Derivačné stromy, redukovaná gramatika

Ing. Viliam Hromada, PhD.

C-510
Ústav informatiky a matematiky
FEI STU

viliam.hromada@stuba.sk

Bezkontextová gramatika č. 1

Je daná gramatika $G = (N, T, P, S)$. $N = \{S, A, B\}$, $T = \{a, b\}$. Pravidlá:

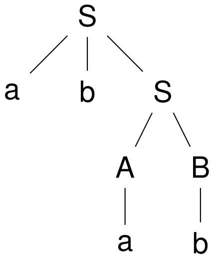
- $S \to abS \mid AB$
- $A \to a \mid aA \mid aBa$
- $B \to b \mid bS$

Pre uvedenú gramatiku splňte nasledovné úlohy:

1. Nájdite derivácie reťazcov *abab*, *aabab* a nakreslite ich derivačné stromy.
2. Zostrojte ľavé a pravé derivácie reťazcov *abab*, *aabab*.
3. Je daná gramatika jednoznačná?

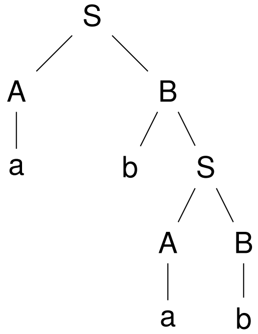
Derivácia *abab* č. 1:

$$S \Rightarrow abS \Rightarrow abAB \Rightarrow abaB \Rightarrow abab$$



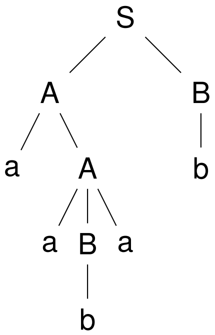
Derivácia *abab* č. 2:

$$S \Rightarrow AB \Rightarrow aB \Rightarrow abS \Rightarrow abAB \Rightarrow abaB \Rightarrow abab$$



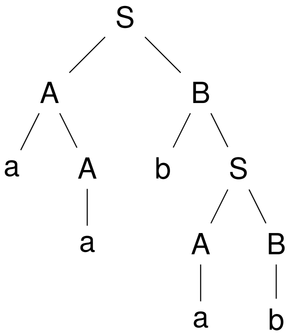
Derivácia *aabab* č. 1:

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow aaBaB \Rightarrow aabaB \Rightarrow aabab$$



Derivácia *aabab* č. 2:

$$S \Rightarrow AB \Rightarrow aAB \Rightarrow aaB \Rightarrow aabS \Rightarrow aabAB \Rightarrow aabaB \Rightarrow aabab$$



Pre obe derivácie *abab* vieme uvažovať ich ľavé (pravé) verzie: Ľavá derivácia *abab* z derivácie č. 1:

$$S \Rightarrow_l abS \Rightarrow_l abAB \Rightarrow_l abaB \Rightarrow_l abab$$

Pravá derivácia *abab* z derivácie č. 1:

$$S \Rightarrow_r abS \Rightarrow_r abAB \Rightarrow_r abAb \Rightarrow_r abab$$

Ľavá derivácia *abab* z derivácie č. 2:

$$S \Rightarrow_l AB \Rightarrow_l aB \Rightarrow_l abS \Rightarrow_l abAB \Rightarrow_l abaB \Rightarrow_l abab$$

Pravá derivácia *abab* z derivácie č. 2:

$$S \Rightarrow_r AB \Rightarrow_r AbS \Rightarrow_r AbAB \Rightarrow_r AbAb \Rightarrow_r Abab \Rightarrow_r abab$$

Podobne pre obe derivácie *aabab* vieme uvažovať ich ľavé (pravé) verzie: Ľavá derivácia *aabab* z derivácie č. 1:

$$S \Rightarrow_l AB \Rightarrow_l aAB \Rightarrow_l aaBaB \Rightarrow_l aabaB \Rightarrow_l aabab$$

Pravá derivácia *aabab* z derivácie č. 1:

$$S \Rightarrow_r AB \Rightarrow_r Ab \Rightarrow_r aAb \Rightarrow_r aaBab \Rightarrow_r aabab$$

Ľavá derivácia *aabab* z derivácie č. 2:

$$S \Rightarrow_l AB \Rightarrow_l aAB \Rightarrow_l aaB \Rightarrow_l aabS \Rightarrow_l aabAB \Rightarrow_l aabaB \Rightarrow_l aabab$$

Pravá derivácia *aabab* z derivácie č. 2:

$$S \Rightarrow_r AB \Rightarrow_r AbS \Rightarrow_r AbAB \Rightarrow_r AbAb \Rightarrow_r Abab \Rightarrow_r aAbab \Rightarrow_r aabab$$

Zároveň sa nám podarilo zodpovedať otázku, či je gramatika jednoznačná:

- Gramatika **nie je jednoznačná**, teda je **nejednoznačná**, pretože
- **Existuje reťazec**, ktorý má aspoň 2 rôzne derivačné stromy - dokonca sme zistili, že existujú minimálne 2 reťazce: *abab* alebo *aabab*, pretože v oboch prípadoch platí, že majú minimálne 2 rôzne derivačné stromy.
- Napr. pre reťazec *abab* vidno, že derivačný strom na slajde č. 3 je **iný** než derivačný strom na slajde č. 4.

Bezkontextová gramatika č. 2

Je daná gramatika $G = (N, T, P, S)$. $N = \{S, A, B\}$, $T = \{a, b\}$. Pravidlá:

- $S \rightarrow AaB \mid BbA$
- $A \rightarrow bAa \mid \varepsilon$
- $B \rightarrow b \mid S$

Dokážte pre uvedenú gramatiku že nie je jednoznačná tým, že nájdete 2 rôzne derivačné stromy pre reťazec *baabb*:

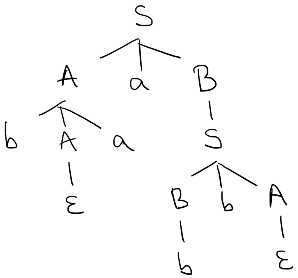
Táto gramatika je **nejednoznačná**, pretože pre reťazec *baabb* existuje derivácia:

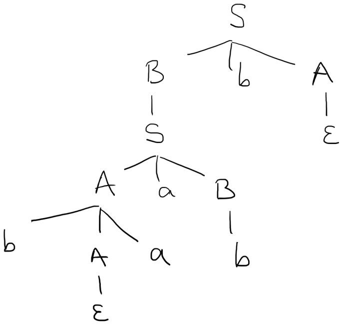
$$S \Rightarrow AaB \Rightarrow bAaaB \Rightarrow baaB \Rightarrow baaS \Rightarrow baaBbA \Rightarrow baabbA \Rightarrow baabb$$

a rovnako existuje iná derivácia:

$$S \Rightarrow BbA \Rightarrow SbA \Rightarrow AaBbA \Rightarrow bAaaBbA \Rightarrow baaBbA \Rightarrow baabbA \Rightarrow baabb$$

pričom **derivačné stromy** oboch derivácií sú rôzne - viď ďalší slajd.





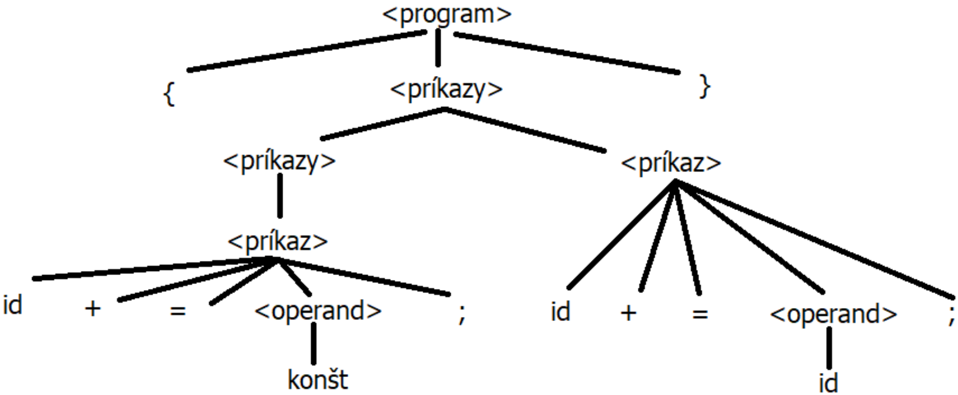
Bezkontextová gramatika č. 3

Je daná bezkontextová gramatika s pravidlami:

1. <program> $\rightarrow$ **{**<príkazy>**}**
2. <príkazy> $\rightarrow$ <príkaz>
3. <príkazy> $\rightarrow$ <príkazy><príkaz>
4. <príkaz> $\rightarrow$ **id+=**<operand>**;**
5. <operand> $\rightarrow$ **id**
6. <operand> $\rightarrow$ **konšt**

kde
$N = \{$<program>, <príkazy>, <príkaz>, <operand>$\}$, $T = \{$**id**, **konšt**, **{**, **}**, **;**, **+**, **=**$\}$ a počiatočný neterminál je <program>. Nájdite derivačný strom reťazca
{ id += konšt ; id += id ;}

Derivačný strom:



Gramatika č. 1 - Úpravy - redukcia gramatiky

Preveďte uvedenú gramatiku na ekvivalentnú redukovanú gramatiku. Daná gramatika $G = (\{S, A, B, C\}, \{a, b, c\}, P, S)$, pravidlá P:

- $S \to aAC \mid bB \mid BB$
- $A \to aAb$
- $B \to \varepsilon \mid aBb \mid SB$
- $C \to CC \mid b \mid c$

Odstránenie nadbytočných neterminálov

V prvom kroku hľadáme neterminály, ktoré **nie sú nadbytočné**, teda také, ktoré patria do množiny N_T neterminálov, z ktorých je možné teoreticky derivovať nejaký reťazec terminálov:

1. $N_T = \emptyset$ (na začiatku)
2. Keďže $B \to \varepsilon$, tak určite $N_T = \{B\}$
3. Keďže $C \to b$, tak určite $N_T = \{B, C\}$
4. Keď už vieme, že $B \in N_T$ a zároveň v pravidlách $S \to BB$, tak potom $N_T = \{B, C, S\}$.
5. Vidíme, že $A \notin N_T$, teda kompletná množina $N_T = \{S, B, C\}$. Neterminál A teda môžeme z gramatiky odstrániť, pretože je nadbytočný

Dostávame gramatiku:

- $S \rightarrow bB \mid BB$
- $B \rightarrow \varepsilon \mid aBb \mid SB$
- $C \rightarrow CC \mid b \mid c$

Teraz nájdeme množinu V_D.

Odstránenie nedostupných symbolov

V druhom kroku hľadáme terminály a neterminály, ktoré **sú dostupné**, teda také, ktoré patria do množiny V_D symbolov, ktoré sú dosiahnuteľné z počiatočného neterminálu S počas nejakej derivácie:

1. Na začiatku $V_D = \{S\}$, pretože počiatočný neterminál je vždy dostupný symbol.
2. Všetky symboly v pravidlách, kde je na ľavej strane S, t.j. $S \to bB, S \to BB$ sú taktiež dosiahnuteľné, t.j. $V_D = \{S, b, B\}$
3. Keďže už vieme, že aj B je dosiahnuteľný symbol, aj symboly v pravidlách, kde je na ľavej strane B sú dosiahnuteľné, t.j. $V_D = \{S, b, B, a\}$ (len pre upozornenie, ε nie je symbol gramatiky, preto ho nedávame do V_D).
4. Keďže do množiny V_D nám už nepribudne ďalší neterminál, výsledok je teda $V_D = \{S, B, a, b\}$.
5. Symboly C, c teda **nie sú dosiahnuteľné** a môžeme ich zmazať z gramatiky!

Výsledná redukovaná gramatika:

- $S \to bB \mid BB$
- $B \to \varepsilon \mid aBb \mid SB$

POZOR!!!

Ak by sme v danej gramatike:

- $S \rightarrow aAC \mid bB \mid BB$
- $A \rightarrow aAb$
- $B \rightarrow \varepsilon \mid aBb \mid SB$
- $C \rightarrow CC \mid b \mid c$

najprv odstránili nedostupné symboly (množina V_D) a potom nadbytočné neterminály (množina N_T), t.j. odstránenie symbolov vykonáme v **opačnom poradí**, **NEDOSTANEME** redukovanú gramatiku!!!

Pre túto gramatiku totiž $V_D = \{S, a, A, C, b, B, c\}$, t.j. všetky symboly sú dosiahnuteľné, čiže by sme neodstránili nič a následne $N_T = \{S, B, C\}$, čiže by sme dostali gramatiku

- $S \rightarrow bB \mid BB$
- $B \rightarrow \varepsilon \mid aBb \mid SB$
- $C \rightarrow CC \mid b \mid c$

ktorá **nie je redukovaná.** Preto na poradí odstraňovania symbolov **záleží!**

Gramatika č. 1 - Úpravy - odstránenie ε-pravidiel

Z gramatiky $G = (\{S, B\}, \{a, b\}, P, S)$ s pravidlami:

- $S \rightarrow bB \mid BB$
- $B \rightarrow \varepsilon \mid aBb \mid SB$

odstráňte ε-pravidlá.

Množina N_ε

Na odstránenie ε-pravidiel najprv vypočítame množinu N_ε:

1. Na začiatku $N_\varepsilon = \emptyset$
2. Keďže v pravidlách je pravidlo $B \rightarrow \varepsilon$, vďaka tomu vidíme, že $B \Rightarrow \varepsilon$ a teda $B \in N_\varepsilon$.
3. Keďže $B \in N_\varepsilon$ a v pravidlách $S \rightarrow BB$, tak vidíme, že $S \Rightarrow BB \Rightarrow B \Rightarrow \varepsilon$, teda $S \Rightarrow^* \varepsilon$ a teda $S \in N_\varepsilon$.
4. Teda: $N_\varepsilon = \{S, B\}$

Keď máme určenú množinu N_ε, môžeme upravovať pravidlá v gramatike. Najprv odstránime všetky ε-pravidlá:

- $S \rightarrow bB \mid BB$
- $B \rightarrow aBb \mid SB$

a podľa množiny N_ε pridáme nasledovné pravidlá:

- Pravidlo: $S \to bB$ má na pravej strane neterminál B pre ktorý platí $B \in N_\varepsilon$.
- Preto do gramatiky okrem pravidla $S \to bB$ by sme pridali aj jeho verziu, kde neterminál B nevystupuje na pravej strane, t.j. pravidlo $S \to b$.

Teda v gramatike:

- $S \to bB \mid BB \mid b$
- $B \to aBb \mid SB$

- Pravidlo: $S \rightarrow BB$ má na pravej strane neterminál B pre ktorý platí $B \in N_\varepsilon$.
- Preto do gramatiky okrem pravidla $S \rightarrow BB$ pridáme **všetky verzie**, kde B na pravej strane vystupuje alebo nevystupuje! To znamená, **teoreticky prichádzajú do úvahy pravidlá** $S \rightarrow BB \mid B \mid \varepsilon$**.** Pridáme však len také, ktoré ešte **nemáme** alebo ktoré **nie sú ε-pravidlá**, t.j. $S \rightarrow B$

Teda v gramatike:

- $S \rightarrow bB \mid BB \mid b \mid B$
- $B \rightarrow aBb \mid SB$

- Pravidlo: $B \to aBb$ má na pravej strane neterminál B pre ktorý platí $B \in N_\varepsilon$.
- Preto do gramatiky okrem pravidla $B \to aBb$ pridáme **všetky verzie**, kde B na pravej strane vystupuje alebo nevystupuje! To znamená, pridáme $B \to ab$

Teda v gramatike:

- $S \to bB \mid BB \mid b \mid B$
- $B \to aBb \mid SB \mid ab$

- Pravidlo: $B \to SB$ má na pravej strane aj neterminál S pre ktorý platí $S \in N_\varepsilon$, aj neterminál B pre ktorý platí $B \in N_\varepsilon$
- Preto do gramatiky okrem pravidla $B \to SB$ pridáme **všetky verzie**, kde alebo S, alebo B na pravej strane vystupujú alebo nevystupujú! To znamená, **teoreticky prichádzajú do úvahy pravidlá** $B \to SB \mid B \mid S \mid \varepsilon$. Pridáme však len také, ktoré ešte **nemáme**, alebo ktoré **nie sú ε-pravidlá**, alebo ktoré nemajú rovnakú ľavú a pravú stranu! Teda pridáme $B \to S$

Teda v gramatike:

- $S \to bB \mid BB \mid b \mid B$
- $B \to aBb \mid SB \mid ab \mid S$

Odstránenie ε-pravidiel

Tým sme vyčerpali všetky pravidlá. Dostali sme gramatiku:

- $S \to bB \mid BB \mid b \mid B$
- $B \to aBb \mid SB \mid ab \mid S$

Na záver, keďže pre počiatočný neterminál S platí, že $S \in N_\varepsilon$, **musíme** do gramatiky doplniť nový počiatočný neterminál $\acute{S}$ a pravidlá $\acute{S} \to \varepsilon \mid S$, teda výsledná gramatika:

- $\acute{S} \to \varepsilon \mid S$
- $S \to bB \mid BB \mid b \mid B$
- $B \to aBb \mid SB \mid ab \mid S$

V tejto gramatike je počiatočný neterminál $\acute{S}$ a neterminály sú $N = \{\acute{S}, S, B\}$.

Gramatika č. 1 - Úpravy - odstránenie jednoduchých pravidiel

Z gramatiky $G = (\{\acute{S}, S, B\}, \{a, b\}, P, S)$ s pravidlami:

- $\acute{S} \rightarrow \varepsilon \mid S$
- $S \rightarrow bB \mid BB \mid b \mid B$
- $B \rightarrow aBb \mid SB \mid ab \mid S$

odstráňte jednoduché pravidlá.

Odstránenie jednoduchých pravidiel

Odstránenie jednoduchých pravidiel začíname tým, že pre neterminály $\acute{S}, S, B$ zostrojíme množiny $N_{\acute{S}}, N_S, N_B$. Začnime s množinou $N_{\acute{S}}$:

- Množina $N_{\acute{S}}$ je množina neterminálov, ktoré je možné odvodiť z neterminálu $\acute{S}$, t.j. $N_{\acute{S}} = \{X \in N \mid \acute{S} \Rightarrow^* X\}$
- Implicitne do tejto množiny určite patrí $\acute{S}$, t.j. $N_{\acute{S}} = \{\acute{S}, ...\}$
- Keďže v gramatike je pravidlo $\acute{S} \rightarrow S$, určite do množiny $N_{\acute{S}}$ patrí aj S, t.j. $N_{\acute{S}} = \{\acute{S}, S, ...\}$, pretože vďaka tomuto pravidlu vieme vyrobiť odvodenie: $\acute{S} \Rightarrow S$
- Keď teraz vieme, že $S \in N_{\acute{S}}$ a vidíme, že v gramatike je pravidlo $S \rightarrow B$, tak potom aj $B \in N_{\acute{S}}$, pretože vďaka tomuto pravidlu máme deriváciu: $\acute{S} \Rightarrow S \Rightarrow B$.
- Teda sme zistili, že $N_{\acute{S}} = \{\acute{S}, S, B\}$

Zostrojenie množiny N_S:

- Množina N_S je množina neterminálov, ktoré je možné odvodiť z neterminálu S, t.j. $N_S = \{X \in N \mid S \Rightarrow^* X\}$
- Implicitne do tejto množiny určite patrí S, t.j. $N_S = \{S, ...\}$
- Keďže v gramatike je pravidlo $S \rightarrow B$, určite do množiny N_S patrí aj B, t.j. $N_S = \{S, B...\}$, pretože vďaka tomuto pravidlu vieme vyrobiť odvodenie: $S \Rightarrow B$
- Hoci nám pribudol nový neterminál B a v gramatike máme pravidlo $B \rightarrow S$, vedeli by sme spraviť deriváciu $S \Rightarrow B \Rightarrow S$. Avšak S nepridáme do množiny N_S, pretože tam už je.
- Pomocou iných pravidiel už nevieme odvodiť iný neterminál z neterminálu S, teda výsledok: $N_S = \{S, B\}$

Zostrojenie množiny N_B:

- Množina N_B je množina neterminálov, ktoré je možné odvodiť z neterminálu B, t.j. $N_B = \{X \in N \mid B \Rightarrow^* X\}$
- Implicitne do tejto množiny určite patrí B, t.j. $N_B = \{B, ...\}$
- Podľa pravidla $B \rightarrow S$ vieme spraviť deriváciu $B \Rightarrow S$, teda $S \in N_B$.
- Ďalšie neterminály do N_B pridať nevieme.
- Výsledok $N_B = \{B, S\}$.

Odstránenie jednoduchých pravidiel

Keď poznáme množiny $N_{\acute{S}}$, N_S a N_B, pristúpime k úprave gramatiky:

- $\acute{S} \rightarrow \varepsilon \mid S$
- $S \rightarrow bB \mid BB \mid b \mid B$
- $B \rightarrow aBb \mid SB \mid ab \mid S$

Najprv z gramatiky môžeme odstrániť všetky jednoduché pravidlá, čím dostaneme gramatiku:

- $\acute{S} \rightarrow \varepsilon$
- $S \rightarrow bB \mid BB \mid b$
- $B \rightarrow aBb \mid SB \mid ab$

Odstránenie jednoduchých pravidiel

Následne do gramatiky doplníme nové pravidlá tak, že do neterminálu X pridáme pravé strany všetkých pravidiel všetkých neterminálov z množiny N_X. Teda napríklad, ak $N_{\acute{S}} = \{\acute{S}, S, B\}$, tak k neterminálu $\acute{S}$ pridáme **pravé strany** neterminálov S a B:

- $\acute{S} \to \varepsilon \mid bB \mid BB \mid b \mid aBb \mid SB \mid ab$
- $S \to bB \mid BB \mid b$
- $B \to aBb \mid SB \mid ab$

Následne urobíme to isté pre neterminál S podľa množiny $N_S = \{S, B\}$, t.j. pridáme k nemu pravé strany pravidiel pre neterminál B:

- $\acute{S} \to \varepsilon \mid bB \mid BB \mid b \mid aBb \mid SB \mid ab$
- $S \to bB \mid BB \mid b \mid aBb \mid SB \mid ab$
- $B \to aBb \mid SB \mid ab$

Odstránenie jednoduchých pravidiel

Následne urobíme to isté pre neterminál B podľa množiny $N_B = \{B, S\}$, t.j. pridáme k nemu pravé strany pravidiel pre neterminál S (tie, ktoré tam ešte nemáme):

- $\acute{S} \to \varepsilon \mid bB \mid BB \mid b \mid aBb \mid SB \mid ab$
- $S \to bB \mid BB \mid b \mid aBb \mid SB \mid ab$
- $B \to aBb \mid SB \mid ab \mid bB \mid BB \mid b$

Teda k zadanej gramatike:

- $\acute{S} \rightarrow \varepsilon \mid S$
- $S \rightarrow bB \mid BB \mid b \mid B$
- $B \rightarrow aBb \mid SB \mid ab \mid S$

sme zostrojili ekvivalentnú gramatiku **bez jednoduchých pravidiel**:

- $\acute{S} \rightarrow \varepsilon \mid bB \mid BB \mid b \mid aBb \mid SB \mid ab$
- $S \rightarrow bB \mid BB \mid b \mid aBb \mid SB \mid ab$
- $B \rightarrow aBb \mid SB \mid ab \mid bB \mid BB \mid b$

Gramatika č. 2 - Úpravy - Redukcia gramatiky

Je daná bezkontextová gramatika $G = (\{S, A, B, C\}, \{a, b, c\}, P, S)$ s pravidlami:

- $S \to AbBb$
- $A \to cB \mid \varepsilon \mid AB$
- $B \to SA \mid \varepsilon \mid BA$
- $C \to aCb \mid cB$

Upravte gramatiku do redukovaného tvaru.

Množina N_T

Množina N_T:

- Keďže $A \to \varepsilon$, tak $N_T = \{A\}$.
- Keďže $B \to \varepsilon$, tak $N_T = \{A, B\}$.
- Keďže $B \in N_T$ a $C \to cB$, tak $N_T = \{A, B, C\}$.
- Keďže $A, B \in N_T$ a $S \to AbBb$, tak $N_T = \{A, B, C, S\}$.

Keďže každý neterminál gramatiky patrí do N_T, tak v tomto kroku žiaden neodstránime.

Množina V_D

Množina V_D:

- Počiatočný neterminál je dosiahnuteľný vždy, $V_D = \{S\}$.
- Z pravidla $S \to AbBb$ máme $V_D = \{S, A, B, b\}$.
- Keďže aj A je dosiahnuteľný, tak z pravidla $A \to cB$ máme $V_D = \{S, A, B, b, c\}$.
- Z ostatných pravidiel by sme dostali ako dosiahnuteľné symboly len také, ktoré už v množine V_D máme.
- Ďalšie dosiahnuteľné symboly nedostaneme - všimnite si, že neexistuje spôsob, ako sa z dosiahnuteľných neterminálov $\{S, A, B\}$ dostať k neterminálu C alebo terminálu a.

Keďže symboly $a, C \notin V_D$, t.j. sú nedosiahnuteľné, odstránime ich z gramatiky.

Gramatika č. 2 - Úpravy - Redukcia gramatiky

Výsledkom je redukovaná gramatika $G = (\{S, A, B\}, \{b, c\}, P, S)$ s pravidlami:

- $S \to AbBb$
- $A \to cB \mid \varepsilon \mid AB$
- $B \to SA \mid \varepsilon \mid BA$

V ďalšom kroku odstráňte ε-pravidlá.

Gramatika č. 2 - Úpravy - odstránenie ε-pravidiel

Množina N_ε:

- Z pravidla $A \to \varepsilon$ máme $N_\varepsilon = \{A\}$.
- Z pravidla $B \to \varepsilon$ máme $N_\varepsilon = \{A, B\}$.
- Keďže pre neterminál S **nemáme** pravidlo, ktoré by malo na pravej strane len kombináciu neterminálov A, B, o ktorých už vieme, že patria do N_ε, tak S do N_ε patriť nebude.

Výsledná množina $N_\varepsilon = \{A, B\}$.

Gramatika č. 2 - Úpravy - odstránenie ε-pravidiel

Z gramatiky najprv odstránime ε-pravidlá, t.j. dostávame:

- $S \rightarrow AbBb$
- $A \rightarrow cB \mid AB$
- $B \rightarrow SA \mid BA$

Následne do gramatiky pridáme nové pravidlá na základe tých pravidiel, ktoré majú na pravej strane neterminály z množiny N_ε

Gramatika č. 2 - Úpravy - odstránenie ε-pravidiel

- Pravidlo $S \rightarrow AbBb$ má na pravej strane neterminály A, B, oba sú v N_ε. Preto do gramatiky pridáme **všetky verzie** pravidla, kde A alebo B vystupuje alebo nevystupuje
- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB$
- $B \rightarrow SA \mid BA$

Gramatika č. 2 - Úpravy - odstránenie ε-pravidiel

- Pravidlo $A \rightarrow cB$ má na pravej strane neterminál B, preto pridáme do gramatiky aj verziu pravidla bez neterminálu B, t.j. $A \rightarrow c$:
- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB \mid c$
- $B \rightarrow SA \mid BA$

Gramatika č. 2 - Úpravy - odstránenie ε-pravidiel

- Pravidlo $A \rightarrow AB$ má na pravej strane neterminály A, B, oba sú v N_ε. Preto do gramatiky pridáme **všetky verzie** pravidla, kde A alebo B vystupuje alebo nevystupuje, avšak **nepridáme** pravidlo $A \rightarrow A$, pretože je zbytočné, nepridáme $A \rightarrow AB$, pretože tam už je a nepridáme $A \rightarrow \varepsilon$, pretože pravidlá tohto typu odstraňujeme. Pridáme teda len $A \rightarrow B$.
- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB \mid c \mid B$
- $B \rightarrow SA \mid BA$

Gramatika č. 2 - Úpravy - odstránenie ε-pravidiel

- Pre pravidlo $B \rightarrow SA$, kde na pravej strane je neterminál A, ktorý je z množiny N_ε, pridáme pravidlo $B \rightarrow S$:
- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB \mid c \mid B$
- $B \rightarrow SA \mid BA \mid S$

Gramatika č. 2 - Úpravy - odstránenie ε-pravidiel

- Pravidlo $B \rightarrow BA$ má na pravej strane neterminály A, B, oba sú v N_ε. Preto do gramatiky pridáme **všetky verzie** pravidla, kde A alebo B vystupuje alebo nevystupuje, avšak **nepridáme** pravidlo $B \rightarrow B$, pretože je zbytočné, nepridáme $B \rightarrow BA$, pretože tam už je a nepridáme $B \rightarrow \varepsilon$, pretože pravidlá tohto typu odstraňujeme. Pridáme teda len $B \rightarrow A$.
- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB \mid c \mid B$
- $B \rightarrow SA \mid BA \mid S \mid A$

Gramatika č. 2 - Úpravy - odstránenie ε-pravidiel

Keďže v tomto prípade $S \notin N_\varepsilon$, nemusíme pridávať nový počiatočný neterminál a teda výsledná gramatika bez ε-pravidiel je:

- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB \mid c \mid B$
- $B \rightarrow SA \mid BA \mid S \mid A$

Gramatika č. 2 - Úpravy - odstránenie jednoduchých pravidiel

Z bezkontextovej gramatiky $G = (\{S, A, B\}, \{b, c\}, P, S)$ s pravidlami:

- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB \mid c \mid B$
- $B \rightarrow SA \mid BA \mid S \mid A$

odstráňte jednoduché pravidlá.

Odstránenie jednoduchých pravidiel

Odstránenie jednoduchých pravidiel začíname tým, že pre neterminály S, A, B zostrojíme množiny N_S, N_A, N_B. Začnime s množinou N_S:

- Množina N_S je množina neterminálov, ktoré je možné odvodiť z neterminálu S, t.j. $N_S = \{X \in N \mid S \Rightarrow^* X\}$
- Implicitne do tejto množiny určite patrí S, t.j. $N_S = \{S...\}$
- Keďže pre neterminál S nemáme pravidlo, kde by na pravej strane bol v tomto momente len 1 neterminál, nič viac do N_S nepridáme, pretože z S nevieme odvodiť len 1 iný neterminál.
- Teda sme zistili, že $N_S = \{S\}$

Zostrojenie množiny N_A:

- Množina N_A je množina neterminálov, ktoré je možné odvodiť z neterminálu A, t.j. $N_A = \{X \in N \mid A \Rightarrow^* X\}$
- Implicitne do tejto množiny určite patrí A, t.j. $N_A = \{A, ...\}$
- Podľa pravidla $A \to B$ vieme spraviť deriváciu $A \Rightarrow B$, teda $B \in N_A$.
- Keď už vieme, že $B \in N_A$ a máme pravidlo $B \to S$, tak potom platí $A \Rightarrow B \Rightarrow S$, teda $A \Rightarrow^* S$ a teda $S \in N_A$.
- Ďalšie neterminály do N_A pridať nevieme.
- Výsledok $N_A = \{A, B, S\}$.

Podobne dostaneme $N_B = \{B, A, S\}$.

Odstránenie jednoduchých pravidiel

Keď poznáme množiny N_S, N_A a N_B, pristúpime k úprave gramatiky:

- $S \to AbBb \mid bBb \mid Abb \mid bb$
- $A \to cB \mid AB \mid c \mid B$
- $B \to SA \mid BA \mid S \mid A$

Najprv z gramatiky môžeme odstrániť všetky jednoduché pravidlá, čím dostaneme gramatiku:

- $S \to AbBb \mid bBb \mid Abb \mid bb$
- $A \to cB \mid AB \mid c$
- $B \to SA \mid BA$

Odstránenie jednoduchých pravidiel

Následne do gramatiky doplníme nové pravidlá tak, že do neterminálu X pridáme pravé strany všetkých pravidiel všetkých neterminálov z množiny N_X. Pre $N_S = \{S\}$ však do pravidiel nedoplníme nič, pretože v množine N_S nie je iný neterminál než S.

- $S \to AbBb \mid bBb \mid Abb \mid bb$
- $A \to cB \mid AB \mid c$
- $B \to SA \mid BA$

Pre $N_A = \{A, B, S\}$ doplníme do pravidiel pre A všetky pravé strany neterminálov S a B:

- $S \to AbBb \mid bBb \mid Abb \mid bb$
- $A \to cB \mid AB \mid c \mid AbBb \mid bBb \mid Abb \mid bb \mid SA \mid BA$
- $B \to SA \mid BA$

Odstránenie jednoduchých pravidiel

Následne urobíme to isté pre neterminál B podľa množiny $N_B = \{B, A, S\}$, t.j. pridáme k nemu pravé strany pravidiel pre neterminály A a S (tie, ktoré tam ešte nemáme):

- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB \mid c \mid AbBb \mid bBb \mid Abb \mid bb \mid SA \mid BA$
- $B \rightarrow SA \mid BA \mid cB \mid AB \mid c \mid AbBb \mid bBb \mid Abb \mid bb$

Teda k zadanej gramatike:

- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB \mid c \mid B$
- $B \rightarrow SA \mid BA \mid S \mid A$

sme zostrojili ekvivalentnú gramatiku **bez jednoduchých pravidiel**:

- $S \rightarrow AbBb \mid bBb \mid Abb \mid bb$
- $A \rightarrow cB \mid AB \mid c \mid AbBb \mid bBb \mid Abb \mid bb \mid SA \mid BA$
- $B \rightarrow SA \mid BA \mid cB \mid AB \mid c \mid AbBb \mid bBb \mid Abb \mid bb$

Gramatika č. 3 - Úprava - Redukcia

Je daná gramatika $G = (\{S, A, B, C\}, \{a, b\}, P, S)$. Upravte gramatiku na ekvivalentnú redukovanú gramatiku:

- $S \rightarrow ABCA$
- $A \rightarrow AB \mid \varepsilon$
- $B \rightarrow AC \mid a$
- $C \rightarrow BA \mid b \mid \varepsilon$

Množina N_T:

- Z pravidla $A \to \varepsilon$: $N_T = \{A\}$.
- Z pravidla $B \to a$: $N_T = \{A, B\}$.
- Z pravidla $C \to b$ alebo $C \to \varepsilon$: $N_T = \{A, B, C\}$.
- Keďže $A, B, C \in N_T$ a $S \to ABCA$, tak aj $N_T = \{A, B, C, S\}$.

Zatiaľ teda neodstránime žiaden neterminál.

Množina V_D:

- Počiatočný neterminál: $V_D = \{S\}$
- Keďže $S \in V_D$, tak z pravidla $S \to ABCA$ aj $A, B, C \in V_D$, $V_D = \{S, A, B, C\}$
- A teda aj všetky symboly na ich pravých stranách, teda $V_D = \{S, A, B, C, a, b\}$

Teda zadaná gramatika **už je redukovaná.**

Gramatika č. 3 - Úprava - Odstránenie ε-pravidiel

Upravte gramatiku na ekvivalentnú gramatiku bez ε-pravidiel:

- $S \rightarrow ABCA$
- $A \rightarrow AB \mid \varepsilon$
- $B \rightarrow AC \mid a$
- $C \rightarrow BA \mid b \mid \varepsilon$

Množina N_ε

1. Na začiatku $N_\varepsilon = \emptyset$
2. V pravidlách $A \rightarrow \varepsilon$, teda určite $A \in N_\varepsilon$
3. V pravidlách $C \rightarrow \varepsilon$, teda určite $C \in N_\varepsilon$
4. Keďže určite $A \in N_\varepsilon$ a $C \in N_\varepsilon$ a zároveň v pravidlách $B \rightarrow AC$, tak určite aj $B \in N_\varepsilon$.
5. Keďže určite $A \in N_\varepsilon$, $B \in N_\varepsilon$ a $C \in N_\varepsilon$ a zároveň v pravidlách $S \rightarrow ABCA$, tak určite aj $S \in N_\varepsilon$.
6. Čiže **z každého neterminálu** gramatiky je možné odvodiť ε, teda $N_\varepsilon = \{S, A, B, C\}$

Odstráňme z gramatiky ε-pravidlá:

- $S \to ABCA$
- $A \to AB$
- $B \to AC \mid a$
- $C \to BA \mid b$

a poďme pridávať pravidlá podľa množiny $N_\varepsilon = \{S, A, B, C\}$

- Pravidlo: $S \rightarrow ABCA$ má na pravej strane neterminály A, B, C, pričom všetky patria aj do množiny N_ε
- Preto do gramatiky okrem pravidla $S \rightarrow ABCA$ pridáme **všetky verzie**, kde A, B, C vystupujú alebo nevystupujú. Keďže na pravej strane sú 4 takéto neterminály (dvakrát A, raz B a C), teoreticky prichádza do úvahy $2^4 = 16$ verzií pravej strany

$$S \rightarrow ABCA \mid \cancel{A}BCA \mid A\cancel{B}CA \mid AB\cancel{C}A \mid ABC\cancel{A}$$

$$S \rightarrow \cancel{A}\cancel{B}CA \mid \cancel{A}B\cancel{C}A \mid \cancel{A}BC\cancel{A} \mid A\cancel{B}\cancel{C}A \mid A\cancel{B}C\cancel{A} \mid AB\cancel{C}\cancel{A}$$

$$S \rightarrow \cancel{A}\cancel{B}\cancel{C}A \mid \cancel{A}\cancel{B}C\cancel{A} \mid \cancel{A}B\cancel{C}\cancel{A} \mid A\cancel{B}\cancel{C}\cancel{A} \mid \cancel{A}\cancel{B}\cancel{C}\cancel{A}$$

Do pravidiel teda na základe pravidla $S \to ABCA$ pridáme:

- $S \to ABCA \mid$
  $BCA \mid ACA \mid ABA \mid ABC \mid CA \mid BA \mid BC \mid AA \mid AC \mid AB \mid A \mid B \mid C$
- $A \to AB$
- $B \to AC \mid a$
- $C \to BA \mid b$

Do pravidiel na základe pravidla $A \to AB$ pridáme:

- $S \to ABCA \mid BCA \mid ACA \mid ABA \mid ABC \mid CA \mid BA \mid BC \mid AA \mid AC \mid AB \mid A \mid B \mid C$
- $A \to AB \mid B$
- $B \to AC \mid a$
- $C \to BA \mid b$

Do pravidiel na základe pravidla $B \to AC$ pridáme:

- $S \to ABCA \mid BCA \mid ACA \mid ABA \mid ABC \mid CA \mid BA \mid BC \mid AA \mid AC \mid AB \mid A \mid B \mid C$
- $A \to AB \mid B$
- $B \to AC \mid a \mid A \mid C$
- $C \to BA \mid b$

Do pravidiel na základe pravidla $C \rightarrow BA$ pridáme:

- $S \rightarrow ABCA \mid BCA \mid ACA \mid ABA \mid ABC \mid CA \mid BA \mid BC \mid AA \mid AC \mid AB \mid A \mid B \mid C$
- $A \rightarrow AB \mid B$
- $B \rightarrow AC \mid a \mid A \mid C$
- $C \rightarrow BA \mid b \mid B \mid A$

A keďže v gramatike platilo pre počiatočný neterminál $S \in N_\varepsilon$, tak do gramatiky pridáme nový počiatočný neterminál $\acute{S}$ a pravidlá $\acute{S} \to \varepsilon \mid S$, t.j. dostaneme výslednú gramatiku:

- $\acute{S} \to \varepsilon \mid S$
- $S \to ABCA \mid BCA \mid ACA \mid ABA \mid ABC \mid CA \mid BA \mid BC \mid AA \mid AC \mid AB \mid A \mid B \mid C$
- $A \to AB \mid B$
- $B \to AC \mid a \mid A \mid C$
- $C \to BA \mid b \mid B \mid A$

Gramatika č. 3 - Úprava - Odstránenie jednoduchých pravidiel

Z gramatiky $G = (\{\acute{S}, S, A, B, C\}, \{a, b\}, P, S)$ s pravidlami

- $\acute{S} \rightarrow \varepsilon \mid S$
- $S \rightarrow ABCA \mid BCA \mid ACA \mid ABA \mid ABC \mid CA \mid BA \mid BC \mid AA \mid AC \mid AB \mid A \mid B \mid C$
- $A \rightarrow AB \mid B$
- $B \rightarrow AC \mid a \mid A \mid C$
- $C \rightarrow BA \mid b \mid B \mid A$

odstráňte jednoduché pravidlá.

Množiny potrebné pre odstránenie jednoduchých pravidiel (už len správne výsledky bez postupu):

- $N_{\acute{S}} = \{\acute{S}, S, A, B, C\}$
- $N_S = \{S, A, B, C\}$
- $N_A = \{A, B, C\}$
- $N_B = \{A, B, C\}$
- $N_C = \{A, B, C\}$

Gramatika č. 3 - Úprava - Odstránenie jednoduchých pravidiel

Výsledná gramatika po odstránení jednoduchých pravidiel:

- $\acute{S} \to \varepsilon \mid ABCA \mid BCA \mid ACA \mid ABA \mid ABC \mid CA \mid BA \mid BC \mid AA \mid AC \mid AB \mid a \mid b$
- $S \to ABCA \mid BCA \mid ACA \mid ABA \mid ABC \mid CA \mid BA \mid BC \mid AA \mid AC \mid AB \mid a \mid b$
- $A \to AB \mid AC \mid a \mid BA \mid b$
- $B \to AC \mid a \mid BA \mid b \mid AB$
- $C \to BA \mid b \mid AB \mid AC \mid a$

Všimnite si, že počas tejto úpravy došlo k tomu, že vo výslednej gramatike sa vyskytuje **nedostupný/nadbytočný symbol**, neterminál S. To sa počas takýchto úprav môže stať.

Gramatika na úpravu č. 4

Najprv upravte uvedenú gramatiku na redukovanú gramatiku. Následne transformujte redukovanú gramatiku na gramatiku bez ε-pravidiel. Následne gramatiku bez ε-pravidiel upravte na gramatiku bez jednoduchých pravidiel.

- $S \rightarrow Aa \mid Bb$
- $A \rightarrow C \mid DaD$
- $B \rightarrow D \mid C$
- $C \rightarrow S \mid a \mid \varepsilon$
- $D \rightarrow S \mid b$
- $E \rightarrow S \mid c \mid \varepsilon$

Množina N_T

- $N_T = \emptyset$
- Keďže v pravidlách $C \to a$, tak $C \in N_T$
- Keďže v pravidlách $D \to b$, tak $D \in N_T$
- Keďže v pravidlách $E \to c$, tak $E \in N_T$
- Keďže v pravidlách $A \to C$ a $C \in N_T$, tak aj $A \in N_T$
- Keďže v pravidlách $B \to D$ a $D \in N_T$, tak aj $B \in N_T$
- Keďže v pravidlách $S \to Aa$ a $A \in N_T$, tak aj $S \in N_T$
- Teda $N_T = \{S, A, B, C, D, E\}$ a žiaden neterminál v tomto kroku neodstránime.

Množina V_D

- $V_D = \{S\}$
- Do množiny V_D pribudnú všetky symboly dosiahnuteľné z neterminálu S, t.j. z pravidiel, kde je S na ľavej strane, $V_D = \{S, A, a, B, b\}$.
- Do množiny V_D pribudnú všetky symboly dosiahnuteľné z neterminálu A, t.j. z pravidiel, kde je A na ľavej strane, $V_D = \{S, A, a, B, b, C, D\}$.
- Do množiny V_D pribudnú všetky symboly dosiahnuteľné z neterminálu B, t.j. z pravidiel, kde je A na ľavej strane, $V_D = \{S, A, a, B, b, C, D\}$.
- Do množiny V_D pribudnú všetky symboly dosiahnuteľné z neterminálu C, t.j. z pravidiel, kde je A na ľavej strane, $V_D = \{S, A, a, B, b, C, D\}$.
- Do množiny V_D pribudnú všetky symboly dosiahnuteľné z neterminálu D, t.j. z pravidiel, kde je A na ľavej strane, $V_D = \{S, A, a, B, b, C, D\}$.
- Vidíme, že $V_D = \{S, A, B, C, D, a, b\}$. Teda symboly E, c sú nedosiahnuteľné a môžeme ich odstrániť.

Redukovaná gramatika

Redukovaná gramatika:

- $S \to Aa \mid Bb$
- $A \to C \mid DaD$
- $B \to D \mid C$
- $C \to S \mid a \mid \varepsilon$
- $D \to S \mid b$

A teraz z nej urobíme ekvivalentnú gramatiku bez ε-pravidiel:

Množina N_ε

Množina N_ε:

- $C \in N_\varepsilon$, pretože $C \to \varepsilon$ je v pravidlách
- $B \in N_\varepsilon$, pretože $C \in N_\varepsilon$ a zároveň $B \to C$ je v pravidlách
- $A \in N_\varepsilon$, pretože $C \in N_\varepsilon$ a zároveň $A \to C$ je v pravidlách
- Celkovo $N_\varepsilon = \{A, B, C\}$

A teraz z nej urobíme ekvivalentnú gramatiku bez ε-pravidiel:

Z gramatiky odstránime ε-pravidlá:

- $S \to Aa \mid Bb$
- $A \to C \mid DaD$
- $B \to D \mid C$
- $C \to S \mid a$
- $D \to S \mid b$

a pridáme pravidlá na základe N_ε

Podľa pravidla $S \to Aa$ pridáme:

- $S \to Aa \mid Bb \mid a$
- $A \to C \mid DaD$
- $B \to D \mid C$
- $C \to S \mid a$
- $D \to S \mid b$

a pridáme pravidlá na základe N_ε

Podľa pravidla $S \to Bb$ pridáme:

- $S \to Aa \mid Bb \mid a \mid b$
- $A \to C \mid DaD$
- $B \to D \mid C$
- $C \to S \mid a$
- $D \to S \mid b$

a pridáme pravidlá na základe N_ε

Podľa iných pravidiel nepribudne nič nové. Preto výsledná gramatika:

- $S \rightarrow Aa \mid Bb \mid a \mid b$
- $A \rightarrow C \mid DaD$
- $B \rightarrow D \mid C$
- $C \rightarrow S \mid a$
- $D \rightarrow S \mid b$

Navyše, keďže $S \notin N_\varepsilon$, nemusíme tentokrát pridávať nový počiatočný neterminál $\acute{S}$.
Následne odstránime jednoduché pravidlá.

Množiny $N_S, \ldots, N_D$

Najprv nájdeme množiny $N_S, \ldots, N_D$. Množiny N_S, N_A, N_B, N_C, N_D:

- $N_S = \{S\}$
- $N_A = \{A, C, S\}$
- $N_B = \{B, D, C, S\}$
- $N_C = \{C, S\}$
- $N_D = \{D, S\}$

Množiny $N_S, \ldots, N_D$

Keď odstránime jednoduché pravidlá:

- $S \to Aa \mid Bb \mid a \mid b$
- $A \to DaD$
- $B \to$
- $C \to a$
- $D \to b$

Ako pravé strany S, A, B, C, D následne doplníme pravé strany tých neterminálov, ktoré sú v príslušných množinách N_S, N_A, N_B, N_C, N_D:

- $S \rightarrow Aa \mid Bb \mid a \mid b$
- $A \rightarrow DaD \mid Aa \mid Bb \mid a \mid b$
- $B \rightarrow Aa \mid Bb \mid a \mid b$
- $C \rightarrow a \mid Aa \mid Bb \mid b$
- $D \rightarrow b \mid Aa \mid Bb \mid a$

Gramatika na úpravu č. 5

Upravte uvedenú gramatiku na redukovanú gramatiku:

- $S \to ABC$
- $A \to abA \mid BbC$
- $B \to Aa \mid BB$
- $C \to S \mid AB \mid \varepsilon$

Množina N_T

- $N_T = \emptyset$
- Keďže v pravidlách $C \to \varepsilon$, tak $C \in N_T$
- Keďže v pravidlách nie sú žiadne ďalšie pravidlá, ktoré by mali napravo alebo reťazec nad terminálmi, reťazec zložený z terminálov a neterminálu C, algoritmus končí a výsledná množina $N_T = \{C\}$
- Odstránili by sme teda neterminály S, A, B, teda **aj počiatočný neterminál!!!**
- Ak nastáva situácia, že **počiatočný neterminál nepatrí do množiny** N_T, tak potom **gramatika generuje len prázdny jazyk** a nemá zmysel pokračovať v jej redukcii!