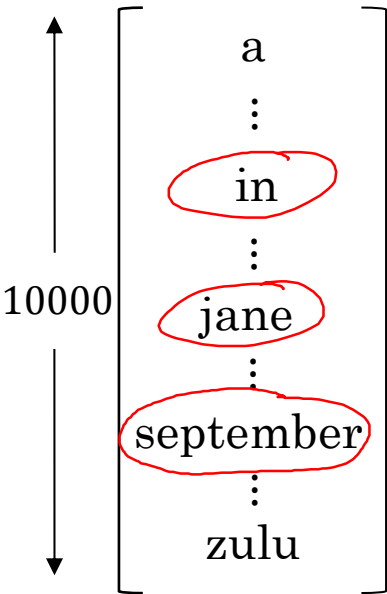# Sequence to sequence models

Beam search

Beam search algorithm

Step 1

$B = 3$ word (beam width)

consider 3 or a the

$\rightarrow P(y^{<1>} | x)$

$\hat{y}^{<1>}$
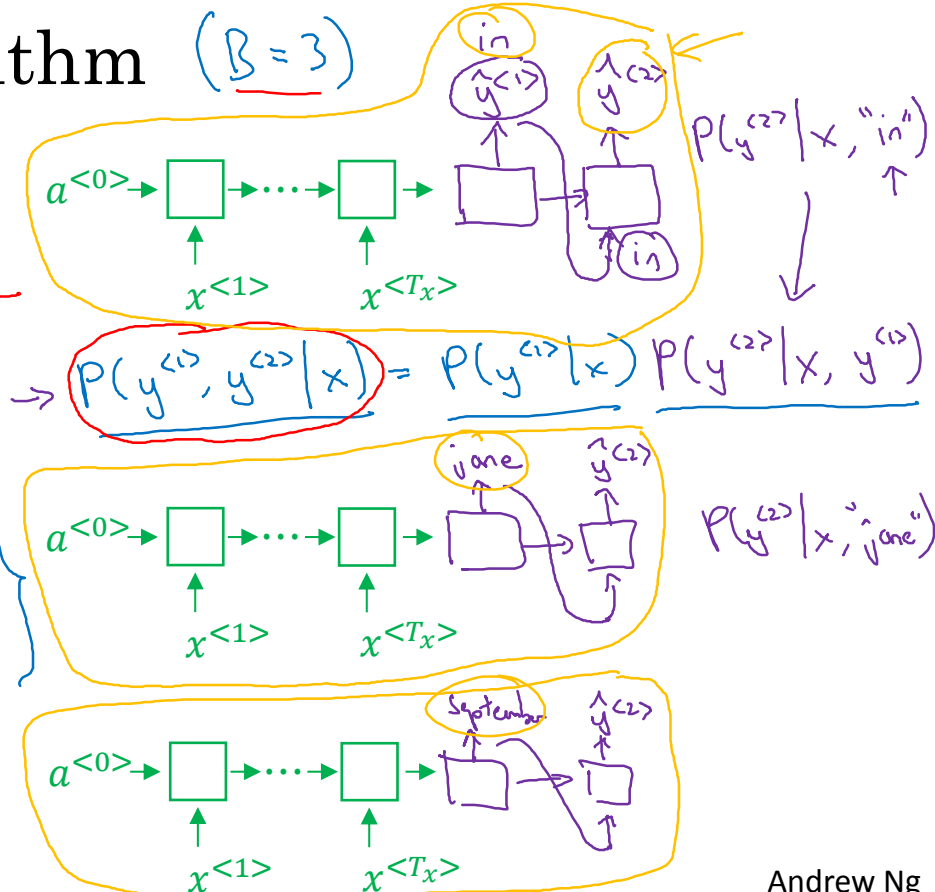
$\begin{bmatrix} a \\ \vdots \\ in \\ \vdots \\ jane \\ \vdots \\ september \\ \vdots \\ zulu \end{bmatrix}$ 10000

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow \square$

$x^{<1>}$ $x^{<T_x>}$

Andrew Ng

Beam search algorithm

Step 1   Step 2 $(B = 3)$

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

$P(y^{<2>} | x, \text{"in"})$

$P(y^{<2>} | x, \text{"jane"})$

10000

$y^{<1>}, y^{<2>}$

Andrew Ng
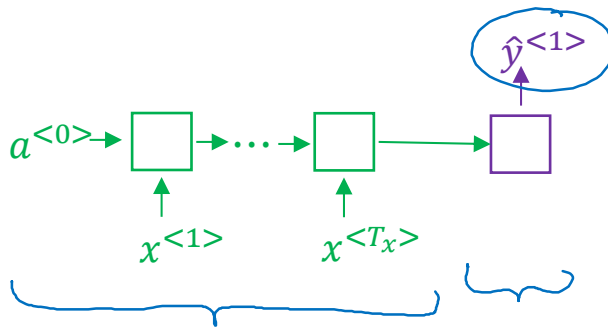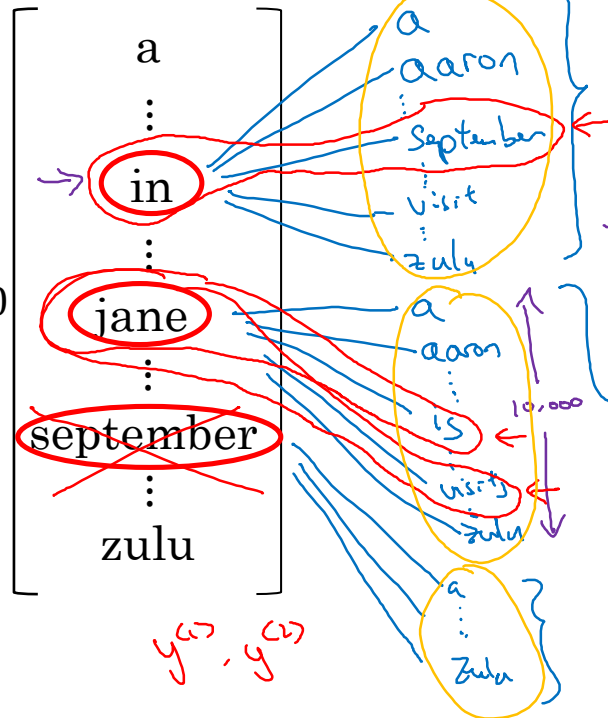
Beam search ($B = 3$)

$B = 1 \rightsquigarrow$ greedy search

in september
— a
aaron
jane
zulu

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow \square \rightarrow \square \rightarrow \square$

$x^{<1>}$ ... $x^{<T_x>}$

in   september $\hat{y}^{<3>}$

$P(y^{<3>} \mid x, \text{"in september"})$

jane   is $\hat{y}^{<3>}$

jane is
— a
visits
zulu

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow \square \rightarrow \square \rightarrow \square$

$x^{<1>}$ ... $x^{<T_x>}$
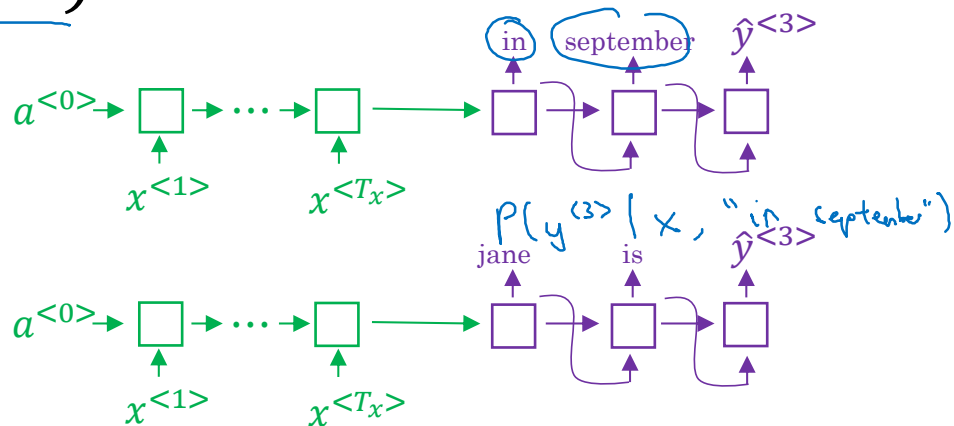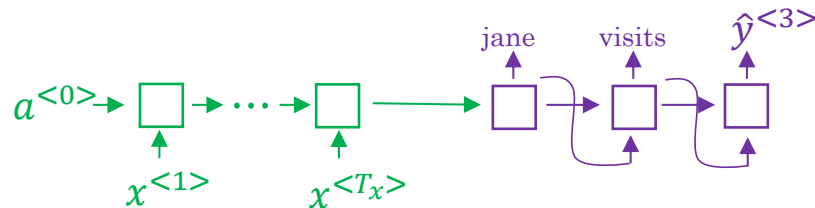
jane   visits $\hat{y}^{<3>}$

jane visits
— a
africa
zulu

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow \square \rightarrow \square \rightarrow \square$

$x^{<1>}$ ... $x^{<T_x>}$

$P(y^{<1>}, y^{<2>} \mid x)$ jane visits africa in september. <EOS>