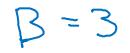


Sequence to sequence models

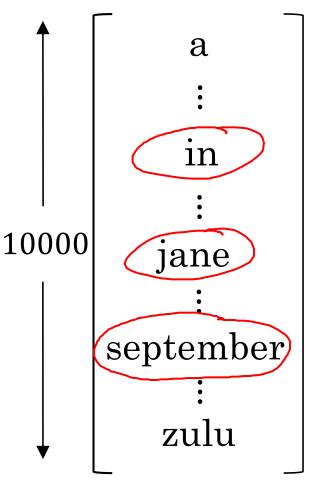
Beam search

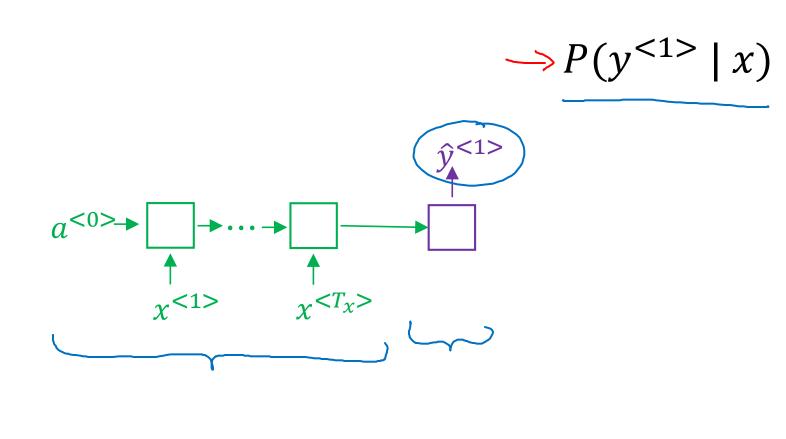
Beam search algorithm

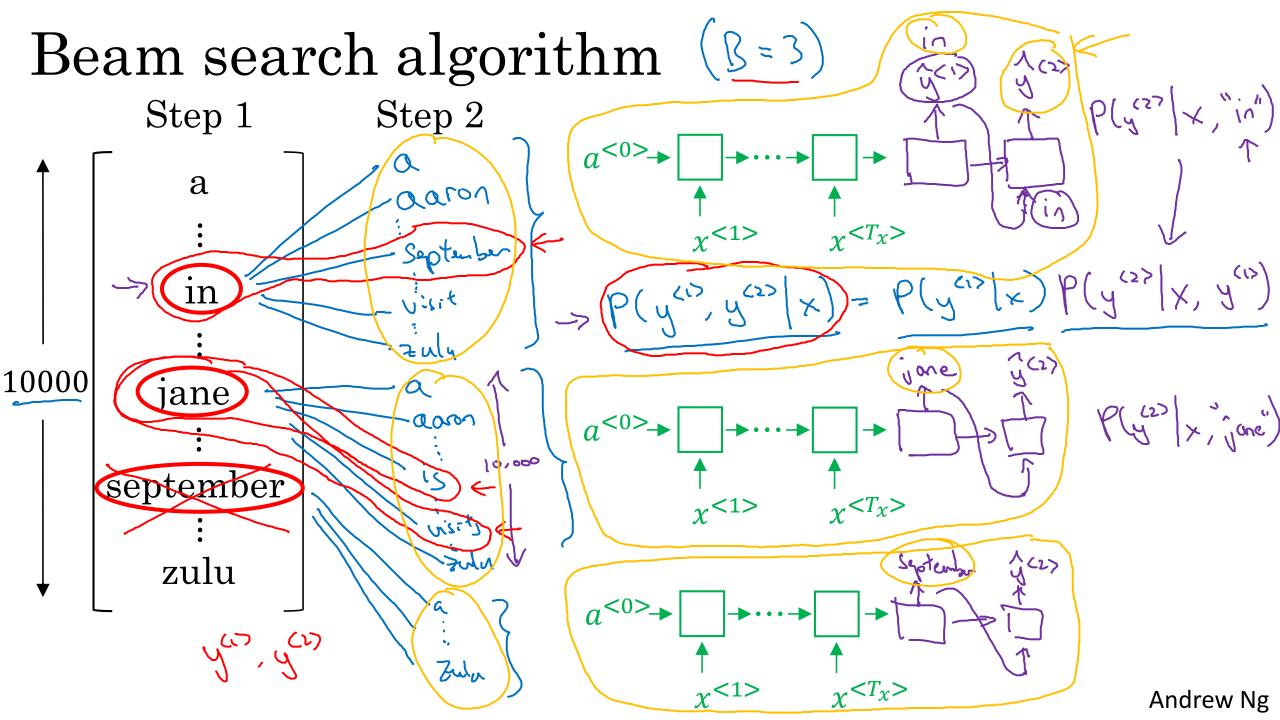


B=3 (bean width)









Beam search (B = 3)



$$P(v^{<1>}, v^{<2>}|x)$$

