Sequence to sequence models
_____

Beam search

deeplearning.ai

Beam search algorithm

$B = 3$ (beam width)

Step 1

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

$$\begin{bmatrix} a \\ \vdots \\ in \\ \vdots \\ jane \\ \vdots \\ september \\ \vdots \\ zulu \end{bmatrix} \quad 10000$$

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow \square \rightarrow \hat{y}^{<1>}$

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

Beam search algorithm

$(B = 3)$

Step 1   Step 2

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

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

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

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

10000

Andrew Ng
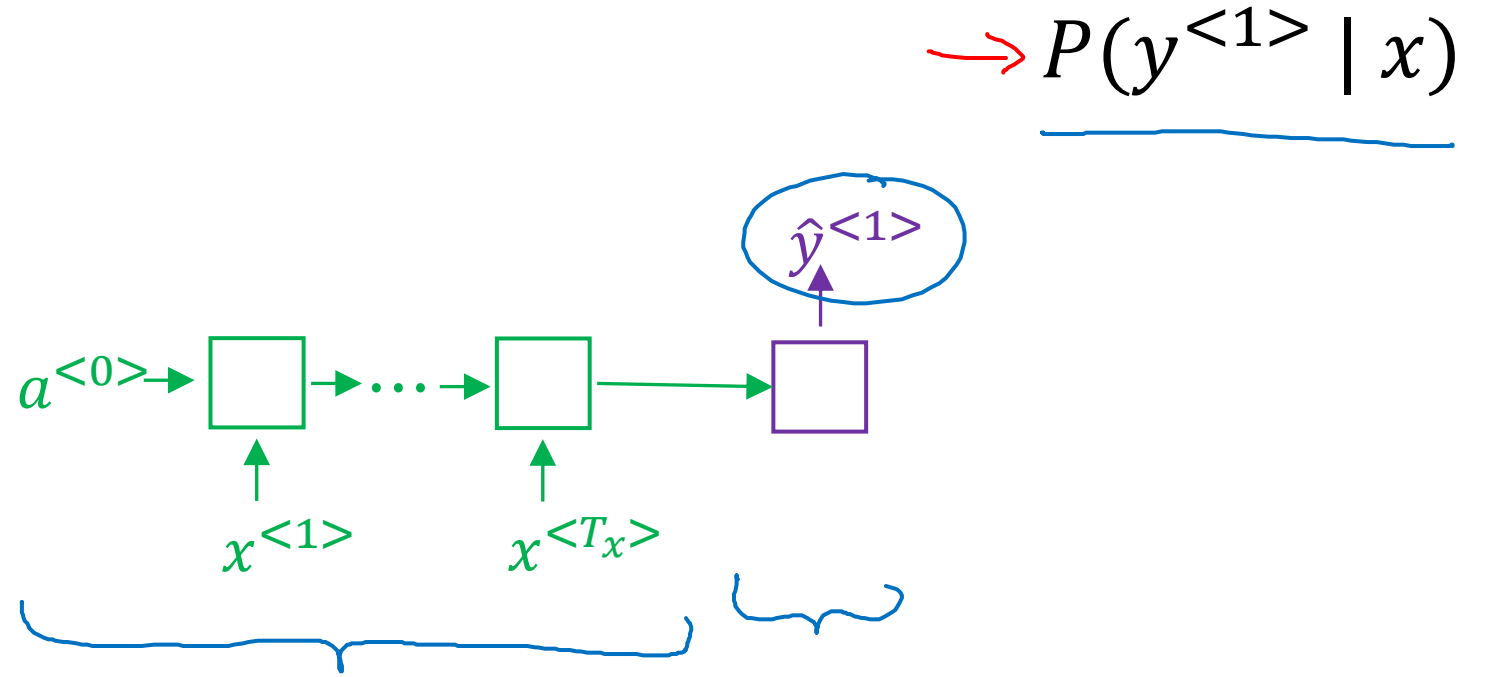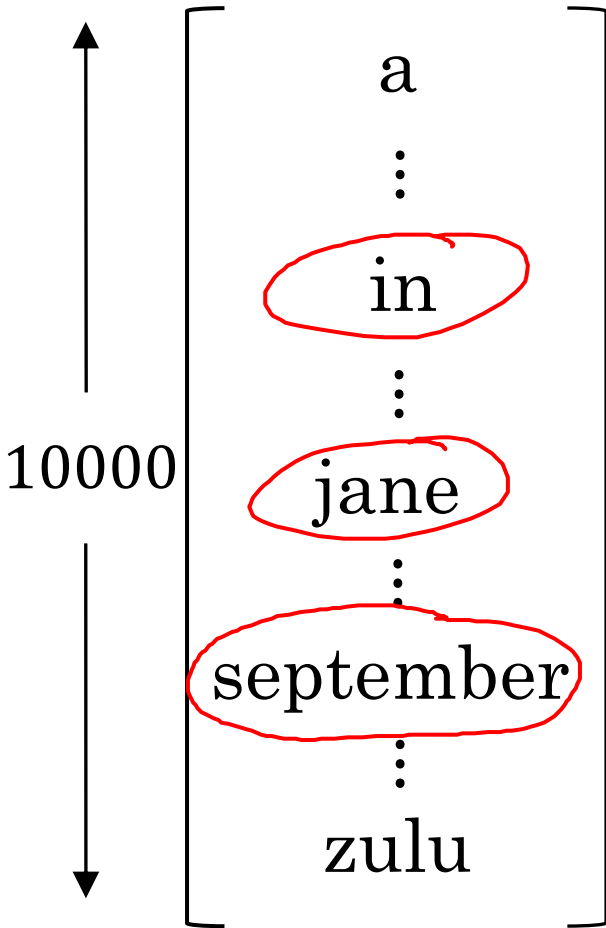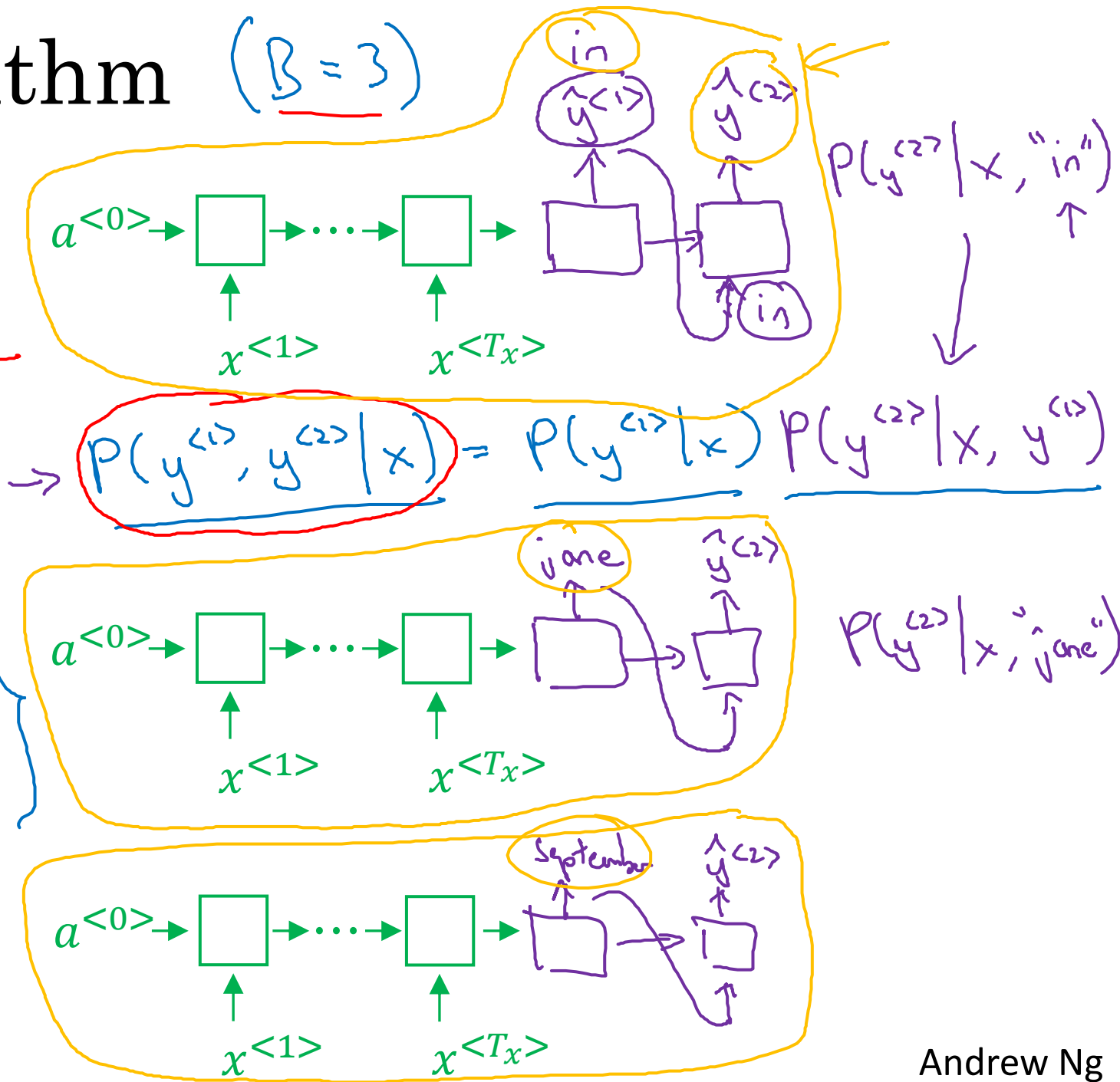
Beam search ($B = 3$)

$B = 1 \rightsquigarrow$ greedy search

in september
a
aaron
jane
zulu

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

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

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

jane is
a
visits
zulu

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

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

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

$x^{<1>}$ $x^{<T_x>}$
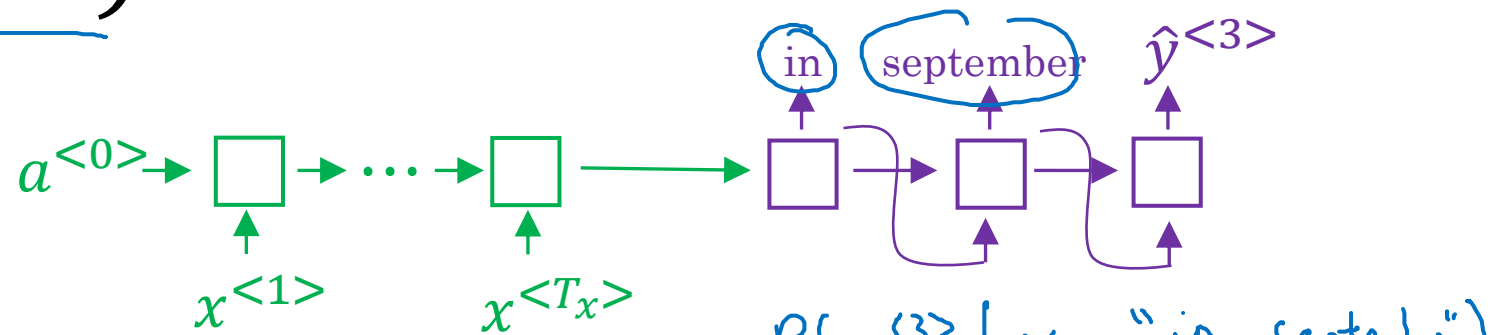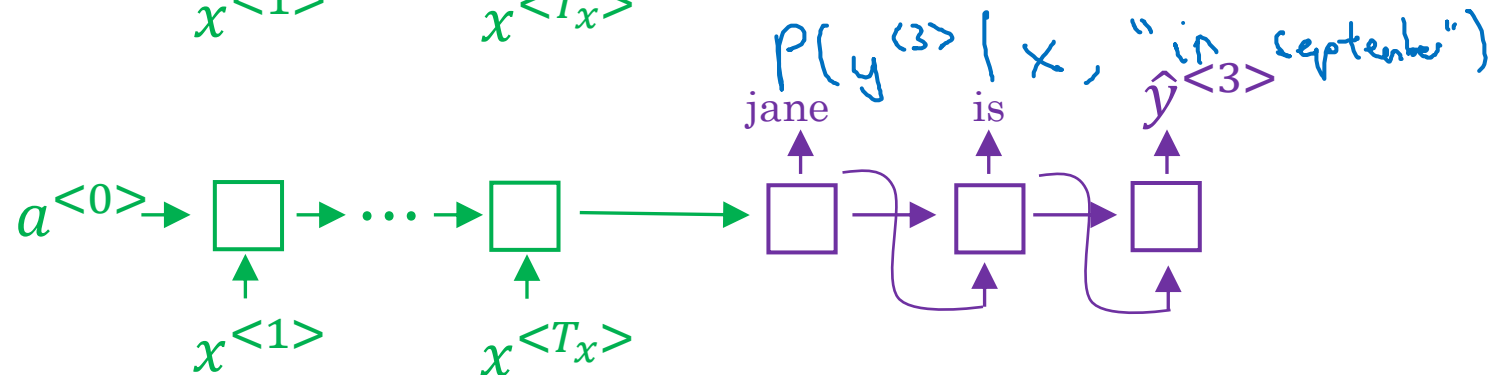
jane visits
a
africa
zulu

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

jane | visits | $\hat{y}^{<3>}$

$x^{<1>}$ $x^{<T_x>}$
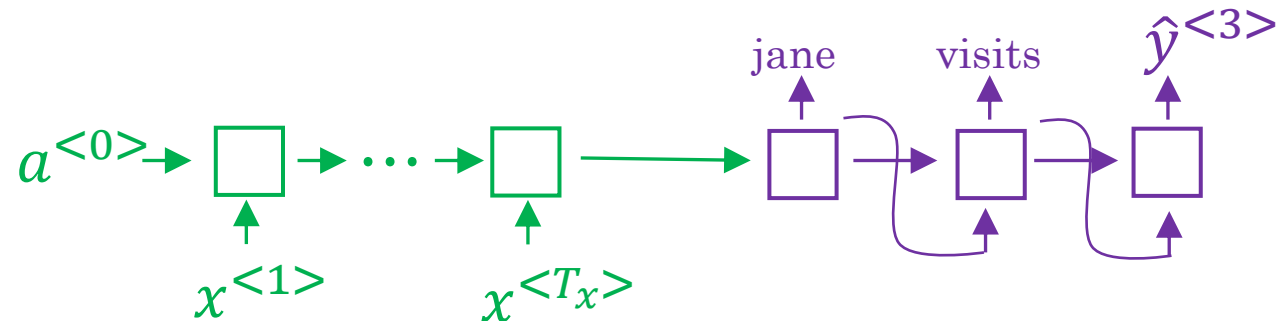
$P(y^{<1>}, y^{<2>} \mid x)$

jane visits africa in september. <EOS>

Andrew Ng