CSEN 1002

Task 4: CFG Epsilon & Unit Rules Elimination [1]

[1] These slides are based on Lecture 10 of CSEN 502 by Assoc. Prof. Haythem O. Ismail

Table of Contents

Table of Contents

Definition

A context-free grammar (CFG) is a 4-tuple $(V; \Sigma; R; S)$, where:

Definition

A context-free grammar (CFG) is a 4-tuple $(V; \Sigma; R; S)$, where:

1. V is the set of variables (Uppercase)

Definition

A context-free grammar (CFG) is a 4-tuple $(V; \Sigma; R; S)$, where:

1. V is the set of variables (Uppercase)
2. Σ is the set of terminals (Lowercase).

Definition

A context-free grammar (CFG) is a 4-tuple $(V; \Sigma; R; S)$, where:

1. V is the set of variables (Uppercase)
2. Σ is the set of terminals (Lowercase).
   - In this case, the names of the lexical categories.

Definition

A context-free grammar (CFG) is a 4-tuple $(V; \Sigma; R; S)$, where:

1. V is the set of variables (Uppercase)
2. Σ is the set of terminals (Lowercase).
   - In this case, the names of the lexical categories.
3. R is the set of production rules.

Definition

A context-free grammar (CFG) is a 4-tuple $(V; \Sigma; R; S)$, where:

1. V is the set of variables (Uppercase)
2. Σ is the set of terminals (Lowercase).
   - In this case, the names of the lexical categories.
3. R is the set of production rules.
   - Represented in the format $V \longrightarrow (\Sigma \cup V)^*$

Definition

A context-free grammar (CFG) is a 4-tuple $(V; \Sigma; R; S)$, where:

1. V is the set of variables (Uppercase)
2. Σ is the set of terminals (Lowercase).
   - In this case, the names of the lexical categories.
3. R is the set of production rules.
   - Represented in the format $V \longrightarrow (\Sigma \cup V)^*$
4. S is the start variable

Table of Contents

ε-Rules Elimination

- For each rule, $r \in R$, of the form $A \longrightarrow \varepsilon$ $(A \neq S)$ do

ε-Rules Elimination

- For each rule, $r \in R$, of the form $A \longrightarrow \varepsilon$ $(A \neq S)$ do
  1. Let $R = R - \{r\}$.

ε-Rules Elimination

- For each rule, $r \in R$, of the form $A \longrightarrow \varepsilon$ $(A \neq S)$ do
  1. Let $R = R - \{r\}$.
  2. If $B \longrightarrow uAv \in R$, (where $u, v \in (V \cup \Sigma)^*$ and A is a proper substring of uAv), let $R = R \cup \{B \longrightarrow uv\}$
     - Note: This should be done for each occurrence of A on the right-hand side of a rule. That is, it should be repeated for each possible choice of u and v.

ε-Rules Elimination

- For each rule, $r \in R$, of the form $A \longrightarrow \varepsilon$ $(A \neq S)$ do
  1. Let $R = R - \{r\}$.
  2. If $B \longrightarrow uAv \in R$, (where $u, v \in (V \cup \Sigma)^*$ and A is a proper substring of uAv), let $R = R \cup \{B \longrightarrow uv\}$
     - Note: This should be done for each occurrence of A on the right-hand side of a rule. That is, it should be repeated for each possible choice of u and v.
  3. If $B \longrightarrow A \in R$, then unless $B \longrightarrow \varepsilon$ has already been removed, let $R = R \cup \{B \longrightarrow \varepsilon\}$.

Example

G_1 :

$$\begin{array}{lcl} S & \longrightarrow & ASA \mid \mathtt{a}B \\ A & \longrightarrow & B \mid S \\ B & \longrightarrow & \mathtt{b} \mid \varepsilon \end{array}$$

Example

G_1 :

$$\begin{array}{lcl} S & \longrightarrow & ASA \mid \mathtt{a}B \\ A & \longrightarrow & B \mid S \\ B & \longrightarrow & \mathtt{b} \mid \varepsilon \end{array}$$

G_2 :

$$\begin{array}{lcl} S & \longrightarrow & ASA \mid \mathbf{a} \mid \mathtt{a}B \\ A & \longrightarrow & B \mid S \mid \varepsilon \\ B & \longrightarrow & \mathtt{b} \end{array}$$

Example

G_1 :

$$\begin{array}{lcl} S & \longrightarrow & ASA \mid \mathtt{a}B \\ A & \longrightarrow & B \mid S \\ B & \longrightarrow & \mathtt{b} \mid \varepsilon \end{array}$$

G_2 :

$$\begin{array}{lcl} S & \longrightarrow & ASA \mid \mathbf{a} \mid \mathtt{a}B \\ A & \longrightarrow & B \mid S \mid \varepsilon \\ B & \longrightarrow & \mathtt{b} \end{array}$$

G_2 :

$$\begin{array}{lcl} S & \longrightarrow & \mathbf{AS} \mid ASA \mid \mathbf{S} \mid \mathbf{SA} \mid \mathtt{a} \mid \mathtt{a}B \\ A & \longrightarrow & B \mid S \\ B & \longrightarrow & \mathtt{b} \end{array}$$

Table of Contents

Unit-Rules Elimination

- For each rule, $r \in R$, of the form $A \longrightarrow B$ (where $B \in V$) do

Unit-Rules Elimination

- For each rule, $r \in R$, of the form $A \longrightarrow B$ (where $B \in V$) do
  1. Let $R = R - \{r\}$

Unit-Rules Elimination

- For each rule, $r \in R$, of the form $A \longrightarrow B$ (where $B \in V$) do
  1. Let $R = R - \{r\}$
  2. For every rule of the form $B \longrightarrow u \in R$ (where $u \in (V_1 \cup \Sigma)^+$ and $u \notin V$), let $R = R \cup \{A \longrightarrow u\}$.

Unit-Rules Elimination

- For each rule, $r \in R$, of the form $A \longrightarrow B$ (where $B \in V$) do
  1. Let $R = R - \{r\}$
  2. For every rule of the form $B \longrightarrow u \in R$ (where $u \in (V_1 \cup \Sigma)^+$ and $u \notin V$), let $R = R \cup \{A \longrightarrow u\}$.
  3. For every rule of the form $B \longrightarrow C \in R$ (where $C \in V$), then unless $A \longrightarrow C$ has already been removed, let $R = R \cup \{A \longrightarrow C\}$.

Example (Cont'd)

G_2 :

$$
\begin{array}{lcl}
S & \longrightarrow & AS \mid ASA \mid S \mid SA \mid \mathtt{a} \mid \mathtt{a}B \\
A & \longrightarrow & B \mid S \\
B & \longrightarrow & \mathtt{b}
\end{array}
$$

Example (Cont'd)

G_2 :

$$\begin{array}{lcl} S & \longrightarrow & AS \mid ASA \mid S \mid SA \mid \mathtt{a} \mid \mathtt{a}B \\ A & \longrightarrow & B \mid S \\ B & \longrightarrow & \mathtt{b} \end{array}$$

G_3 :

$$\begin{array}{lcl} S & \longrightarrow & AS \mid ASA \mid SA \mid \mathtt{a} \mid \mathtt{a}B \\ A & \longrightarrow & \mathbf{AS} \mid \mathbf{ASA} \mid \mathbf{SA} \mid \mathbf{a} \mid \mathbf{aB} \mid \mathbf{b} \\ B & \longrightarrow & \mathtt{b} \end{array}$$

Table of Contents

CFG String Representation

G_1 :

$$\begin{array}{lcl} S & \longrightarrow & ASA \mid \mathtt{a}B \\ A & \longrightarrow & B \mid S \\ B & \longrightarrow & \mathtt{b} \mid \varepsilon \end{array}$$

String Representation

`S;A;B#a;b#S/ASA,aB;A/B,S;B/b,e`

CFG String Representation

G_3 :

$$
\begin{array}{lcl}
S & \longrightarrow & AS \mid ASA \mid SA \mid \texttt{a} \mid \texttt{a}B \\
A & \longrightarrow & AS \mid ASA \mid SA \mid a \mid aB \mid b \\
B & \longrightarrow & \texttt{b}
\end{array}
$$

String Representation

`S;A;B#a;b#S/AS,ASA,SA,a,aB;A/AS,ASA,SA,a,aB,b;B/b`