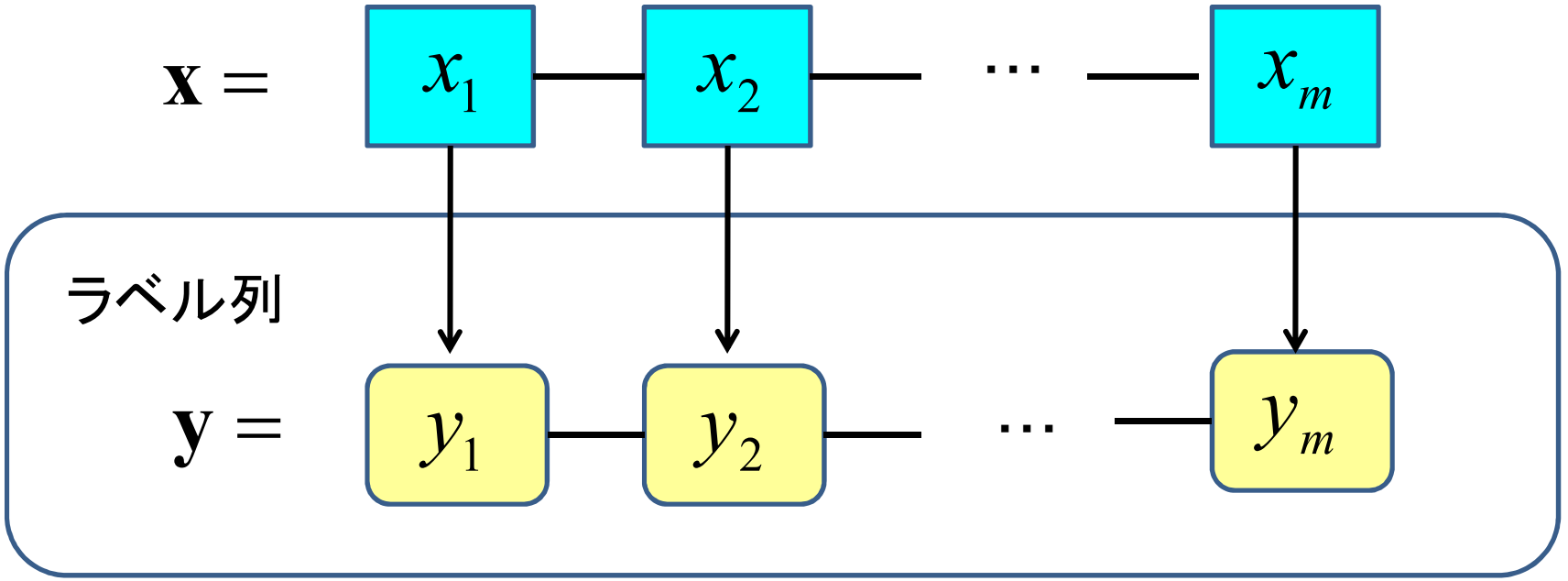
言語処理のための機械学習入門

5.4 条件付き確率場
5.5 チャンキングへの適用の仕方

新納浩幸

系列ラベリング問題

観測系列



$y_i \in Y$（ラベル集合）

観測系列からラベル列を推定する問題

品詞タガー

NLP における系列ラベリング問題の代表例

$\mathbf{x} =$ This is a pen.

↓ ↓ ↓ ↓

$\mathbf{y} =$ PN V D N

対数線形モデル

$$P(\mathbf{y} \mid \mathbf{x}) = \frac{1}{Z_{x,w}} \exp\left(\mathbf{w} \cdot \phi(\mathbf{x}, \mathbf{y})\right)$$

$$Z_{x,w} = \sum_{\mathbf{y}} \exp\left(\mathbf{w} \cdot \phi(\mathbf{x}, \mathbf{y})\right)$$

CRF

$$\phi(\mathbf{x},\mathbf{y}) = \left(\phi_1(\mathbf{x},\mathbf{y}), \phi_2(\mathbf{x},\mathbf{y}), \cdots, \phi_n(\mathbf{x},\mathbf{y})\right)$$

$$\phi_k(\mathbf{x},\mathbf{y}) = \sum_t \phi_k(\mathbf{x}, y_t, y_{t-1})$$

$$\mathbf{w}\cdot\phi(\mathbf{x},\mathbf{y}) = \sum_t \mathbf{w}\cdot\phi(\mathbf{x}, y_t, y_{t-1})$$

識別（ラベリング）

$$\mathbf{y}^* = \arg\max_y \frac{1}{Z_{x,w}} \exp\left(\mathbf{w} \cdot \phi(\mathbf{x}, \mathbf{y})\right)$$

$$= \arg\max_y \mathbf{w} \cdot \phi(\mathbf{x}, \mathbf{y})$$

$$= \arg\max_y \sum_t \mathbf{w} \cdot \phi(\mathbf{x}, y_t, y_{t-1})$$

ビタビアルゴリズムで解ける

学習（1）

重みベクトル w が学習の対象

最急勾配法で求める

$$\mathbf{w}^{new} = \mathbf{w}^{old} + \varepsilon \nabla_w L(\mathbf{w}^{old})$$

$$\nabla_w L(\mathbf{w}) = \sum_D \left(\phi(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}) - \sum_y P(\mathbf{y} \mid \mathbf{x}^{(i)}) \phi(\mathbf{x}^{(i)}, \mathbf{y}) \right)$$

計算困難

学習（2）

$$\sum_{y} P(\mathbf{y} \mid \mathbf{x}) \phi(\mathbf{x}, \mathbf{y}) = \sum_{y} P(\mathbf{y} \mid \mathbf{x}) \sum_{t} \phi(\mathbf{x}, y_t, y_{t-1})$$

$$= \sum_{t} \sum_{y_t, y_{t-1}} P(y_{t-1}, y_t \mid \mathbf{x}) \phi(\mathbf{x}, y_t, y_{t-1})$$

$P(y_{t-1}, y_t \mid \mathbf{x})$　が計算できれば求まる

学習（3）

$$P(y_{t-1}, y_t \mid \mathbf{x}) = \frac{1}{Z_{x,w}} \varphi_t(y_t, y_{t-1}) \alpha(y_{t-1}, t-1) \beta(y_t, t)$$

s.t.

$$\varphi_t(y_t, y_{t-1}) = \exp(\mathbf{w}, \phi(\mathbf{x}, y_t, y_{t-1}))$$

$$\alpha(y_t, t) = \sum_{y_{0:t-1}} \prod_{t'=1}^{t} \varphi_{t'}(y_{t'}, y_{t'-1}) \qquad \alpha(y_0, 0) = 1$$

$$\beta(y_t, t) = \sum_{y_{t+1:T+1}} \prod_{t'=t+1}^{T+1} \varphi_{t'}(y_{t'}, y_{t'-1})$$

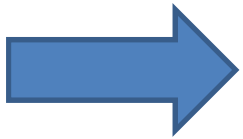
$$\beta(y_{T+1}, T+1) = 1$$

前向き・後ろ向きアルゴリズム

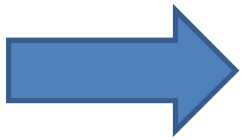
$Z_{x,w} = \sum_{y_T} \alpha(y_T, T)$　なので

求め方は簡単

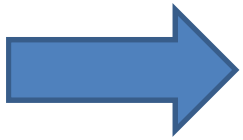
$\alpha(y_t, t)$　と　$\beta(y_t, t)$　が求まれば



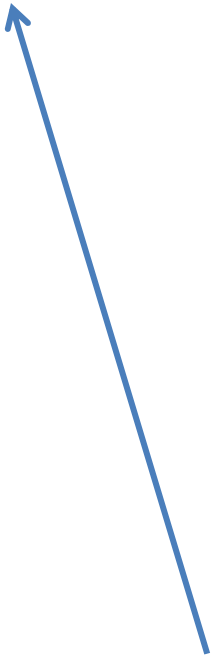
$P(y_{t-1}, y_t \mid \mathbf{x})$　が求まる



$\sum_y P(\mathbf{y} \mid \mathbf{x}) \phi(\mathbf{x}, \mathbf{y})$　が求まる



$\mathbf{w}$　が求まる（学習完了）

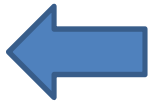


この形にもってゆく手法が

前向き・後ろ向きアルゴリズム

チャンキング

言語表現の意味的あるいは文法的にまとまった部分を発見するタスク

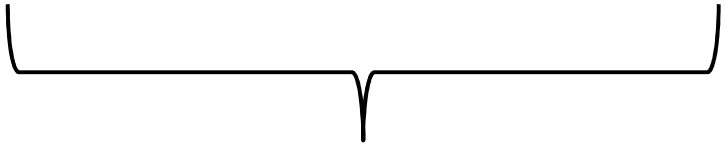


系列ラベリング問題の応用

IOB タグ

Suddenly, the tall German guy talked to me.

O B I I I O O O



抽出箇所（人を表す表現）

CRF のツール

CRF++ ： http://crfpp.sourceforge.net/

訓練データと素性を記したテンプレートを
用意すれば、簡単に利用可能

系列ラベリング問題にどんどん利用してみよう！