

## Key Assumption in N-gram LM:

$$p(w_m | w_1, \dots, w_{m-n+1}, \dots, w_{m-1}) = p(w_m | w_{m-n+1}, \dots, w_{m-1})$$

Ignored

Does this assumption hold?