas que involucren simplemente números sumas y restas. ¿Que significa $3 - 4 + 2$?
- ¿Que pasa si a esa expresión le añadimos paréntesis de este modo $3 - (4 + 2)$?
- En JAVA, ¿que significa $i + j$?, ¿es lo mismo que las variables i, j sean strings que sean enteros ó números en punto flotante?
- Pensemos ahora en alguna lengua humana, como español, ¿es fácil detectar ironía?

Ejemplos

- Las sumas expresiones matemáticas que involucren simplemente números sumas y restas. ¿Que significa $3 - 4 + 2$?
- ¿Que pasa si a esa expresión le añadimos paréntesis de este modo $3 - (4 + 2)$?
- En JAVA, ¿que significa $i + j$?, ¿es lo mismo que las variables i, j sean strings que sean enteros ó números en punto flotante?
- Pensemos ahora en alguna lengua humana, como español, ¿es fácil detectar ironía?

Ejemplos

- Las sumas expresiones matemáticas que involucren simplemente números sumas y restas. ¿Que significa $3 - 4 + 2$?
- ¿Que pasa si a esa expresión le añadimos paréntesis de este modo $3 - (4 + 2)$?
- En JAVA, ¿que significa $i + j$?, ¿es lo mismo que las variables i, j sean strings que sean enteros ó números en punto flotante?
- Pensemos ahora en alguna lengua humana, como español, ¿es fácil detectar ironía?

Jerarquía de Chomsky

Definición (Gramáticas Regulares o de Tipo 3)

Definiremos las gramáticas con producciones lineales del modo siguiente:

Llamaremos gramática lineal por la izquierda a toda

$G := (V, \Sigma, Q_0, P)$ gramática tal que todas las producciones de P son de uno de los dos tipos siguientes:

- *$A \mapsto a$, donde $A \in V$ y $a \in \Sigma \cup \{\lambda\}$.*
- *$A \mapsto aB$, donde $A, B \in V$ y $a \in \Sigma \cup \{\lambda\}$.*

Ejemplo de Gramática Linear

Sea $G := (\{Q_0, Op\}, \{Num, +, -\}, Q_0, P)$ gramática la siguiente gramática:

$$P = \{Q_0 \mapsto Num Op \mid Num, Op \mapsto +Q_0 \mid -Q_0\}.$$

Esta gramática genera el lenguaje de todas las operaciones aritméticas que involucren solo sumas y restas.

Jerarquía de Chomsky

Definición (Gramáticas Libres de Contexto o de Tipo 2)

Llamaremos *gramática libre de contexto* a toda $G = (V, \Sigma, Q_0, P)$ gramática tal que todas las producciones de P son del tipo siguiente:

$$A \mapsto \omega, \text{ donde } A \in V \text{ y } \omega \in (\Sigma \cup V)^*.$$

Un lenguaje **libre de contexto** es un lenguaje generado por una gramática libre de contexto.

Ejemplo

Tomemos la gramática G que habíamos definido anteriormente y añadamos una variable

$G := (\{Q_0, Q_1, Op\}, \{Num, +, -, (,)\}, Q_1, P)$. Las producciones van a ser ahora

$$P = \{Q_0 \mapsto Q_1 Op \mid Q_1, Op \mapsto +Q_1 \mid -Q_1, Q_1 \mapsto Num \mid (Q_0)\}.$$

Jerarquía de Chomsky

Definición (Gramáticas sensibles al contexto o de Tipo 1)

Llamaremos *gramática sensible al contexto* a toda $G = (V, \Sigma, Q_0, P)$ gramática tal que todas las producciones de P son del tipo siguiente:

$$\gamma A \delta \mapsto \gamma \omega \delta, \text{ donde } A \in V \text{ y } \omega, \gamma, \delta \in (\Sigma \cup V)^*, \omega \neq \lambda.$$

Un lenguaje **sensible al contexto** es un lenguaje generado por una gramática libre de contexto.

Jerarquía de Chomsky

Supongamos que queremos definir el lenguaje que **solo** acepta programas donde las variables hayan sido especificadas. Por simplicidad supongamos que solo vamos a utilizar programas con una sola variable.

Jerarquía de Chomsky

Boolean a

a=true

Es un programa válido, pero

Boolean a

b=true

no lo es.

Jerarquía de Chomsky

Una posible gramática (solo en el caso de los identificadores estén formados por las letras a,b) :

$$\begin{aligned} S &\mapsto \textit{Boolean CD=true}, \\ C &\mapsto \textit{aCA|bCB|\lambda}, D \mapsto \lambda \\ AD &\mapsto \textit{aD}, BD \mapsto \textit{bD} \\ Aa &\mapsto \textit{aA}, Ab \mapsto \textit{bA} \\ Ba &\mapsto \textit{aB}, Bb \mapsto \textit{bB} \end{aligned}$$

Jerarquía de Chomsky

Definición (Gramáticas de Tipo 0)

Llamaremos *gramática de tipo 0* a toda $G = (V, \Sigma, Q_0, P)$ gramática tal que todas las producciones de P son del tipo siguiente:

$$\gamma \mapsto \omega, \text{ donde } \gamma, \omega \in (\Sigma \cup V)^*.$$

Jerarquía de Chomsky

Notar que un lenguaje regular es siempre un lenguaje libre de contexto. Nuestra preocupación será encontrar cual es la mínima estructura que tiene un lenguaje.