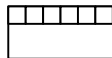
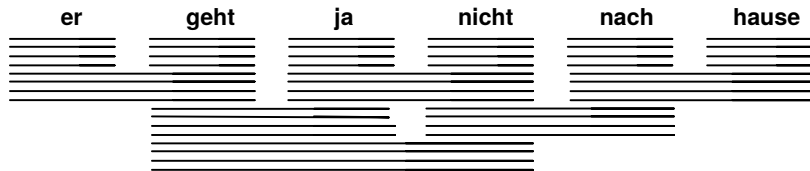


## Stage 2: Decoding (Beam Search)



initial hypothesis: no input words covered, no output produced