

# Beam Search — Algorithm

1. Begin with a single empty hypothesis in the beam.
2. In each time step:
  - 2.1 Extend all hypotheses in the beam by  $k$  most probable words (we call these candidate hypotheses).
  - 2.2 Sort the candidate hypotheses by their score.
  - 2.3 Put the best  $k$  hypotheses in the new beam.
  - 2.4 If a hypothesis from the beam reaches the end-of-sentence symbol, we move it to the list of finished hypotheses.
3. Finish (1) at the final time step or (2) all  $k$ -best hypotheses end with  $\text{</s>}$ .
4. Sort the hypotheses by their score and output the best one.