Sequence to sequence models

Beam search

deeplearning.ai

Beam search algorithm

$B = 3$ (beam width)

Step 1

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

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

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

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

Andrew Ng

Beam search algorithm

$(B = 3)$

Step 1   Step 2

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

Step 2 lists:
a, aaron, September, visit, zulu
a, aaron, is, visit, zulu   10,000
a, ... , zulu

$y^{(1)}, y^{(2)}$

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

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow$ in, $y^{<1>}$, $y^{<2>}$, in

$P(y^{(2)} | x, "in")$

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow$ jane, $y^{<2>}$

$P(y^{(2)} | x, "jane")$

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow$ September, $y^{<2>}$
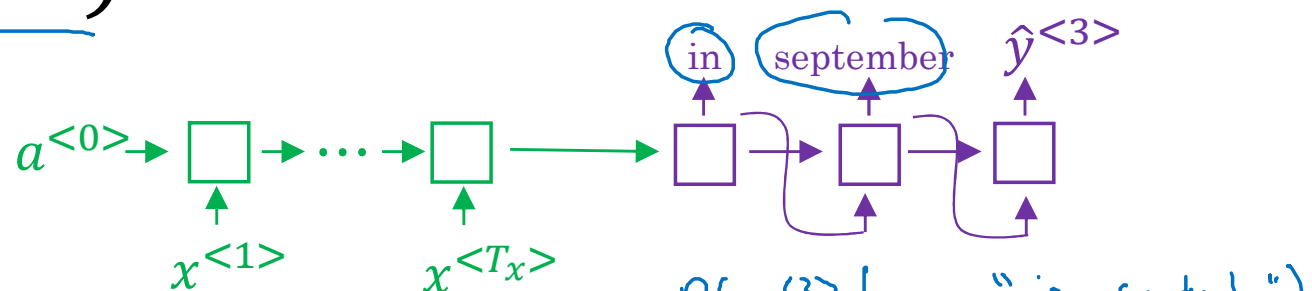
$x^{<1>} \quad x^{<T_x>}$

Andrew Ng
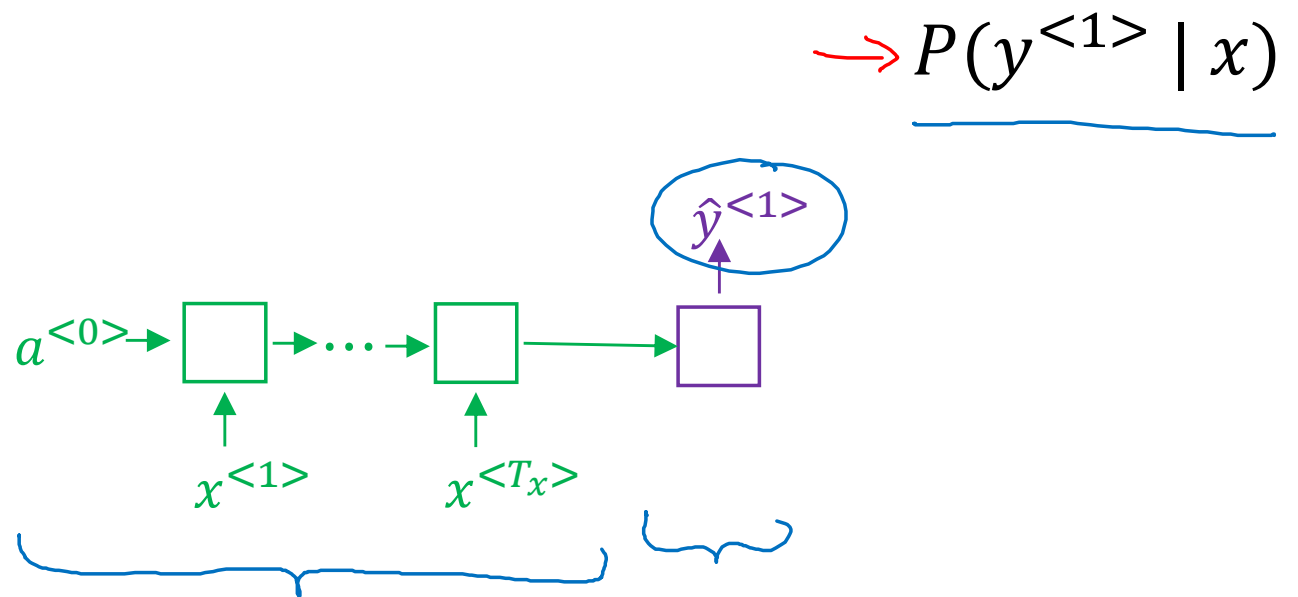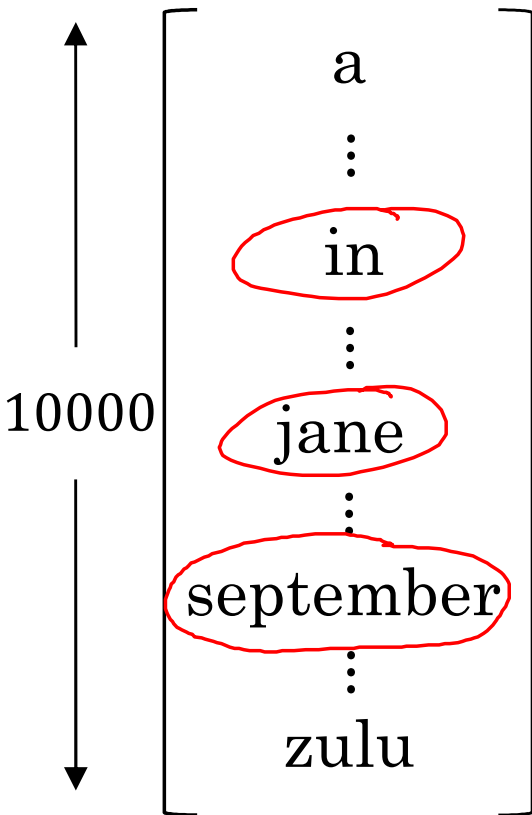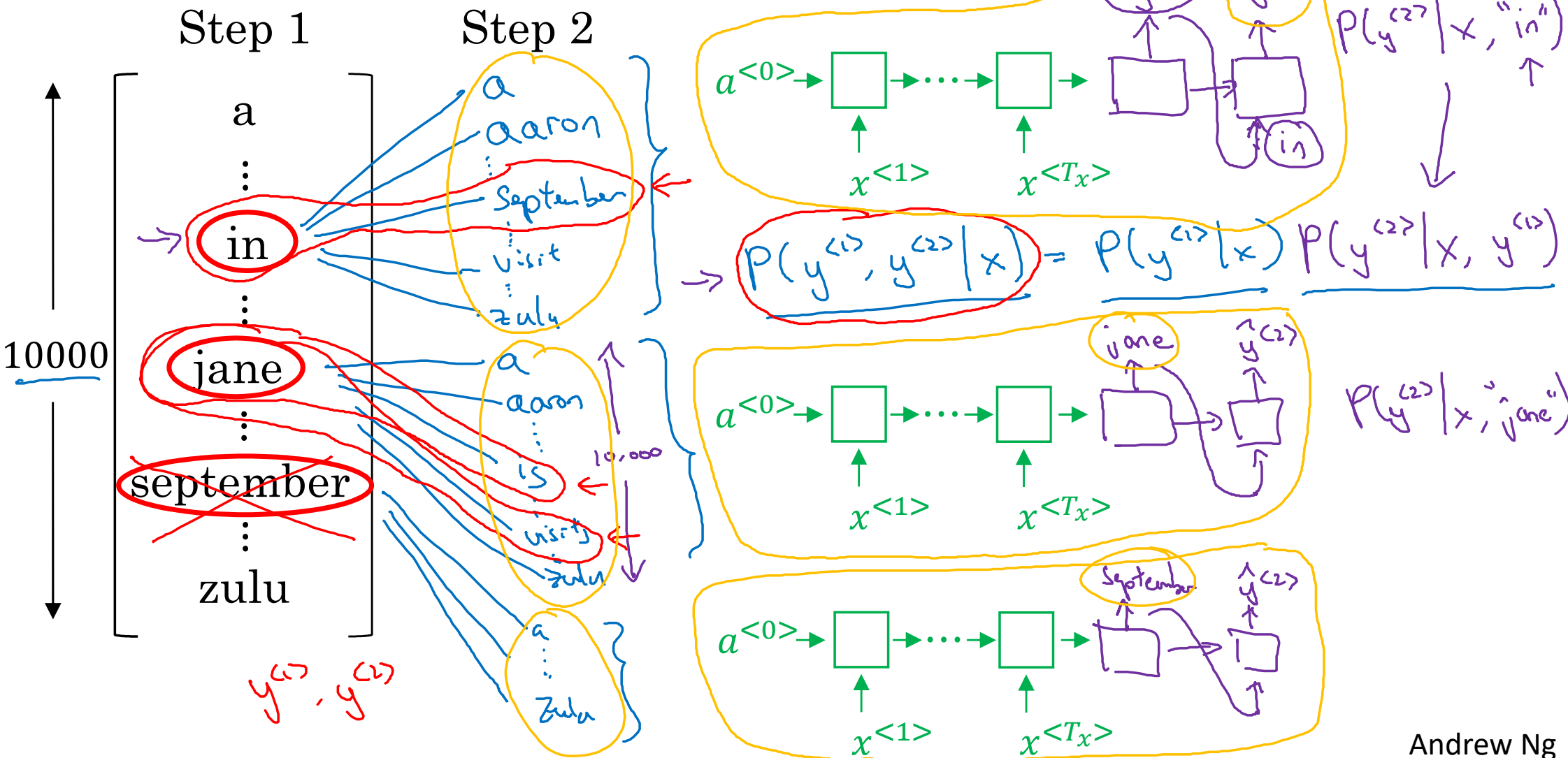
Beam search ($B = 3$)

$B = 1 \longrightarrow$ greedy search

in september

a
aaron
jane
zulu

jane is

a
visits
zulu
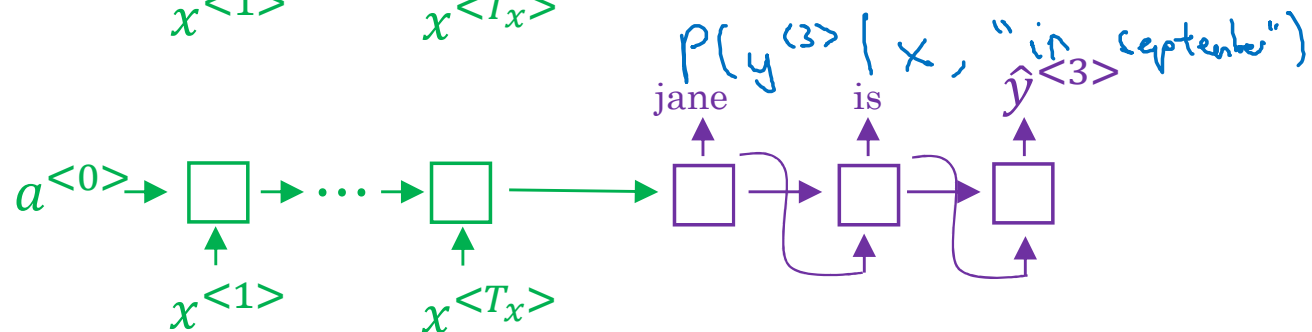
jane visits

a
africa
zulu

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

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

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

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

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

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

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

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

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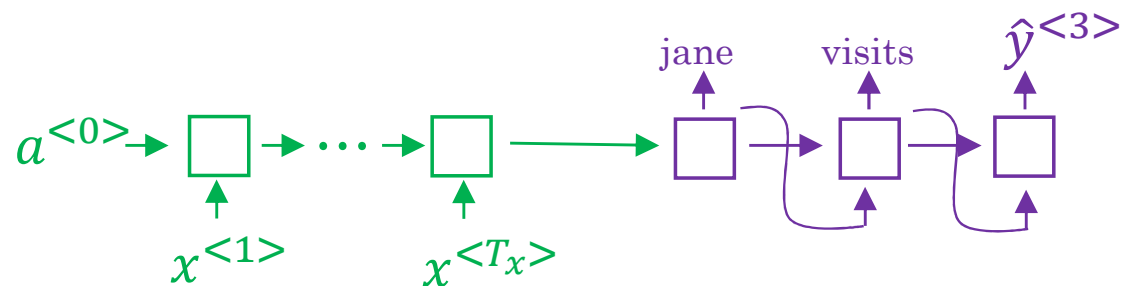
jane visits $\hat{y}^{<3>}$

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

jane visits africa in september. <EOS>

Andrew Ng