Simplificación de gramáticas

Clase 21

IIC2223 / IIC2224

Prof. Cristian Riveros

Outline

Variables inútiles (clase ant.)

Producciones inútiles

Outline

Variables inútiles (clase ant.)

Producciones inútiles

Variables inútiles (recordatorio)

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una CFG.

Definición

Diremos que una variable $X \in V$ es **útil** si existe una derivación:

$$S \overset{\star}{\Rightarrow} \alpha X \beta \overset{\star}{\Rightarrow} w$$

Al contrario, diremos que una variable X es **inútil** si NO es útil.

¿qué variables son inútiles?

$$\begin{array}{ccl} S & \rightarrow & aAa \mid aBC \\ A & \rightarrow & aS \mid bD \\ B & \rightarrow & aBa \mid b \\ C & \rightarrow & abb \mid DD \\ D & \rightarrow & aDa \end{array}$$

Desde el lado positivo: símbolos útiles (recordatorio)

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una CFG.

Definición

Para una variable $X \in V$:

1. Decimos que X es **alcanzable** si existe una derivación:

$$S \stackrel{\star}{\Rightarrow} \alpha X \beta$$

2. Decimos que X es **generadora** si existe una derivación:

$$X \stackrel{\star}{\Rightarrow} w$$

¿cómo determinamos si una variable es **alcanzable** o **generadora**?

Variables **alcanzables** (recordatorio)

input : Gramática $\mathcal{G} = (V, \Sigma, P, S)$

output: Conjunto C de variables alcanzables

Function `alcanzables` *($\mathcal{G}$)*

- **let** $C_0 := \{S\}$
- **let** $C := \varnothing$
- **while** $C_0 \neq \varnothing$ **do**
  - **take** $Y \in C_0$
  - $C_0 := C_0 - \{Y\}$
  - $C := C \cup \{Y\}$
  - **foreach** $X \in V - C$ *tal que existe una regla* $\left(Y \to \alpha X \beta\right) \in P$ **do**
    - $C_0 := C_0 \cup \{X\}$
- **return** C

Variables **generadoras** (recordatorio)

input : Gramática $\mathcal{G} = (V, \Sigma, P, S)$

output: Conjunto G de variables generadoras

```
Function Generadores (𝒢)
    let G₀ := {X ∈ V | (X → w) ∈ P}
    let G := ∅
    while G₀ ≠ G do
        G := G₀
        foreach (X → α) ∈ P do
            if todas las variables en α estan en G then
                G₀ := G₀ ∪ {X}
    return G
```

Eliminación de variables inútiles

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una CFG.

Teorema

Sea $\mathcal{G}''$ una gramática creada a partir de $\mathcal{G}$ después de:

- eliminar todos la variables y reglas **NO generadoras**.
- eliminar todas las variables y reglas **NO alcanzables**.

Entonces, $\mathcal{L}(\mathcal{G}'') = \mathcal{L}(\mathcal{G})$ y $\mathcal{G}''$ no contiene variables inútiles.

¿qué falla al eliminar primero las no alcanzables y después las no generadoras?

Ejemplo

$$
\mathcal{G}: \quad \begin{array}{rcl} \mathsf{S} & \to & \mathsf{AB} \mid \mathsf{b} \\ \mathsf{A} & \to & \mathsf{a} \\ \mathsf{B} & \to & \mathsf{B} \end{array}
$$

Al eliminar variables **no alcanzables** en $\mathcal{G}$:

$$
\mathcal{G}': \quad \begin{array}{rcl} \mathsf{S} & \to & \mathsf{AB} \mid \mathsf{b} \\ \mathsf{A} & \to & \mathsf{a} \\ \mathsf{B} & \to & \mathsf{B} \end{array}
$$

Al eliminar variables **no generadoras** en $\mathcal{G}'$:

$$
\mathcal{G}'': \quad \begin{array}{rcl} \mathsf{S} & \to & \mathsf{b} \\ \mathsf{A} & \to & \mathsf{a} \end{array}
$$

Eliminación de variables inútiles

Teorema

Sea $\mathcal{G}''$ una gramática creada a partir de $\mathcal{G}$ después de: ...
Entonces, $\mathcal{L}(\mathcal{G}'') = \mathcal{L}(\mathcal{G})$ y $\mathcal{G}''$ no contiene variables inútiles.

Demostración

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una CFG.

Sea $\mathcal{G}' = (V', \Sigma, P', S)$ al eliminar las variables **no generadoras** de $\mathcal{G}$:

$$
\begin{array}{rcl}
V' & = & \{\, X \in V \mid \exists w.\ X \overset{\star}{\underset{\mathcal{G}}{\Rightarrow}} w \,\} \\
P' & = & \{\, X \to \alpha \in P \mid X \in V' \wedge \alpha \in (V' \cup \Sigma)^* \,\}
\end{array}
$$

Sea $\mathcal{G}'' = (V'', \Sigma, P'', S)$ al eliminar las variables **no alcanzables** de $\mathcal{G}'$:

$$
\begin{array}{rcl}
V'' & = & \{\, X \in V' \mid \exists \alpha, \beta.\ S \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} \alpha X \beta \,\} \\
P'' & = & \{\, X \to \alpha \in P' \mid X \in V'' \wedge \alpha \in (V'' \cup \Sigma)^* \,\}
\end{array}
$$

Eliminación de variables inútiles

Demostración

Sea $\mathcal{G}' = (V', \Sigma, P', S)$ tal que $V' = \{\, X \in V \mid \exists w.\ X \overset{\star}{\underset{\mathcal{G}}{\Rightarrow}} w \,\}$.

Sea $\mathcal{G}'' = (V'', \Sigma, P'', S)$ tal que $V'' = \{\, X \in V' \mid \exists \alpha, \beta.\ S \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} \alpha X \beta \,\}$.

Considere las siguientes propiedades de $\mathcal{G}$, $\mathcal{G}'$ y $\mathcal{G}''$.

1. Para todo $\alpha \in (V \cup \Sigma)^*$, si $\alpha \overset{\star}{\underset{\mathcal{G}}{\Rightarrow}} w$ entonces $\alpha \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} w$.
2. Para todo $\alpha \in (V' \cup \Sigma)^*$, si $S \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} \alpha$ entonces $S \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} \alpha$.
3. Para todo $\alpha \in (V'' \cup \Sigma)^*$, si $\alpha \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} w$ entonces $\alpha \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} w$.

Ejercicio: demuestre las propiedades.

Eliminación de variables inútiles

Demostración

1. Para todo $\alpha \in (V \cup \Sigma)^*$, si $\alpha \overset{\star}{\underset{\mathcal{G}}{\Rightarrow}} w$ entonces $\alpha \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} w$.
2. Para todo $\alpha \in (V' \cup \Sigma)^*$, si $S \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} \alpha$ entonces $S \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} \alpha$.
3. Para todo $\alpha \in (V'' \cup \Sigma)^*$, si $\alpha \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} w$ entonces $\alpha \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} w$.

PD: $\mathcal{L}(\mathcal{G}'') = \mathcal{L}(\mathcal{G})$.

Como $V'' \subseteq V$ y $P'' \subseteq P$, entonces $\mathcal{L}(\mathcal{G}'') \subseteq \mathcal{L}(\mathcal{G})$.

PD: $\mathcal{L}(\mathcal{G}) \subseteq \mathcal{L}(\mathcal{G}'')$.

Sea $w \in \mathcal{L}(\mathcal{G})$ tal que $S \overset{\star}{\underset{\mathcal{G}}{\Rightarrow}} w$.

- Por la propiedad 1. tenemos que $S \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} w$.
- Por la propiedad 2. tenemos que $S \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} w$.

Por lo tanto $w \in \mathcal{L}(\mathcal{G}'')$ y concluimos que $\mathcal{L}(\mathcal{G}) \subseteq \mathcal{L}(\mathcal{G}'')$.

Eliminación de variables inútiles

Demostración

PD: Para todo $X \in V''$, X es **útil en** $\mathcal{G}''$.

Como $X \in V''$, entonces $S \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} \alpha X \beta$ para algún $\alpha, \beta \in (V' \cup \Sigma)^*$.

Por la propiedad 2. se tiene que: $S \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} \alpha X \beta$ y $\alpha, \beta \in (V'' \cup \Sigma)^*$.

Como $X \in V'$ y $\alpha, \beta \in (V' \cup \Sigma)^*$, entonces existen u, v, w tal que:

$$\alpha \overset{\star}{\underset{\mathcal{G}}{\Rightarrow}} u \ , \quad X \overset{\star}{\underset{\mathcal{G}}{\Rightarrow}} v \ , \quad \beta \overset{\star}{\underset{\mathcal{G}}{\Rightarrow}} w$$

Por la propiedad 1. se tiene que: $\alpha \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} u \ , \ X \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} v \ , \ \beta \overset{\star}{\underset{\mathcal{G}'}{\Rightarrow}} w$.

Por la propiedad 3. se tiene que: $\alpha \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} u \ , \ X \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} v \ , \ \beta \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} w$.

Juntando todo $S \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} \alpha X \beta \overset{\star}{\underset{\mathcal{G}''}{\Rightarrow}} uvw$ y X es útil en $\mathcal{G}''$. ■

Lenguaje vació de una gramática

¿cómo podemos verificar si $\mathcal{L}(\mathcal{G}) \neq \varnothing$?

Algoritmo para verificar lenguaje vacío

- Ejecuto G = Generadores($\mathcal{G}$).
- Si $S \in G$ acepto.
- Si no, rechazo.

Outline

¿cómo podemos simplificar está gramática?

$$
\begin{array}{lcl}
S & \rightarrow & aAa \mid aB \\
A & \rightarrow & B \mid C \\
B & \rightarrow & aBa \mid b \\
C & \rightarrow & bC \mid Da \\
D & \rightarrow & bCa \mid \epsilon
\end{array}
$$

¿cuáles producciones son **poco útiles**?

Producciones en vacío y unitarias

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una CFG.

Definición

- Decimos que una producción de la forma: $X \to \epsilon$ es **en vacío**.
- Decimos que una producción de la forma: $X \to Y$ es **unitaria**.

¿cuáles producciones son **en vacío** y cuáles **unitarias**?

$$
\begin{array}{lcl}
\mathrm{S} & \to & \mathrm{aAa} \mid \mathrm{aBC} \\
\mathrm{A} & \to & \mathrm{aD} \mid \mathrm{S} \\
\mathrm{B} & \to & \mathrm{aBa} \mid \epsilon \\
\mathrm{C} & \to & \mathrm{abb} \mid \mathrm{D} \\
\mathrm{D} & \to & \mathrm{aDa} \mid \mathrm{a}
\end{array}
$$

¿cómo eliminamos las producciones **en vacío** y **unitarias**?

¿és siempre posible eliminar las producciones en vacío?

Ejemplo

$$S \quad \rightarrow \quad a\ S\ b \mid \epsilon$$

Conclusión

Si $\epsilon \in \mathcal{L}(\mathcal{G})$, entonces
NO se pueden borrar las producciones en vacío sin alterar el lenguaje $\mathcal{G}$.

Desde ahora supondremos que $\epsilon \notin \mathcal{L}(\mathcal{G})$

¿es razonable?

¿cómo eliminamos las producciones en vacío y unitarias?

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una CFG tal que $\epsilon \notin \mathcal{L}(\mathcal{G})$.

Observaciones

Suponga las reglas $X \to Y$ y $Y \to \gamma$ en P.

- Si $\mathcal{G}' = (V, \Sigma, P \cup \{X \to \gamma\}, S)$, entonces: ¿$\mathcal{L}(\mathcal{G}') = \mathcal{L}(\mathcal{G})$? ✓
- Si $\mathcal{G}'' = (V, \Sigma, (P \cup \{X \to \gamma\}) \setminus \{X \to Y\}, S)$: ¿$\mathcal{L}(\mathcal{G}'') = \mathcal{L}(\mathcal{G})$? ✗

Suponga las reglas $X \to \epsilon$ y $Z \to \alpha X \beta$ en P.

- Si $\mathcal{G}' = (V, \Sigma, P \cup \{Z \to \alpha\beta\}, S)$, entonces: ¿$\mathcal{L}(\mathcal{G}') = \mathcal{L}(\mathcal{G})$? ✓
- Si $\mathcal{G}'' = (V, \Sigma, (P \cup \{Z \to \alpha\beta\}) \setminus \{X \to \epsilon\}, S)$: ¿$\mathcal{L}(\mathcal{G}'') = \mathcal{L}(\mathcal{G})$? ✗

Conclusión: debemos agregar todas las posibles "composiciones" de reglas antes de remover las producciones en vacío o unitarias.

¿cómo eliminamos las producciones en vacío y unitarias?

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una CFG tal que $\epsilon \notin \mathcal{L}(\mathcal{G})$.

Clausura de producciones unitarias y en vacío

Sea P^* el **menor conjunto de producciones** que contiene a P y **cerrado bajo** las siguientes reglas:

1. Si $X \to Y \in P^*$ y $Y \to \gamma \in P^*$, entonces $X \to \gamma \in P^*$.
2. Si $X \to \epsilon \in P^*$ y $Z \to \alpha X \beta \in P^*$, entonces $Z \to \alpha\beta \in P^*$.

Defina $\mathcal{G}^* = (V, \Sigma, P^*, S)$. Entonces:

- P^* es finito (¿por qué?)
- $\mathcal{L}(\mathcal{G}^*) = \mathcal{L}(\mathcal{G})$ (¿por qué?)

¿cómo eliminamos las producciones en vacío y unitarias?

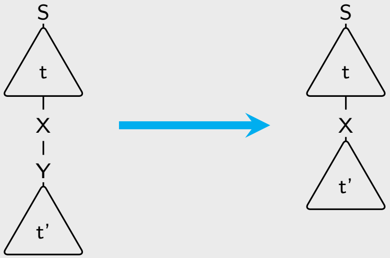
Para cualquier palabra $w \in \mathcal{L}(\mathcal{G}^*)$,
sea $\mathcal{T}$ un árbol de derivación de w en $\mathcal{G}^*$ de **tamaño mínimo**.

Propiedad 1

El árbol de derivación $\mathcal{T}$ NO usa una **producción unitaria**.

Demostración (por contradicción)

Suponemos que $\mathcal{T}$ usa una producción unitaria:



¿cómo eliminamos las producciones en vacío y unitarias?

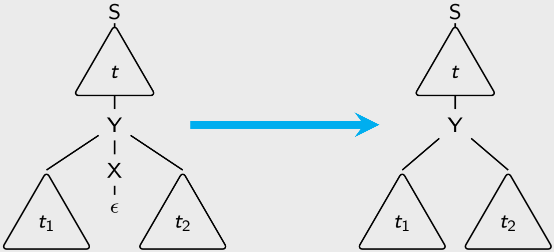
Para cualquier palabra $w \in \mathcal{L}(\mathcal{G}^*)$,
sea $\mathcal{T}$ un árbol de derivación de w en $\mathcal{G}^*$ de **tamaño mínimo**.

Propiedad 2

El árbol de derivación $\mathcal{T}$ NO usa una producción **en vacío**.

Demostración (por contradicción)

Suponemos que $\mathcal{T}$ usa una producción en vacío:



¿cómo eliminamos las producciones en vacío y unitarias?

Por la **Propiedad 1** y **Propiedad 2** tenemos que:

> *Para todo $w \in \mathcal{L}(\mathcal{G}^*)$, existe una derivación de w en $\mathcal{G}^*$ que NO usa producciones* **en vacío** *ni producciones* **unitarias***.*

Podemos eliminar las producciones en vacío y unitarias de $\mathcal{G}^*$!

Teorema

Para toda CFG $\mathcal{G}$ tal que $\epsilon \notin \mathcal{L}(\mathcal{G})$, sea:

- $\mathcal{G}^*$ la clausura de producciones unitarias y en vacío.
- $\hat{\mathcal{G}}$ el resultado de remover toda producción unitaria o en vacío de $\mathcal{G}^*$.

Entonces $\mathcal{L}(\hat{\mathcal{G}}) = \mathcal{L}(\mathcal{G})$ y $\hat{\mathcal{G}}$ no tiene producciones unitarias o en vacío.

¿cómo eliminamos las producciones en vacío y unitarias?

Sea $\mathcal{G} = (V, \Sigma, P, S)$ una CFG tal que $\epsilon \notin \mathcal{L}(\mathcal{G})$.

Para eliminar las producciones en vacío o unitarias de $\mathcal{G}$:

- construimos $\mathcal{G}^*$ haciendo la **clausura** de prod. unitarias y en vacío,
- construimos $\hat{\mathcal{G}}$ **removiendo** todas las prod. unitarias o en vacío de $\mathcal{G}^*$.

Por el resultado anterior sabemos que $\mathcal{L}(\mathcal{G}) = \mathcal{L}(\hat{\mathcal{G}})$.

Importante: es posible que $\hat{\mathcal{G}}$ contiene variables inútiles.

Cierre de clase

En esta clase vimos:

- Eliminación de variables inútiles.
- Eliminación de producciones inútiles.

Próxima clase: Forma normal de Chomsky.