

Definition of a SLM

- Vocabulary set: $V = \{t_1, t_2, \dots, t_N\}$, N terms
- Sequence of M terms: $s = w_1 w_2 \dots w_M$, $w_i \in V$
- Probability of sequence s :
 - $p(s) = p(w_1 w_2 \dots w_M) = ?$
- How do we compute this probability? How do we “generate” a sequence using a probabilistic model?
 - Option 1: Assume each sequence is generated as a “whole unit”
 - Option 2: Assume each sequence is generated by generating one word each time
 - Each word is generated independently
 - Option 3: ??