# Automata Theory :: Chomsky Normal Form

Jörg Endrullis

Vrije Universiteit Amsterdam

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

A variable *A* is called **erasable** if  $A \Rightarrow^+ \lambda$ .

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

A variable *A* is called **erasable** if  $A \Rightarrow^+ \lambda$ .

The set of erasable variables can be computed as follows:

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

A variable *A* is called **erasable** if  $A \Rightarrow^+ \lambda$ .

The set of erasable variables can be computed as follows:

• If  $A \rightarrow \lambda$ , then A is erasable.

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

## A variable *A* is called **erasable** if $A \Rightarrow^+ \lambda$ .

The set of erasable variables can be computed as follows:

- If  $A \rightarrow \lambda$ , then A is erasable.
- If  $A \rightarrow B_1 \cdots B_n$  and  $B_1, \ldots, B_n$  are erasable, then so is A.

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

### A variable A is called **erasable** if $A \Rightarrow^+ \lambda$ .

The set of erasable variables can be computed as follows:

- If  $A \rightarrow \lambda$ , then A is erasable.
- If  $A \rightarrow B_1 \cdots B_n$  and  $B_1, \ldots, B_n$  are erasable, then so is A.

$$S o AcB$$
  $A o CBC$   $B o abB$   $C o cCd$   $B o \lambda$   $C o BB$ 

We determine the set of erasable variables:

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

### A variable A is called **erasable** if $A \Rightarrow^+ \lambda$ .

The set of erasable variables can be computed as follows:

- If  $A \rightarrow \lambda$ , then A is erasable.
- If  $A \rightarrow B_1 \cdots B_n$  and  $B_1, \ldots, B_n$  are erasable, then so is A.

$$S o AcB$$
  $A o CBC$   $B o abB$   $C o cCd$   $B o \lambda$   $C o BB$ 

We determine the set of erasable variables:

■ *B* is erasable because of the rule  $B \rightarrow \lambda$ 

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

#### A variable A is called **erasable** if $A \Rightarrow^+ \lambda$ .

The set of erasable variables can be computed as follows:

- If  $A \rightarrow \lambda$ , then A is erasable.
- If  $A \rightarrow B_1 \cdots B_n$  and  $B_1, \ldots, B_n$  are erasable, then so is A.

$$egin{array}{lll} S 
ightarrow AcB & A 
ightarrow CBC & B 
ightarrow abB & C 
ightarrow cCd \ & B 
ightarrow \lambda & C 
ightarrow BB \ \end{array}$$

We determine the set of erasable variables:

- *B* is erasable because of the rule  $B \rightarrow \lambda$
- C is erasable because of  $C \rightarrow BB$  and B is erasable

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

### A variable *A* is called **erasable** if $A \Rightarrow^+ \lambda$ .

The set of erasable variables can be computed as follows:

- If  $A \rightarrow \lambda$ , then A is erasable.
- If  $A \rightarrow B_1 \cdots B_n$  and  $B_1, \ldots, B_n$  are erasable, then so is A.

$$S o AcB \qquad A o CBC \qquad B o abB \qquad C o cCd \ B o \lambda \qquad C o BB$$

We determine the set of erasable variables:

- *B* is erasable because of the rule  $B \rightarrow \lambda$
- C is erasable because of  $C \rightarrow BB$  and B is erasable
- A is erasable because of  $A \rightarrow CBC$  and B, C are erasable

A production rule  $A \rightarrow \lambda$  is called  $\lambda$ -production rule.

### A variable *A* is called **erasable** if $A \Rightarrow^+ \lambda$ .

The set of erasable variables can be computed as follows:

- If  $A \rightarrow \lambda$ , then A is erasable.
- If  $A \rightarrow B_1 \cdots B_n$  and  $B_1, \dots, B_n$  are erasable, then so is A.

$$S o AcB$$
  $A o CBC$   $B o abB$   $C o cCd$   $B o \lambda$   $C o BB$ 

We determine the set of erasable variables:

- *B* is erasable because of the rule  $B \rightarrow \lambda$
- C is erasable because of  $C \rightarrow BB$  and B is erasable
- lacksquare A is erasable because of A o CBC and B, C are erasable

So the variables A, B, C are erasable.

#### **Theorem**

For every context-free language L there exists a context-free grammar G without  $\lambda$ -rules such that  $L(G) = L \setminus \{\lambda\}$ .

#### Theorem

For every context-free language L there exists a context-free grammar G without  $\lambda$ -rules such that  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let *G* be a context-free grammar with L(G) = L.

#### **Theorem**

For every context-free language L there exists a context-free grammar G without  $\lambda$ -rules such that  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let G be a context-free grammar with L(G) = L.

■ Determine all erasable variables (that is, variables  $A \Rightarrow^* \lambda$ ).

#### **Theorem**

For every context-free language L there exists a context-free grammar G without  $\lambda$ -rules such that  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let *G* be a context-free grammar with L(G) = L.

- Determine all erasable variables (that is, variables  $A \Rightarrow^* \lambda$ ).
- For every rule  $A \rightarrow xBy$  with  $B \Rightarrow^* \lambda$ , add a rule  $A \rightarrow xy$ .

#### **Theorem**

For every context-free language L there exists a context-free grammar G without  $\lambda$ -rules such that  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let G be a context-free grammar with L(G) = L.

- Determine all erasable variables (that is, variables  $A \Rightarrow^* \lambda$ ).
- For every rule  $A \rightarrow xBy$  with  $B \Rightarrow^* \lambda$ , add a rule  $A \rightarrow xy$ .
- Remove all λ-production rules.

#### **Theorem**

For every context-free language L there exists a context-free grammar G without  $\lambda$ -rules such that  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let G be a context-free grammar with L(G) = L.

- Determine all erasable variables (that is, variables  $A \Rightarrow^* \lambda$ ).
- For every rule  $A \rightarrow xBy$  with  $B \Rightarrow^* \lambda$ , add a rule  $A \rightarrow xy$ .
- Remove all λ-production rules.

The resulting grammar G has the property  $L(G) = L \setminus \{\lambda\}$ .

Consider the following grammar

$$S o ABaC$$
  $A o BC$   $B o b \mid \lambda$   $D o d$ 

$$\textbf{\textit{C}} \rightarrow \textbf{\textit{D}} \mid \lambda$$

Consider the following grammar

$$S 
ightarrow ABaC$$
  $A 
ightarrow BC$   $B 
ightarrow b \mid \lambda$   $D 
ightarrow d$   $C 
ightarrow D \mid \lambda$ 

What variables are erasable?

Consider the following grammar

$$egin{aligned} \mathcal{S} 
ightarrow \mathit{ABaC} & A 
ightarrow \mathit{BC} & B 
ightarrow \mathit{b} \mid \lambda & D 
ightarrow \mathit{d} \ & C 
ightarrow \mathit{D} \mid \lambda & \end{aligned}$$

What variables are erasable?

■ *A*, *B* and *C* 

Consider the following grammar

$$S 
ightarrow ABaC$$
  $A 
ightarrow BC$   $B 
ightarrow b \mid \lambda$   $D 
ightarrow d$   $C 
ightarrow D \mid \lambda$ 

What variables are erasable?

A, B and C

Consider the following grammar

$$egin{aligned} S 
ightarrow ABaC & A 
ightarrow BC & B 
ightarrow b \mid \lambda & D 
ightarrow d \ C 
ightarrow D \mid \lambda & \end{aligned}$$

What variables are erasable?

A, B and C

$$egin{aligned} S &
ightarrow ABaC \ A &
ightarrow BC \ B &
ightarrow b \mid \lambda \ \end{array} \qquad \qquad C &
ightarrow D \mid \lambda \ \end{array} \qquad \qquad D &
ightarrow d \end{aligned}$$

Consider the following grammar

$$egin{aligned} S 
ightarrow ABaC & A 
ightarrow BC & B 
ightarrow b \mid \lambda & D 
ightarrow d \ C 
ightarrow D \mid \lambda & \end{aligned}$$

What variables are erasable?

A, B and C

Construct the resulting grammar after removing all  $\lambda$ -rules:

$$S
ightarrow ABaC\mid BaC$$
 $A
ightarrow BC$ 
 $B
ightarrow b\mid \lambda$ 
 $C
ightarrow D\mid \lambda$ 

Consider the following grammar

$$egin{aligned} \mathcal{S} 
ightarrow \mathit{ABaC} & A 
ightarrow \mathit{BC} & B 
ightarrow \mathit{b} \mid \lambda & D 
ightarrow \mathit{d} \ & C 
ightarrow \mathit{D} \mid \lambda & \end{aligned}$$

What variables are erasable?

■ *A*, *B* and *C* 

Construct the resulting grammar after removing all  $\lambda$ -rules:

$$S
ightarrow ABaC \mid BaC \mid AaC$$
 $A
ightarrow BC$ 
 $B
ightarrow b\mid \lambda$ 
 $C
ightarrow D\mid \lambda$ 

Consider the following grammar

$$egin{aligned} S 
ightarrow ABaC & A 
ightarrow BC & B 
ightarrow b \mid \lambda & D 
ightarrow d \ C 
ightarrow D \mid \lambda & \end{aligned}$$

What variables are erasable?

■ *A*, *B* and *C* 

Construct the resulting grammar after removing all  $\lambda$ -rules:

$$S
ightarrow ABaC \mid BaC \mid AaC \mid ABa$$
  $A
ightarrow BC$   $B
ightarrow b \mid \lambda$   $C
ightarrow D \mid \lambda$ 

Consider the following grammar

$$egin{aligned} \mathcal{S} 
ightarrow \mathit{ABaC} & A 
ightarrow \mathit{BC} & B 
ightarrow \mathit{b} \mid \lambda & D 
ightarrow \mathit{d} \ & C 
ightarrow \mathit{D} \mid \lambda & \end{aligned}$$

What variables are erasable?

■ *A*, *B* and *C* 

Construct the resulting grammar after removing all  $\lambda$ -rules:

$$S 
ightarrow ABaC \mid BaC \mid AaC \mid ABa \mid aC$$
 $A 
ightarrow BC$ 

$$B o b \mid \lambda$$
  $C o D \mid \lambda$ 

Consider the following grammar

$$egin{aligned} S 
ightarrow ABaC & A 
ightarrow BC & B 
ightarrow b \mid \lambda & D 
ightarrow d \ C 
ightarrow D \mid \lambda & \end{aligned}$$

What variables are erasable?

■ *A*, *B* and *C* 

$$S o ABaC\mid BaC\mid AaC\mid ABa\mid aC\mid Ba$$
 $A o BC$ 
 $B o b\mid \lambda$   $C o D\mid \lambda$   $D o d$ 

#### Consider the following grammar

$$S 
ightarrow ABaC$$
  $A 
ightarrow BC$   $B 
ightarrow b \mid \lambda$   $D 
ightarrow d$   $C 
ightarrow D \mid \lambda$ 

What variables are erasable?

■ A, B and C

$$S o ABaC \mid BaC \mid AaC \mid ABa \mid aC \mid Ba \mid Aa$$
  
 $A o BC$   
 $B o b \mid \lambda$   $C o D \mid \lambda$   $D o d$ 

Consider the following grammar

$$egin{aligned} \mathcal{S} 
ightarrow egin{aligned} \mathcal{A} eta \mathcal{B} \mathcal{C} & \mathcal{B} 
ightarrow \mathcal{B} \mid \lambda & \mathcal{D} 
ightarrow \mathcal{D} \\ \mathcal{C} 
ightarrow \mathcal{D} \mid \lambda & \mathcal{C} \end{aligned}$$

What variables are erasable?

■ *A*, *B* and *C* 

$$S o ABaC \mid BaC \mid AaC \mid ABa \mid aC \mid Ba \mid Aa \mid a$$
  $A o BC$   $B o b \mid \lambda$   $C o D \mid \lambda$   $D o d$ 

Consider the following grammar

$$S o ABaC$$
  $A o BC$   $B o b \mid \lambda$   $D o d$   $C o D \mid \lambda$ 

What variables are erasable?

■ *A*, *B* and *C* 

$$S o ABaC\mid BaC\mid AaC\mid ABa\mid aC\mid Ba\mid Aa\mid a$$
  
 $A o BC\mid C$   
 $B o b\mid \lambda$   $C o D\mid \lambda$   $D o d$ 

Consider the following grammar

$$egin{aligned} \mathcal{S} 
ightarrow egin{aligned} \mathcal{A} 
ightarrow BC & B 
ightarrow b \mid \lambda & D 
ightarrow d \ C 
ightarrow D \mid \lambda & \end{aligned}$$

What variables are erasable?

■ *A*, *B* and *C* 

$$S o ABaC\mid BaC\mid AaC\mid ABa\mid aC\mid Ba\mid Aa\mid a$$
  
 $A o BC\mid C\mid B$   
 $B o b\mid \lambda$   $C o D\mid \lambda$   $D o d$ 

#### Consider the following grammar

$$S o ABaC$$
  $A o BC$   $B o b \mid \lambda$   $D o d$   $C o D \mid \lambda$ 

What variables are erasable?

■ *A*, *B* and *C* 

$$S 
ightarrow ABaC \mid BaC \mid AaC \mid ABa \mid aC \mid Ba \mid Aa \mid a$$
  
 $A 
ightarrow BC \mid C \mid B \mid \lambda$   
 $B 
ightarrow b \mid \lambda$   $C 
ightarrow D \mid \lambda$   $D 
ightarrow d$ 

Consider the following grammar

$$egin{aligned} \mathcal{S} 
ightarrow egin{aligned} \mathcal{A} eta \mathcal{B} \mathcal{C} & \mathcal{B} 
ightarrow \mathcal{B} \mid \lambda & \mathcal{D} 
ightarrow \mathcal{D} \\ \mathcal{C} 
ightarrow \mathcal{D} \mid \lambda & \mathcal{C} \end{aligned}$$

What variables are erasable?

A, B and C

$$S 
ightarrow ABaC \mid BaC \mid AaC \mid ABa \mid aC \mid Ba \mid Aa \mid a$$
  
 $A 
ightarrow BC \mid C \mid B$   
 $B 
ightarrow b$   $C 
ightarrow D$   $D 
ightarrow d$ 



# Removal of Unit Production Rules

A rule  $A \rightarrow B$  is called **unit production rule** (here  $B \in V$ ).

A rule  $A \rightarrow B$  is called **unit production rule** (here  $B \in V$ ).

#### **Theorem**

For every context-free language L there is a context-free grammar G without  $\lambda$ - and unit-productions with  $L(G) = L \setminus \{\lambda\}$ .

A rule  $A \rightarrow B$  is called **unit production rule** (here  $B \in V$ ).

#### Theorem

For every context-free language L there is a context-free grammar G without  $\lambda$ - and unit-productions with  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let *G* be context-free, without  $\lambda$ -rules, and  $L(G) = L \setminus {\lambda}$ .

A rule  $A \rightarrow B$  is called **unit production rule** (here  $B \in V$ ).

#### **Theorem**

For every context-free language L there is a context-free grammar G without  $\lambda$ - and unit-productions with  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let *G* be context-free, without  $\lambda$ -rules, and  $L(G) = L \setminus {\lambda}$ .

A rule  $A \rightarrow B$  is called **unit production rule** (here  $B \in V$ ).

#### **Theorem**

For every context-free language L there is a context-free grammar G without  $\lambda$ - and unit-productions with  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let *G* be context-free, without  $\lambda$ -rules, and  $L(G) = L \setminus {\lambda}$ .

- Determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ .
- Whenever  $A \Rightarrow^+ B$  and  $B \rightarrow y$  is a rule, add a rule  $A \rightarrow y$ .

A rule  $A \rightarrow B$  is called **unit production rule** (here  $B \in V$ ).

#### **Theorem**

For every context-free language L there is a context-free grammar G without  $\lambda$ - and unit-productions with  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let *G* be context-free, without  $\lambda$ -rules, and  $L(G) = L \setminus {\lambda}$ .

- Determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ .
- Whenever  $A \Rightarrow^+ B$  and  $B \rightarrow y$  is a rule, add a rule  $A \rightarrow y$ .
- Remove all unit production rules.

A rule  $A \rightarrow B$  is called **unit production rule** (here  $B \in V$ ).

#### **Theorem**

For every context-free language L there is a context-free grammar G without  $\lambda$ - and unit-productions with  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let *G* be context-free, without  $\lambda$ -rules, and  $L(G) = L \setminus {\lambda}$ .

- Determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ .
- Whenever  $A \Rightarrow^+ B$  and  $B \rightarrow y$  is a rule, add a rule  $A \rightarrow y$ .
- Remove all unit production rules.

The resulting grammar G has no  $\lambda$ - and unit-productions and it has the property  $L(G) = L \setminus \{\lambda\}$ .

Remove all unit production rules from

$$S o Aa \mid B$$
  $A o a \mid bc \mid B$   $B o A \mid bb$ 

$$B \rightarrow A \mid bb$$

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

$$B \rightarrow A \mid bb$$

Note that there are no  $\lambda$ -productions. (So no need to first remove  $\lambda$ -productions.)

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

$$A
ightarrow a\,|\,bc\,|\,a$$

$$B \rightarrow A \mid bb$$

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

$$B \rightarrow A \mid bb$$

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

$$S \Rightarrow^+ B$$

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

$$A \rightarrow a \mid bc \mid B$$

$$B \rightarrow A \mid bb$$

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

$$S \Rightarrow^+ B \qquad A \Rightarrow^+ B$$

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

$$B \rightarrow A \mid bb$$

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$ 

$$4 \Rightarrow ' E$$

$$B \Rightarrow^+ \prime$$

Remove all unit production rules from

$$S \rightarrow Aa \mid E$$

$$S o Aa \mid B$$
  $A o a \mid bc \mid B$   $B o A \mid bb$ 

$$B \rightarrow A \mid bb$$

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$4 \Rightarrow ' E$$

$$B \Rightarrow ' P$$

$$S \Rightarrow^{\scriptscriptstyle +} A$$

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

$$B \rightarrow A \mid bb$$

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$A \Rightarrow^+ B$$

$$B \Rightarrow^+ A$$

$$S \Rightarrow^+ A$$

$$S \rightarrow Aa \mid B$$

$$A \rightarrow a \mid bc \mid B$$

$$B \rightarrow A \mid bb$$

Remove all unit production rules from

$$S 
ightarrow Aa \mid B$$
  $A 
ightarrow a \mid bc \mid B$   $B 
ightarrow A \mid bb$ 

Note that there are no  $\lambda$ -productions. (So no need to first remove  $\lambda$ -productions.)

(So no need to first remove x-productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$S \rightarrow Aa \mid B \mid a$$
  
 $A \rightarrow a \mid bc \mid B$   
 $B \rightarrow A \mid bb$ 

Remove all unit production rules from

$$S 
ightarrow Aa \mid B$$
  $A 
ightarrow a \mid bc \mid B$   $B 
ightarrow A \mid bb$ 

Note that there are no  $\lambda$ -productions. (So no need to first remove  $\lambda$ -productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$S \rightarrow Aa \mid B \mid a \mid bc$$
  
 $A \rightarrow a \mid bc \mid B$   
 $B \rightarrow A \mid bb$ 

Remove all unit production rules from

$$S 
ightarrow Aa \mid B$$
  $A 
ightarrow a \mid bc \mid B$   $B 
ightarrow A \mid bb$ 

Note that there are no  $\lambda$ -productions. (So no need to first remove  $\lambda$ -productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$S \rightarrow Aa \mid B \mid a \mid bc \mid A$$
  
 $A \rightarrow a \mid bc \mid B$   
 $B \rightarrow A \mid bb$ 

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

$$A
ightarrow a\,|\,bc\,|\,B$$

$$B \rightarrow A \mid bb$$

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$A \Rightarrow^+ B$$

$$B \Rightarrow^+ A$$

$$S \Rightarrow^+ A$$

$$S \rightarrow Aa \mid B \mid a \mid bc \mid A \mid bb$$

$$A \rightarrow a \mid bc \mid B$$

$$B \rightarrow A \mid bb$$

Remove all unit production rules from

$$S 
ightarrow Aa \mid B$$
  $A 
ightarrow a \mid bc \mid B$   $B 
ightarrow A \mid bb$ 

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$S \rightarrow Aa \mid B \mid a \mid bc \mid A \mid bb$$
  
 $A \rightarrow a \mid bc \mid B \mid A$   
 $B \rightarrow A \mid bb$ 

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

Note that there are no  $\lambda$ -productions. (So no need to first remove  $\lambda$ -productions.)

(do no need to mot remove x productions.)

We determine all pairs 
$$A \neq B$$
 with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$S \rightarrow Aa \mid B \mid a \mid bc \mid A \mid bb$$
  
 $A \rightarrow a \mid bc \mid B \mid A \mid bb$   
 $B \rightarrow A \mid bb$ 

Remove all unit production rules from

$$S 
ightarrow Aa \mid B$$
  $A 
ightarrow a \mid bc \mid B$   $B 
ightarrow A \mid bb$ 

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$S \rightarrow Aa \mid B \mid a \mid bc \mid A \mid bb$$
  
 $A \rightarrow a \mid bc \mid B \mid A \mid bb$   
 $B \rightarrow A \mid bb \mid a$ 

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$S \rightarrow Aa \mid B \mid a \mid bc \mid A \mid bb$$
  
 $A \rightarrow a \mid bc \mid B \mid A \mid bb$   
 $B \rightarrow A \mid bb \mid a \mid bc$ 

Remove all unit production rules from

$$S \rightarrow Aa \mid B$$
  $A \rightarrow a \mid bc \mid B$   $B \rightarrow A \mid bb$ 

Note that there are no  $\lambda$ -productions.

(So no need to first remove  $\lambda$ -productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

$$S \rightarrow Aa \mid B \mid a \mid bc \mid A \mid bb$$
  
 $A \rightarrow a \mid bc \mid B \mid A \mid bb$   
 $B \rightarrow A \mid bb \mid a \mid bc \mid B$ 

Remove all unit production rules from

$$S 
ightarrow Aa \mid B$$
  $A 
ightarrow a \mid bc \mid B$   $B 
ightarrow A \mid bb$ 

Note that there are no  $\lambda$ -productions. (So no need to first remove  $\lambda$ -productions.)

We determine all pairs  $A \neq B$  with  $A \Rightarrow^+ B$ :

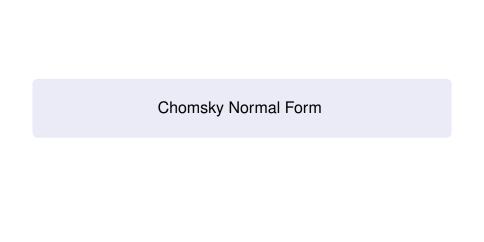
$$S \Rightarrow^+ B$$
  $A \Rightarrow^+ B$   $B \Rightarrow^+ A$   $S \Rightarrow^+ A$ 

Thus we add the following rules:

$$S \rightarrow Aa \mid B \mid a \mid bc \mid A \mid bb$$
  
 $A \rightarrow a \mid bc \mid B \mid A \mid bb$   
 $B \rightarrow A \mid bb \mid a \mid bc \mid B$ 

Removing all unit production rules yields the final result:

$$S 
ightarrow a \mid bb \mid bc \mid Aa$$
  $A 
ightarrow a \mid bb \mid bc$   $B 
ightarrow a \mid bb \mid bc$ 



In a grammar in Chomsky normal form all rules have the form

$$A \rightarrow BC$$
 or  $A \rightarrow a$ 

In a grammar in Chomsky normal form all rules have the form

$$A \rightarrow BC$$
 or  $A \rightarrow a$ 

Note that a grammar in Chomsky normal form contains

- no λ-production rules,
- no unit production rules.

In a grammar in **Chomsky normal form** all rules have the form

$$A \rightarrow BC$$
 or  $A \rightarrow a$ 

Note that a grammar in Chomsky normal form contains

- no λ-production rules,
- no unit production rules.

#### Theorem

For every context-free language L there is a grammar G in Chomsky normal form with  $L(G) = L \setminus \{\lambda\}$ .

#### Construction

Let G be a context-free grammar without  $\lambda$ - and unit-productions and  $L(G) = L \setminus \{\lambda\}$  .

#### Construction

Let G be a context-free grammar without  $\lambda$ - and unit-productions and  $L(G) = L \setminus {\lambda}$ .

■ Introduce variables  $C_a$  and rules  $C_a \rightarrow a$  for every  $a \in T$ .

#### Construction

Let G be a context-free grammar without  $\lambda$ - and unit-productions and  $L(G) = L \setminus \{\lambda\}$ .

- Introduce variables  $C_a$  and rules  $C_a \rightarrow a$  for every  $a \in T$ .
- Replace every rule  $A \rightarrow x_1 \cdots x_n \ (x_i \in V \cup T)$  with  $n \geq 2$  by

$$A \to \sigma(x_1) \cdots \sigma(x_n)$$
 where  $\sigma(x) = \begin{cases} x, & \text{if } x \in V \\ C_x, & \text{if } x \in T \end{cases}$ 

#### Construction

Let G be a context-free grammar without  $\lambda$ - and unit-productions and  $L(G) = L \setminus \{\lambda\}$ .

- Introduce variables  $C_a$  and rules  $C_a \rightarrow a$  for every  $a \in T$ .
- Replace every rule  $A \rightarrow x_1 \cdots x_n \ (x_i \in V \cup T)$  with  $n \geq 2$  by

$$A \to \sigma(x_1) \cdots \sigma(x_n)$$
 where  $\sigma(x) = \begin{cases} x, & \text{if } x \in V \\ C_x, & \text{if } x \in T \end{cases}$ 

■ Replace every  $A \rightarrow B_1 \cdots B_n$  with  $n \ge 3$  by

$$A \rightarrow B_1 \cdots B_{n-2}C$$
  $C \rightarrow B_{n-1}B_n$ 

where *C* is a fresh variable.

#### Construction

Let G be a context-free grammar without  $\lambda$ - and unit-productions and  $L(G) = L \setminus \{\lambda\}$ .

- Introduce variables  $C_a$  and rules  $C_a \rightarrow a$  for every  $a \in T$ .
- Replace every rule  $A \rightarrow x_1 \cdots x_n$  ( $x_i \in V \cup T$ ) with  $n \ge 2$  by

$$A \to \sigma(x_1) \cdots \sigma(x_n)$$
 where  $\sigma(x) = \begin{cases} x, & \text{if } x \in V \\ C_x, & \text{if } x \in T \end{cases}$ 

■ Replace every  $A \rightarrow B_1 \cdots B_n$  with  $n \ge 3$  by

$$A \rightarrow B_1 \cdots B_{n-2}C$$
  $C \rightarrow B_{n-1}B_n$ 

where *C* is a fresh variable.

Repeat the last step until all rules are in Chomsky normal form.

Transform the following context-free grammar

$$S \rightarrow aSbX \mid a$$

$$extbf{X} 
ightarrow extbf{X} a \, | \, aba$$

into Chomsky normal form.

The  $\lambda$ -rules and unit-rules have already been removed.

Transform the following context-free grammar

$$S \rightarrow aSbX \mid a$$
  
 $X \rightarrow Xa \mid aba$ 

into Chomsky normal form.

The  $\lambda$ -rules and unit-rules have already been removed.

$$S \rightarrow aSbX \mid a$$
  
 $X \rightarrow Xa \mid aba$ 

Transform the following context-free grammar

$$S \rightarrow aSbX \mid a$$
  
 $X \rightarrow Xa \mid aba$ 

into Chomsky normal form.

The  $\lambda$ -rules and unit-rules have already been removed.

$$S 
ightarrow aSbX \mid a$$
 $X 
ightarrow Xa \mid aba$ 
 $C_a 
ightarrow a$ 
 $C_b 
ightarrow b$ 

Transform the following context-free grammar

$$S \rightarrow aSbX \mid a$$
  
 $X \rightarrow Xa \mid aba$ 

into Chomsky normal form.

The  $\lambda$ -rules and unit-rules have already been removed.

$$S
ightarrow C_aSC_bX\mid a$$
  $X
ightarrow Xa\mid aba$   $C_a
ightarrow a$   $C_b
ightarrow b$ 

Transform the following context-free grammar

$$S \rightarrow aSbX \mid a$$
  
 $X \rightarrow Xa \mid aba$ 

into Chomsky normal form.

The  $\lambda$ -rules and unit-rules have already been removed.

$$S
ightarrow C_aSC_bX\mid a$$
  $X
ightarrow XC_a\mid aba$   $C_a
ightarrow a$   $C_b
ightarrow b$ 

Transform the following context-free grammar

$$S \rightarrow aSbX \mid a$$
  
 $X \rightarrow Xa \mid aba$ 

into Chomsky normal form.

The  $\lambda$ -rules and unit-rules have already been removed.

$$S
ightarrow C_aSC_bX\mid a \ X
ightarrow XC_a\mid C_aC_bC_a \ C_a
ightarrow a \ C_b
ightarrow b$$

Transform the following context-free grammar

$$S \rightarrow aSbX \mid a$$
  
 $X \rightarrow Xa \mid aba$ 

into Chomsky normal form.

The  $\lambda$ -rules and unit-rules have already been removed.

$$S
ightarrow C_aS_2 \mid a$$
  $S_2
ightarrow SS_3$   $S_3
ightarrow C_bX$   $X
ightarrow XC_a \mid C_aC_bC_a$   $C_a
ightarrow a$   $C_b
ightarrow b$ 

Transform the following context-free grammar

$$S \rightarrow aSbX \mid a$$
  
 $X \rightarrow Xa \mid aba$ 

into Chomsky normal form.

The  $\lambda$ -rules and unit-rules have already been removed.

$$egin{array}{lll} S 
ightarrow C_a S_2 \mid a & S_2 
ightarrow SS_3 & S_3 
ightarrow C_b X \ X 
ightarrow X C_a \mid C_a X_2 & X_2 
ightarrow C_b C_a \ C_b 
ightarrow b & \end{array}$$