

CS 7650 - Lec 7 scribe notes

System #1
 I-PER I-PER 0 ...
 Barack Obama will ...

System #2
 B-LOC 0-0-0 B-ORG ...
 Hangzhou ... G20 ...

Sequential Models
 $x = (x_1, x_2, \dots, x_n)$ $y = (y_1, y_2, \dots, y_n)$
 $\hat{y} = \underset{y \in \mathcal{Y}}{\operatorname{argmax}} \phi(x, y)$
 $\phi(x, y) = \sum_{j=1}^n w_j f(x^{(j)}, y^{(j)})$
 Scoring func. on pair of refer

$\mathcal{Y}^n \times \mathcal{Y}^n \rightarrow \mathbb{R}$

HMM $\hat{y} = \underset{y}{\operatorname{argmax}} P(y|x)$

CRT $\hat{y} = \underset{y}{\operatorname{argmax}} P(y|x) = \underset{y}{\operatorname{argmax}} \frac{\exp(\phi(x, y))}{\sum_{y' \in \mathcal{Y}} \exp(\phi(x, y'))}$

LR training

$$L(x, y^*) = \sum_{j=1}^n \log P(y^{(j)*} | x^{(j)})$$

$$= \sum_{j=1}^n (w^T f(x^{(j)}, y^{(j)*}) - \log \sum_{y \in \mathcal{Y}} \exp(w^T f(x^{(j)}, y)))$$

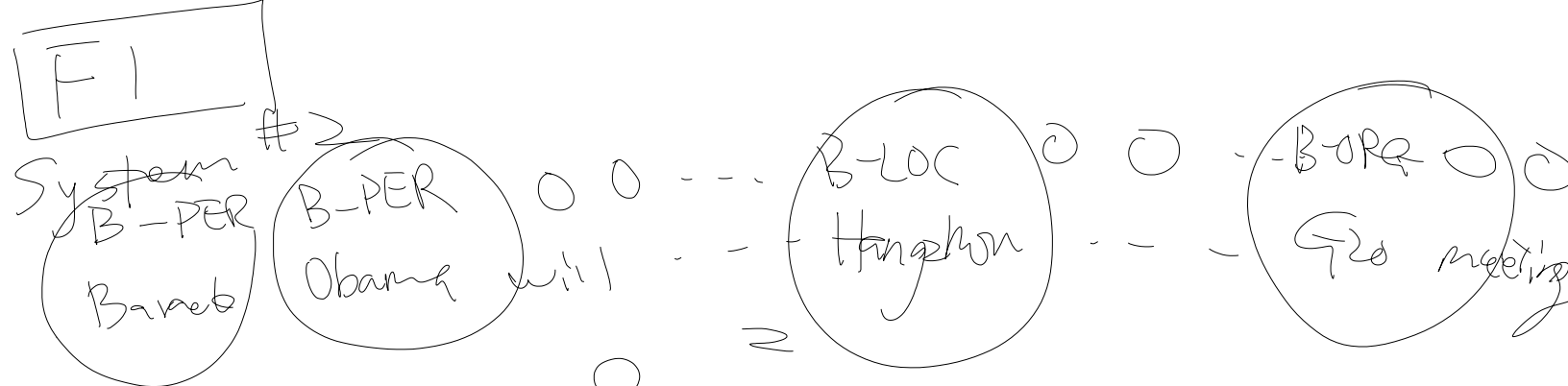
$$\frac{\partial}{\partial w} L(x, y^*) = f(x^{(j)}, y^{(j)*}) - \sum_{y \in \mathcal{Y}} f(x^{(j)}, y) P(y | x^{(j)})$$

CRT training

$$L(x, y^*) = \sum_{j=1}^n \log P(y^{(j)*} | x^{(j)})$$

$$\frac{\partial}{\partial w} L(x, y^*) = f(x^{(j)}, y^{(j)*}) - \sum_{y \in \mathcal{Y}_n} f(x^{(j)}, y) P(y | x^{(j)})$$

exponential!



precision = $\frac{1}{4}$

recall = $\frac{2}{3}$