Max Likelihood Estimation for CRFs

Parameter Estimation

$$P_\theta(V|x) = \frac{1}{Z_x(\theta)} \tilde{P}_\theta(x, y)$$

$$Z_x(\theta) = \sum_y \tilde{P}_\theta(x, y)$$

M instances

$$D = \left\{ (x[m], y[m]) \right\}_{m=1}^{M}$$

Data is now a set of pairs

log conditional likelihood

$$\ell_{y|x}(\theta : D) = \sum_{m=1}^{M} \ln P_\theta\big(y[m] \mid x[m], \theta\big)$$

$$\ell_{y|x}\big(\theta : (x[m], y[m])\big) = \left(\sum_i \theta_i f_i\big(x[m], y[m]\big) \right) - \ln Z_{x[m]}(\theta)$$