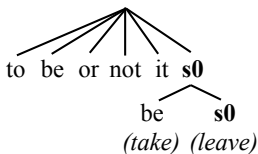


Original Corpus

to be or not to be
take it or leave it

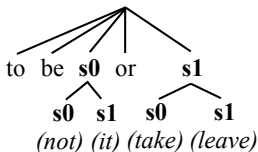
Repeat-All Encoding



to be or not to be
s0 be it or s0 s0 it

13 encoded tokens

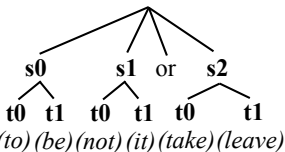
Repeat-Symbol Encoding



to be or s0 s0 to be
s1 s0 s0 s1 or s1 s1 s0 s1

16 encoded tokens

No-Repeats Encoding



s0 t0 s0 t1 or s1 t0 s0 t0 s0 t1
s2 t0 s1 t1 or s2 t1 s1 t1

20 encoded tokens