# TDT4205 PS 2

## Øyvind Skaaden (oyvindps)

### February 21, 2022

#### Problem 1.

**1.1**) As discussed in Piazza, it is not possible to rewrite the following grammar in table 1 into LL(1) just by using left factoring and/or left recursion elimination.

Table 1: Original grammar, supplied in PS2 description.

We can already see a problem when for instance we have the input sequences ending in ebz and embz. By just looking at e we can't determine that the e is part of the X tree or Y tree, without looking at the next symbol b or m.

The solution is to rewrite the grammar so that there is only one type of tree, X. The X tree always start with MB, and can be followed by an arbitrary number of eMB and then an optimal eB.

This can be done by rewriting the X to  $X \to MBX'$ , where  $X' \to eX'' | \epsilon$  and  $X'' \to X | B$ . Then the Y is unused and we can remove that and the Y from S. The new S is therefore  $S \to wXz$ .

We can not set  $B \to b | \epsilon$  because then we could have an eB block in the middle, which is not allowed by the FORTRAN language.

The new grammar is written down in table 2.

Table 2: Rewritten grammar to comply with LL(1).

## **1.2**) The FIRST and FOLLOW sets are described in table 3.

Table 3: FIRST and FOLLOW sets.

NT	FIRST	FOLLOW	$\rightarrow \epsilon$ ?
S	W	-	no
X	m	z	no
X X' X"	е	z	yes
X''	m,b	z	no
M	m	b	no
В	b	e,z	no

The predictive parsing table is shown in table 4.

Table 4: Predictive parsing table for the new grammar.

	W	m	е	b	z
S	$S \to wXz$				
X		$X \to MBX'$			
X'			$X' \to eX''$		$X' \to \epsilon$
X''		$X'' \to X$		$X'' \to B$	
M		$M \to m$			
В				$B \rightarrow b$	