Sequence to sequence models

Beam search

deeplearning.ai

Beam search algorithm

$B = 3$ (beam width)

Step 1

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

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

10000

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

$\hat{y}^{<1>}$

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

Andrew Ng

Beam search algorithm

$(B = 3)$

Step 1

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

10000

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

Step 2

a
aaron
September
visit
zulu

a
aaron
is
visit
zulu

10,000

a
:
zulu

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

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

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

in
$\hat{y}^{<1>}$ $\hat{y}^{<2>}$
in

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

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

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

jane $\hat{y}^{<2>}$

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

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow$
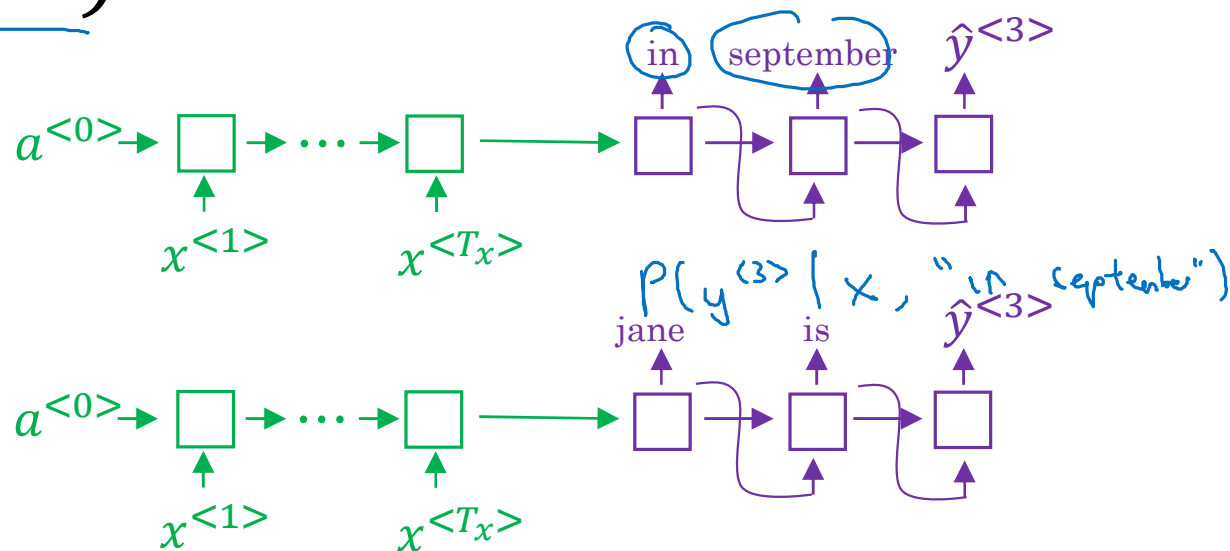
$x^{<1>}$ $x^{<T_x>}$

September $\hat{y}^{<2>}$

Andrew Ng
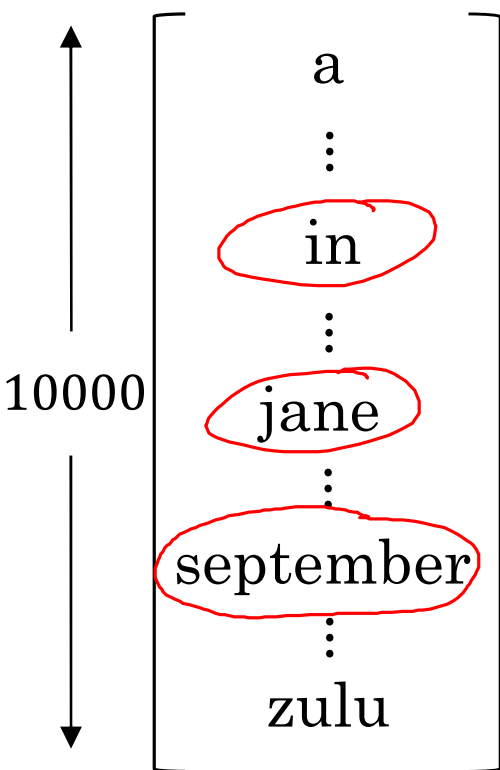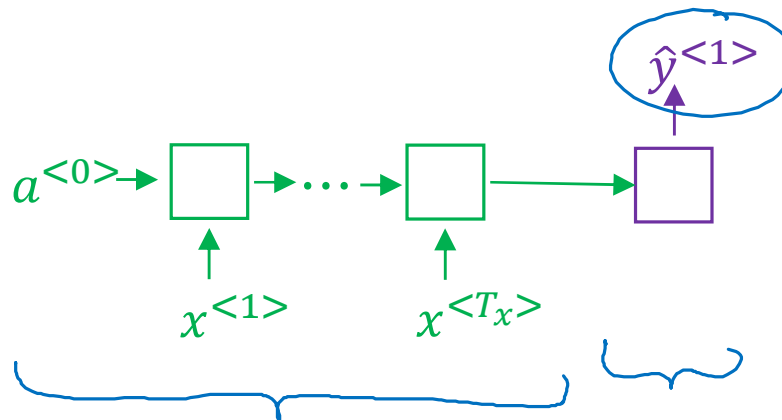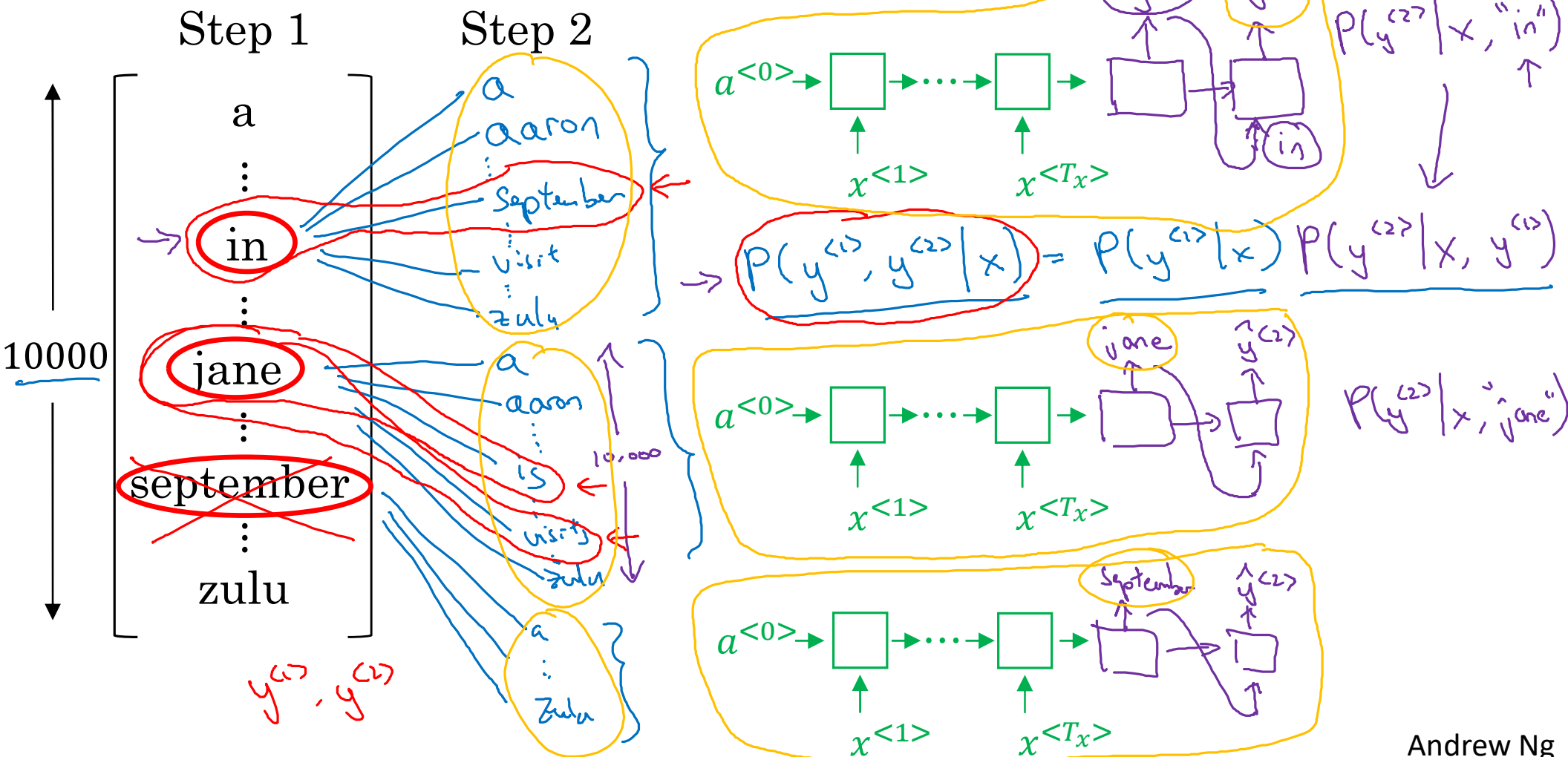
Beam search ($B = 3$)

$B = 1 \longrightarrow$ greedy search

in september

a
aaron
jane
zulu

jane is

a
visits
zulu
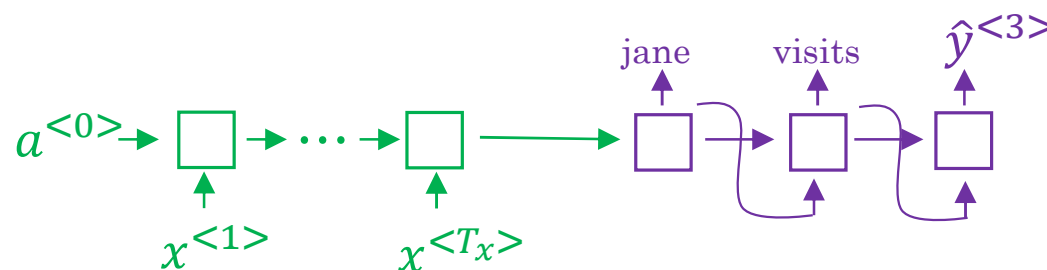
jane visits

a
africa
zulu

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

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

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

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

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

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

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

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

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

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

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

jane visits africa in september. <EOS>

Andrew Ng