

FM-Index for pattern matching with BWT

1. The table below shows the BWT for $S = \text{barbara}\$$. Complete the columns to create the FM-Index, which consists of:

- (a) The BWT of S , i.e. the L column of $\pi^{\text{sorted}}(S)$
- (b) $M[c]$, the first index of c in F (note that F is actually not part of the FM-Index)
- (c) $\text{occ}(c, i)$, the number of times c occurs in $L[1 \dots i]$, inclusive

i	F	L	$\text{occ}(\$)$	$\text{occ}(a)$	$\text{occ}(b)$	$\text{occ}(r)$
1		a				
2		r				
3		b				
4		b				
5		r				
6		\$				
7		a				
8		a				

c	$M[c]$
\$	
a	
b	
r	

Extra practice: compute $BW(S)$ for this string and check your answer with the table above. Then reverse the BWT and make sure you can obtain the original string.

2. Use the FM-Index above to find the start indices of the pattern $P = ba$ in the original string.