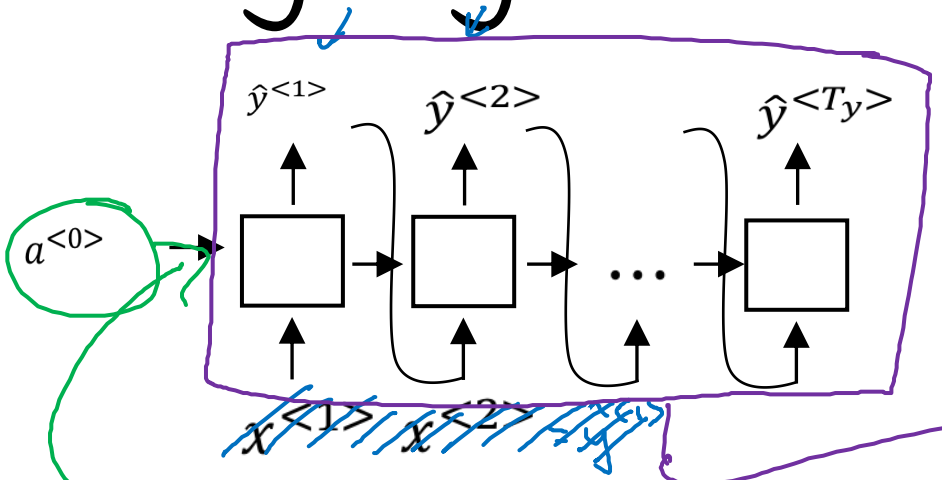
deeplearning.ai

Sequence to sequence models

Picking the most likely sentence

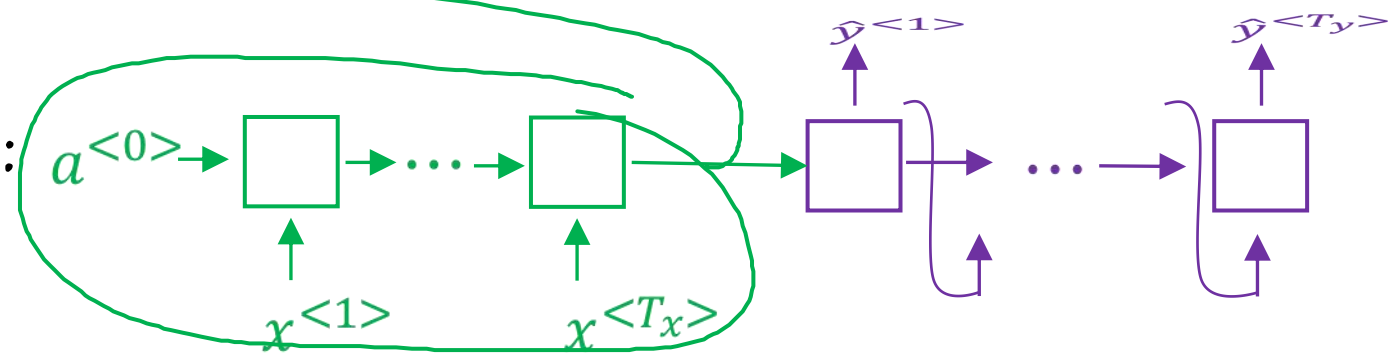
Machine translation as building a conditional language model

Language model:



$P(y^{<1>}, \ldots, y^{<T_y>})$

Machine translation:



"Conditional language model"

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

Finding the most likely translation

Jane visite l'Afrique en septembre.

$$P(y^{<1>}, \ldots, y^{<T_y>} \mid x)$$

English

French

→ Jane is visiting Africa in September.

→ Jane is going to be visiting Africa in September.

→ In September, Jane will visit Africa.

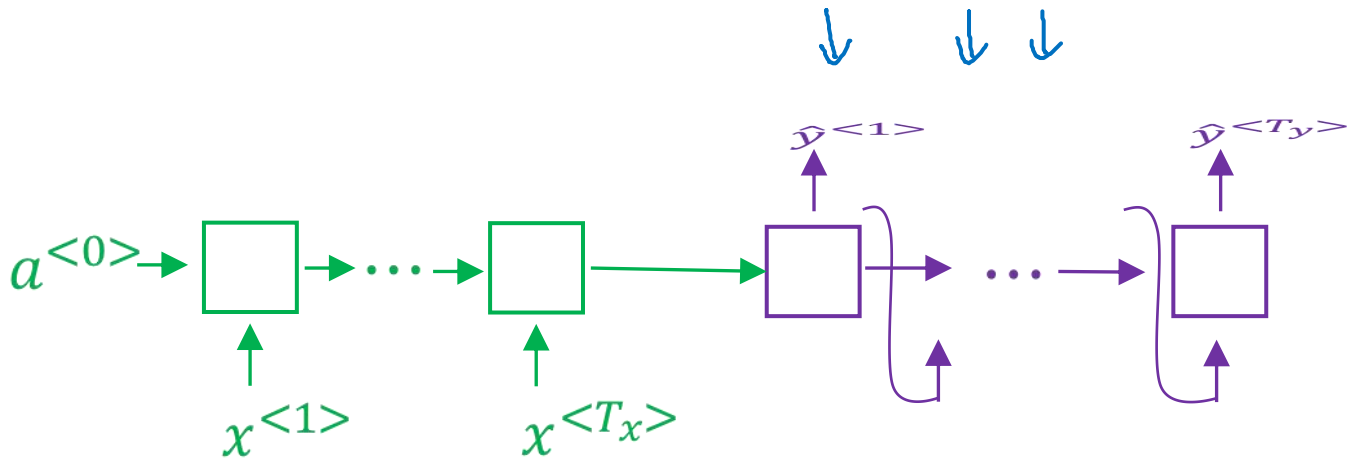
→ Her African friend welcomed Jane in September.

$$\arg\max_{y^{<1>},\ldots,y^{<T_y>}} P(y^{<1>}, \ldots, y^{<T_y>} \mid x)$$

Andrew Ng

Why not a greedy search?

$p(\hat{y}^{<1>} | x)$



$\arg\max_y P(\hat{y}^{<1>}, \hat{y}^{<2>}, \ldots, \hat{y}^{<T_y>} | x)$

10,000

10

$10{,}000^{10}$

$P(y|x)$

→ Jane is visiting Africa in September.

→ Jane is going to be visiting Africa in September.

$P(\text{Jane is going} | x) > P(\text{Jane is visiting} | x)$

Andrew Ng