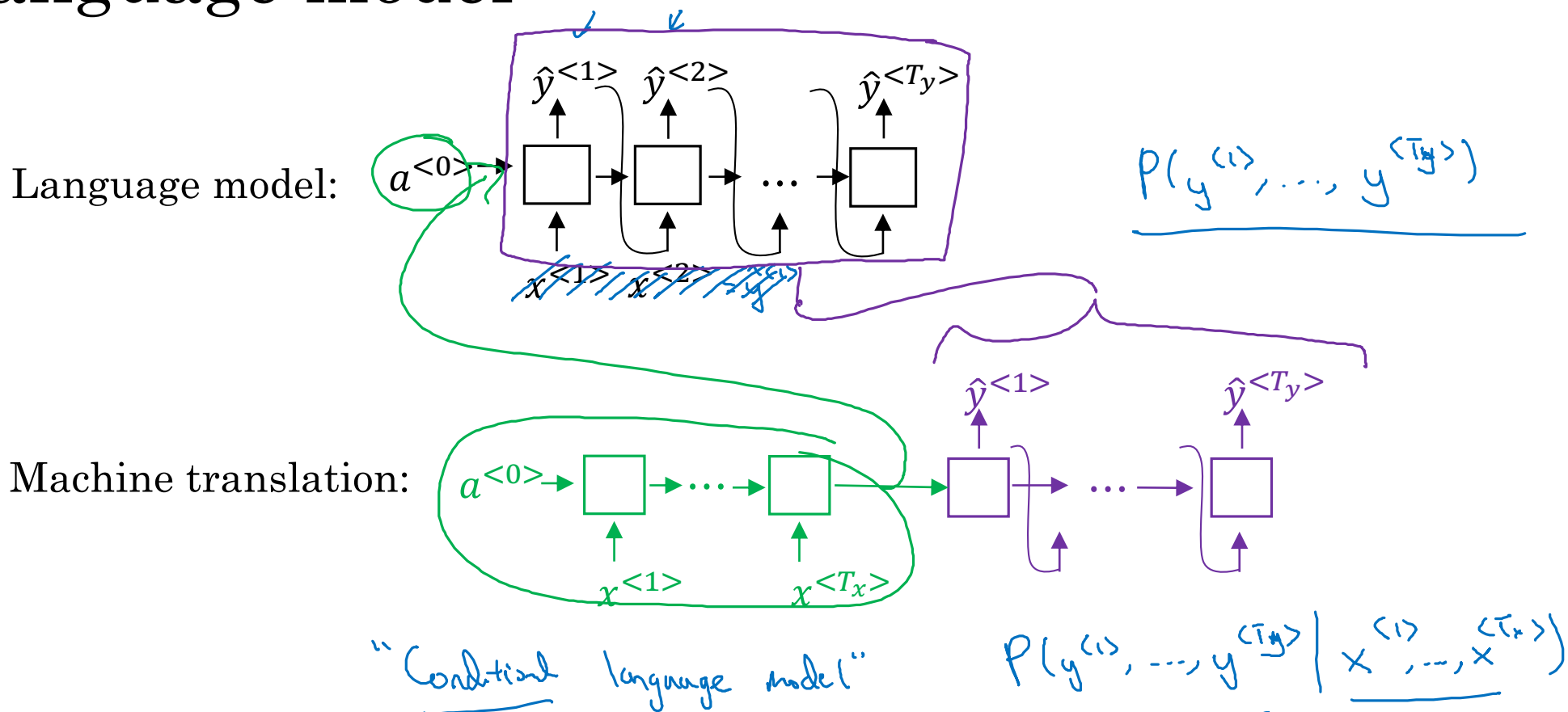deeplearning.ai

Sequence to sequence models

Picking the most likely sentence

Machine translation as building a conditional language model

Language model:

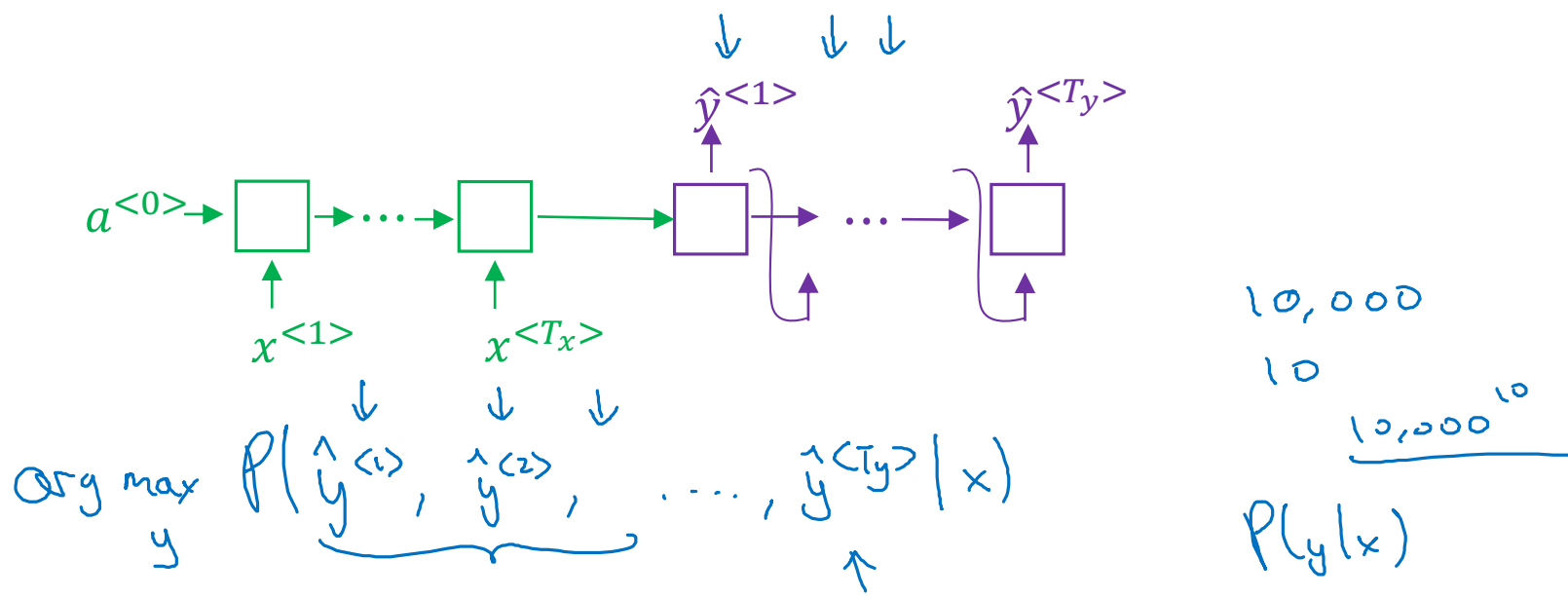$a^{<0>}$ $\hat{y}^{<1>}$ $\hat{y}^{<2>}$ $\hat{y}^{<T_y>}$

$x^{<1>}$ $x^{<2>}$ $x^{<3>}$

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

Machine translation:

$a^{<0>}$ $x^{<1>}$ $x^{<T_x>}$ $\hat{y}^{<1>}$ $\hat{y}^{<T_y>}$

"Conditional language model"

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

Andrew Ng

Finding the most likely translation

Jane visite l'Afrique en septembre.

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

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.

$$\underset{y^{<1>}, \dots, y^{<T_y>}}{\arg\max} \; P(y^{<1>}, \dots, y^{<T_y>} \mid x)$$

Andrew Ng

Why not a greedy search?

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

$$\underset{y}{\arg\max}\ P(\hat{y}^{<1>},\ \hat{y}^{<2>},\ \dots,\ \hat{y}^{<T_y>}|x)$$

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