## Parsing

## Homework 2 (CFG, PDA), due 19 April 2022 before the course

Laura Kallmeyer

SS 2022, Heinrich-Heine-Universität Düsseldorf

## Question 1 (CFG)

1. Consider the $C F G G_{1}$ with non-terminals $\{S, A, B\}$, terminals $\{a, b, c\}$, start symbol $S$ and productions

$$
S \rightarrow A S B B|A S B| c \quad A \rightarrow a A \mid a \quad B \rightarrow b
$$

Transform $G_{1}$ into an equivalent $C F G G_{1}^{\prime}$ in Chomsky Normal Form.
2. Consider the $C F G G_{2}$ with non-terminals $\{S, T\}$, terminals $\{a, b, c, e\}$, start symbol $S$ and productions

$$
S \rightarrow T b \quad S \rightarrow a \quad T \rightarrow S a
$$

Transform $G_{2}$ into an equivalent $C F G G_{2}^{\prime}$ without left recursion. Do so with the algorithm from the course. Perform two transformation, one with $S$ having index 1, $T$ index 2, and the other one with the reverse indexing. Which grammars do you obtain?
Make sure that you remove useless symbols after each transformation step (if there are any).

Solution:

1. $S \rightarrow A X|A Y| c \quad X \rightarrow S Z \quad Y \rightarrow S B \quad Z \rightarrow B B \quad A \rightarrow C_{a} A \mid a \quad C_{a} \rightarrow a \quad B \rightarrow b$
2. (a) First indexing:

$$
S_{1} \rightarrow T_{2} b \quad S_{1} \rightarrow a \quad T_{2} \rightarrow S_{1} a
$$

I Replace $T_{2} \rightarrow S_{1} a$ with $T_{2} \rightarrow T_{2} b a$ and $T_{2} \rightarrow a a$
$\sim$ new set of productions: $\quad S_{1} \rightarrow T_{2} b \quad S_{1} \rightarrow a \quad T_{2} \rightarrow T_{2} b a \quad T_{2} \rightarrow a a$
II Take a new non-terminal $X$ and replace the $T_{2}$-productions with $T_{2} \rightarrow a a \mid a a X$ and $X \rightarrow$ $b a X \mid b a$
$\leadsto$ new set of productions: $\quad S_{1} \rightarrow T_{2} b\left|a \quad T_{2} \rightarrow a a\right| a a X \quad X \rightarrow b a X \mid b a$
(b) Second indexing:
$S_{2} \rightarrow T_{1} b \quad S_{2} \rightarrow a \quad T_{1} \rightarrow S_{2} a$
I Replace $S_{2} \rightarrow T_{1} b$ with $S_{2} \rightarrow S_{2} a b$
$\sim$ new set of productions: $\quad S_{2} \rightarrow S_{2} a b \quad S_{2} \rightarrow a \quad T_{1} \rightarrow S_{2} a$
Now, $T_{1}$ has become useless and can be removed.
$\sim$ new set of productions: $\quad S_{2} \rightarrow S_{2} a b \quad S_{2} \rightarrow a$
II Take a new non-terminal $X$ and replace the $S_{2}$ productions with $S_{2} \rightarrow a \mid a X$ and $X \rightarrow$ $a b X \mid a b$
$\leadsto$ new set of productions: $\quad S_{2} \rightarrow a|a X \quad X \rightarrow a b X| a b$

## Question 2 (PDA)

Give a PDA that recognizes the following language, acceptance with final state: $\left\{(a b)^{n} b^{m} \mid m \geq n \geq 0\right\}$.

Solution: Initial stack symbol is \#.


## Question 3 (PDA)

Consider the following PDA $M$, initial stack symbol is \#.


1. Which of the following words are elements of $L(M)$ ?
(i) $a$
(ii) $a b a b$
(iii) babba
(iv) babbaa
(v) $\varepsilon$
2. Give $L(M)$ and $N(M)$.

Solution:

1. (ii) and (iv)
2. $L(M)=N(M)=\left\{\left.w \in\{a, b\}^{+}| | w\right|_{a}=|w|_{b}\right\}$
