



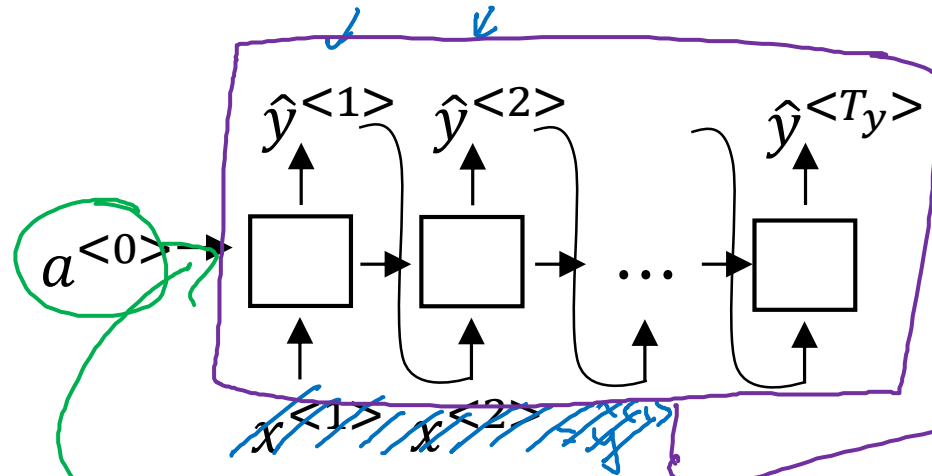
deeplearning.ai

Sequence to sequence models

Picking the most likely sentence

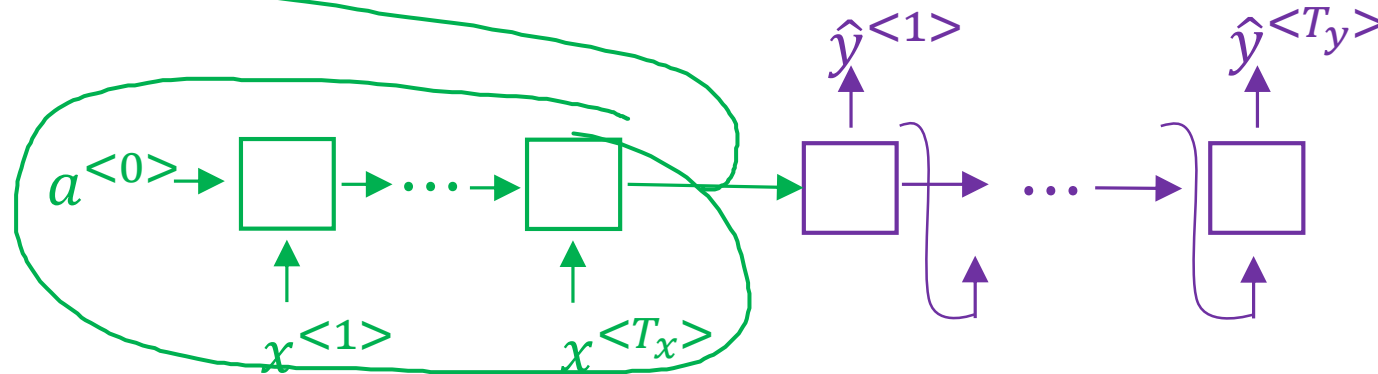
Machine translation as building a conditional language model

Language model:



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

Machine translation:



"Conditional language model"

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

Finding the most likely translation

Jane visite l'Afrique en septembre.

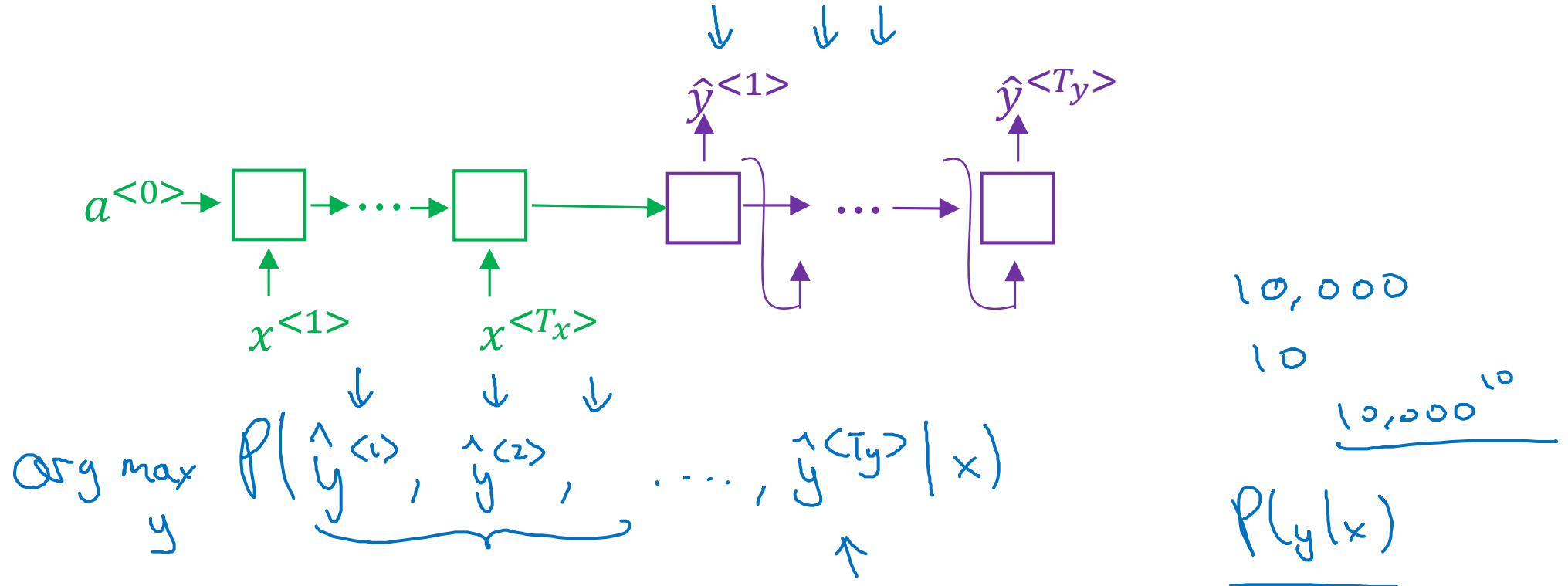
$$P(y^{<1>}, \dots, y^{<T_y>} | x)$$

- 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.

$$\arg \max_{y^{<1>}, \dots, y^{<T_y>}} \underline{P(y^{<1>}, \dots, y^{<T_y>} | x)}$$

Why not a greedy search?

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



→ Jane is visiting Africa in September.

→ Jane is going to be visiting Africa in September.

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