

$$P(x_1, x_2, \dots, x_n, y_1, \dots, y_n) = p(y_1) * p(y_2|y_1) \cdots * p(y_n|y_{n-1}, y_{n-2}, \dots, y_1) * p(x_1|y_n, y_{n-1}, \dots, y_1) * p(x_n|x_{n-1}, \dots, x_1, y_n, y_{n-1}, \dots, y_1)$$

$\prod_{i=1}^n q(y_i|y_{i-2}, y_{i-1})$ equivalente a una LM de trigramas sobre las etiquetas,
 Es la probabilidad Prior de nuestras etiquetas

$\prod_{i=1}^n e(x_i|y_i)$ $e(\text{the}|\text{DT})$ Probabilidad de observar la palabra
 "the" dado que la etiqueta es "DT"

$\mathcal{V} = \{\text{the, dog, cat, a, barks, ..}\}$

$\mathcal{S} = \{\text{DT, NN, VB, P, ADV, ..}\}$

Sequence Classification -> Input (El perro ladra), output (mascotas)

Sequence Labeling or Tagging -> Input(El perro ladra), output(DET, NOUN, VERB)

Sequence to Sequence -> Input (El perro ladra mucho), out(The dog barks a lot)

Forma de la HMM

$$p(x_1 \dots x_n, y_1 \dots y_{n+1}) = \prod_{i=1}^{n+1} q(y_i | y_{i-2}, y_{i-1}) \prod_{i=1}^n e(x_i | y_i)$$

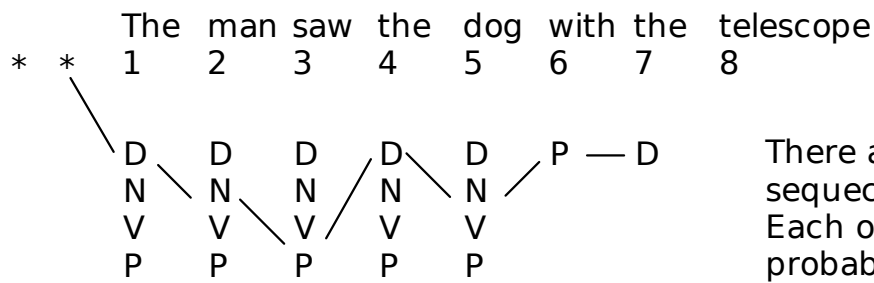
$$y_{n+1} = \text{STOP}$$

Problem: for an input $x_1 \dots x_n$, find

$$\arg \max_{y_1 \dots y_{n+1}} p(x_1 \dots x_n, y_1 \dots y_{n+1})$$

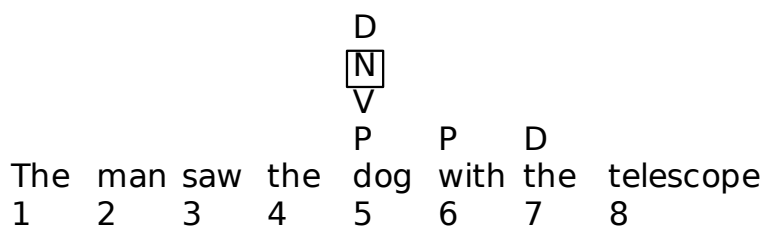
Because of the Limited Horizon of the HMM, we don't need to keep a complete record of how we arrived at a certain state.

$$S = \{D, N, P, V\}$$



There are many possible sequences of tags. Each of these will have a probability calculated from parameters q and e .

$\Pi(7, P, D) =$ This is the maximum probability of any of those tag sequences ending in P D at position 7, the path represents the sequence with the maximum probability.



$$\mathcal{S}_5 = \mathcal{S} = \{D, N, V, P\}$$

$$\Pi(7, P, D) = \max_{w \in \mathcal{S}_5} (\Pi(6, w, P) \times q(D|w, P) \times e(\text{the}|D))$$

If we think in any tag sequence ending in tags P and D in position 7, it has to include some tag at position 5. We are basically searching the tag that maximizes the probability at position 5.