Mod 3 - CO3 - 2024-08-01 - Apply various POS Tagging Techniques

01 August 2024 12:21

-> Simplified Formula :

$$\hat{T} = \arg\max_T \prod_i P(w_i | t_i) \times P(t_i | t_{i-1})$$

-> Tag Transition Probabilities : $P(t_i | t_{i-1})$

$$P(t_i | t_{i-1}) = \frac{c(t_{i-1}, t_i)}{c(t_{i-1})}$$

Eg : $P(NN | DT) = \frac{c(DT, NN)}{c(DT)}$

-> Word Likelihood Probabilities : $P(w_i | t_i)$

$$P(w_i | t_i) = \frac{c(t_i, w_i)}{c(t_i)}$$

Eg : $P(is | VBZ) = \frac{c(VBZ, is)}{c(VBZ)}$

-> Disambiguating (Brown Corpus) :

-> Thief is expected to race tomorrow.

-> Thief is expected to race tomorrow.

-> **Sentence Formation:**

a) Thief | NNP -> is | VBZ -> expected | VBN -> to | TO -> race | VB -> tomorrow | RB.

b) Thief | NNP -> is | VBZ -> expected | VBN -> to | TO -> race | NN -> tomorrow | RB.

a) $P(VB \mid TO) \times P(RB \mid VB) \times P(Race \mid VB)$

b) $P(NN \mid TO) \times P(RB \mid NN) \times P(Race \mid NN)$

Graph Style System:

$$
\begin{array}{cccccccc}
 & \text{Thief} & \text{is} & \text{expected} & \text{to} & \text{race} & \text{tomorrow.} & \\
 & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \\
S \rightarrow & NNP \rightarrow & VBZ \rightarrow & VBN \rightarrow & TO \rightarrow & VB \rightarrow & RB \rightarrow & E
\end{array}
$$

-> **Hidden Markov Model:**

-> Tag Transitioning Probabilities : $P(t_i \mid t_{i-1})$

-> Word likelihood Probabilities : (Emissions)

→ Word Likelihood Probabilities : (Emissions)

: $P(w_i | t_i)$

→ First-Order Markov Model :

→ Weather Example : Sunny, Rainy, Foggy.

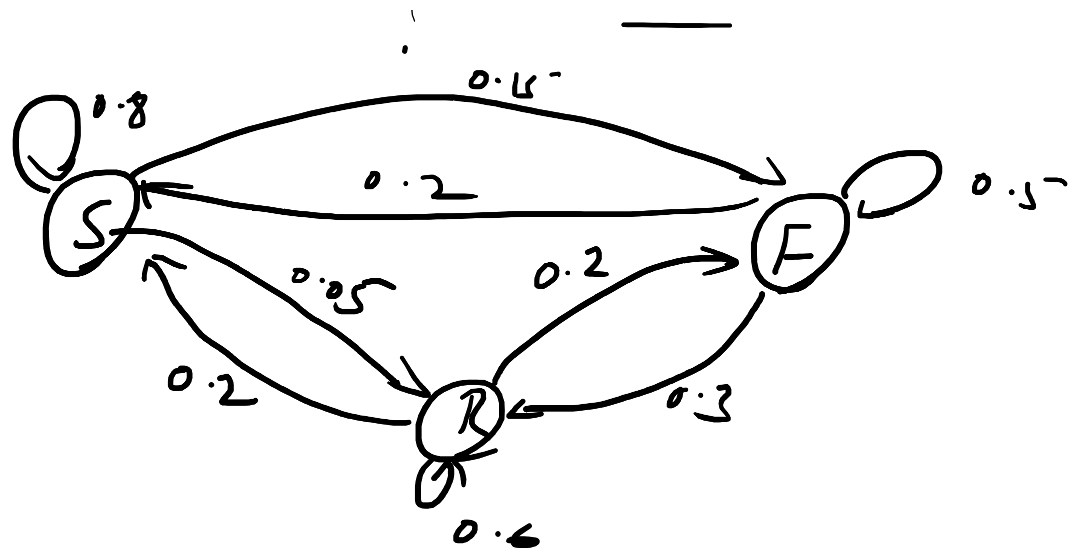
q_n : Variable denoting the weather on the n^{th} day.

$\therefore P(q_n \mid q_{n-1}, q_{n-2} \cdots\cdots q_1)$

$\Rightarrow$ F-O-M-M $= P(q_n | q_{n-1})$

	Tomorrow		
Today	S	R	F
S	0.8	0.05	0.15
R	0.2	0.6	0.2
F	0.2	0.3	0.5

0.15



→ Given that today the weather is sunny, what is the probability that tomorrow is sunny & day after tomorrow is <u>rainy</u>.

$$= P(q_{n+1} = \text{Sunny} \mid q_n = \text{Sunny}) \times P(q_{n+2} = \text{Rainy} \mid q_n = \text{Sunny}, q_{n+1} = \text{Sunny})$$

$$= P(q_{n+1} = \text{Sunny} \mid q_n = \text{Sunny}) \times P(q_{n+2} = \text{Rainy} \mid q_{n+1} = \text{Sunny})$$

$$= 0.8 \times 0.05$$

$$= 0.04$$

→ the probability of raining day after tomorrow is 4 % iff tomorrow is sunny.