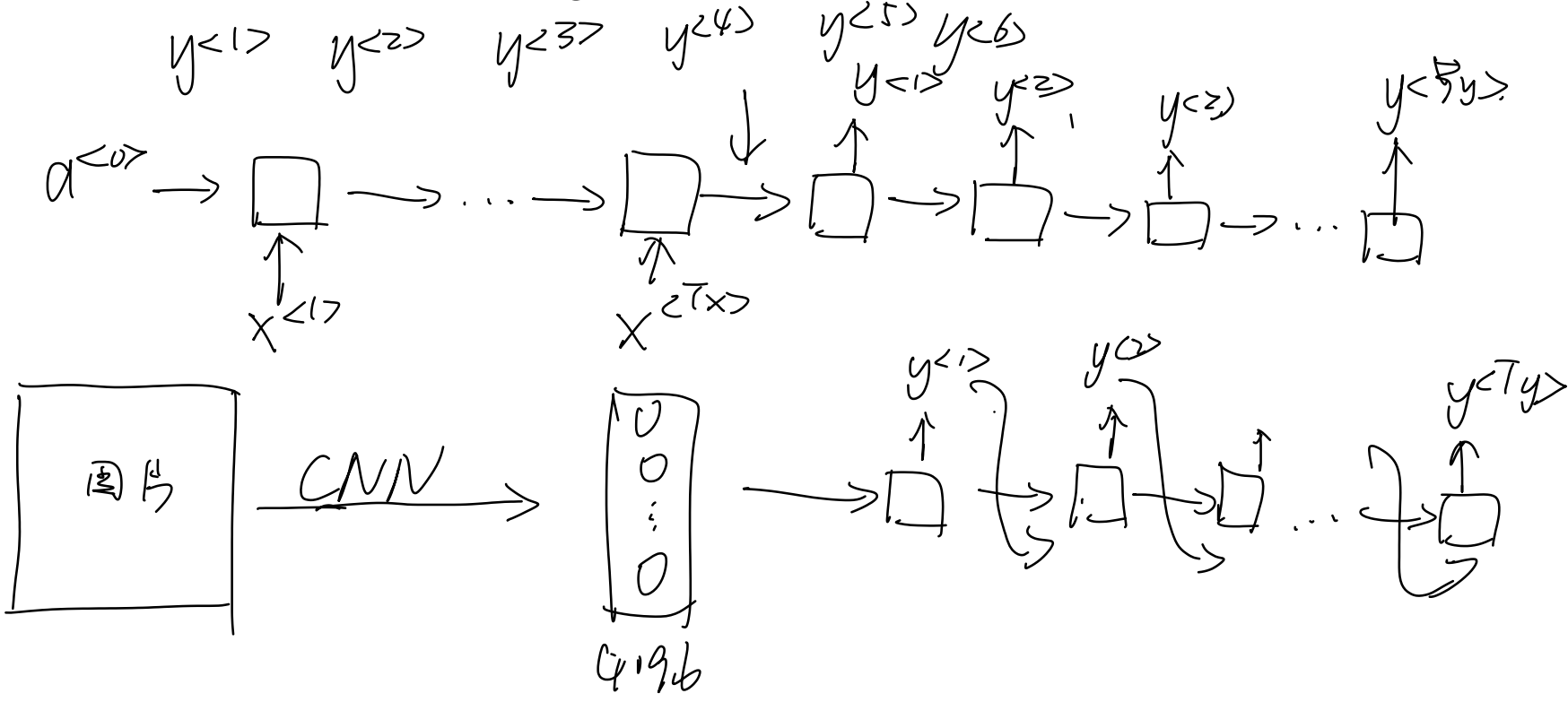


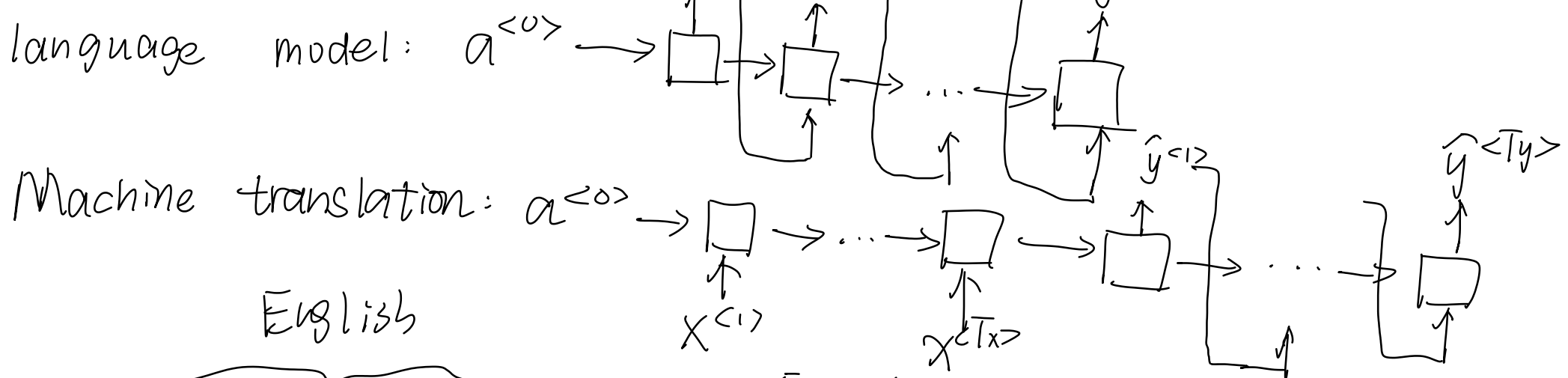
Basic Models

→ Jane is visiting Africa in September

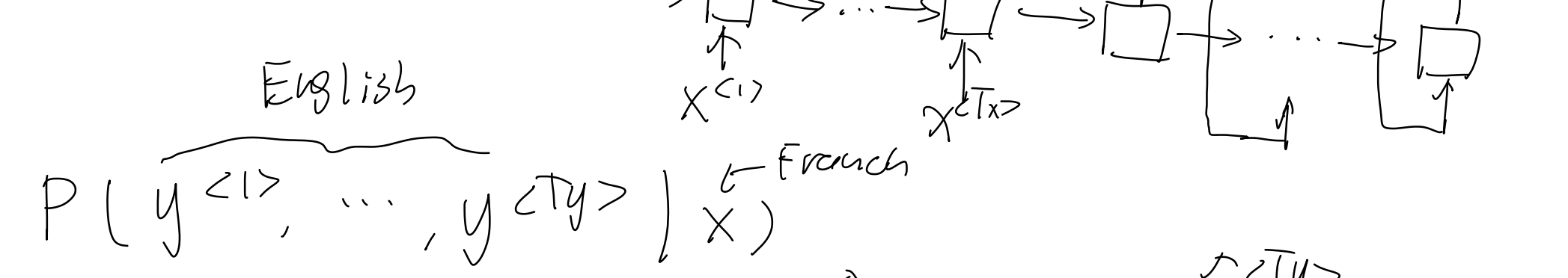


Picking the Most Likely Sentence.

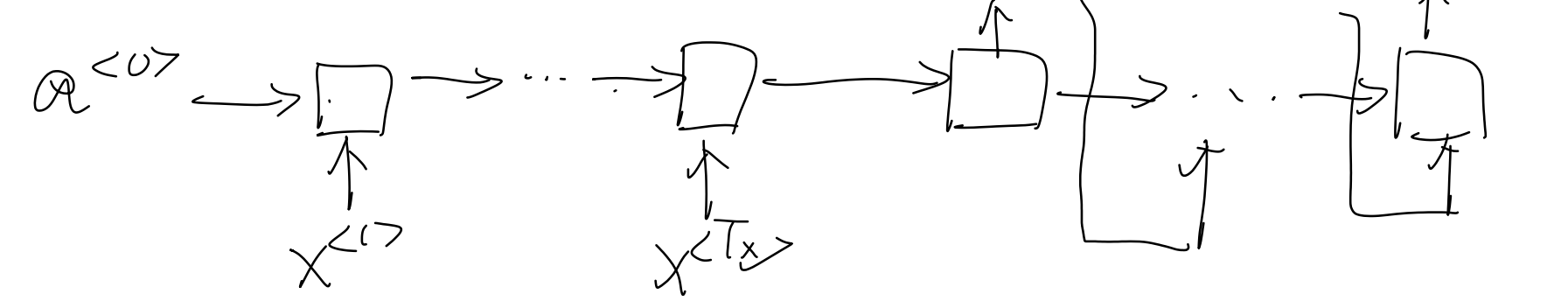
$P(y^{<1>}, \dots, y^{<Ty>})$



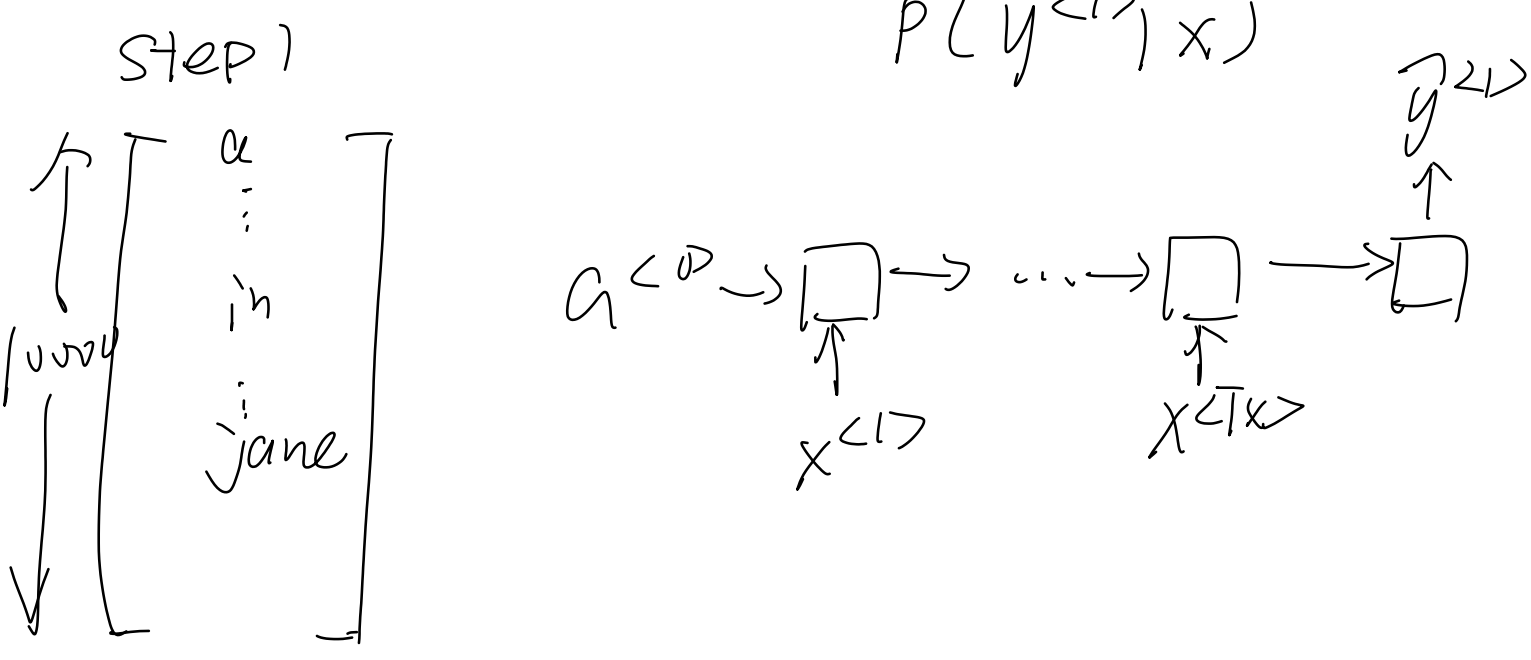
Machine translation:



冗余搜索



Beam Search



$B=3$ (beam width)

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

“一次考虑多个值的搜索”

Refinements to Beam Search

长度规范:

$$\arg \max_y \prod_{t=1}^{Ty} P(y^{<t>} | x, y^{<1>}, \dots, y^{<t-1>})$$
$$\arg \max_y \sum_{t=1}^{Ty} \log P(y^{<t>} | x, y^{<1>}, \dots, y^{<t-1>})$$
$$\frac{1}{Ty} \sum_{t=1}^{Ty} \log P(y^{<t>} | x, y^{<1>}, \dots, y^{<t-1>})$$

large B: better result, slower
small B: worse result, faster

Beam width B ?
 $1 \rightarrow 3 \rightarrow 10, 100, 1000, \rightarrow 2000$
large B: better result, slower
small B: worse result, faster

Error Analysis in Beam Search

RNN
Beam Search
Case 1: $P(y^* | x) > P(\hat{y} | x)$
Case 2: $P(y^* | x) \leq P(\hat{y} | x)$

Belu 得分

对于句子中单词出现的次数相应给分,
 $\exp(\frac{1}{4} \sum_{n=1}^N p_n)$
BP 惩罚因子来调节:

Attention Model Intuition