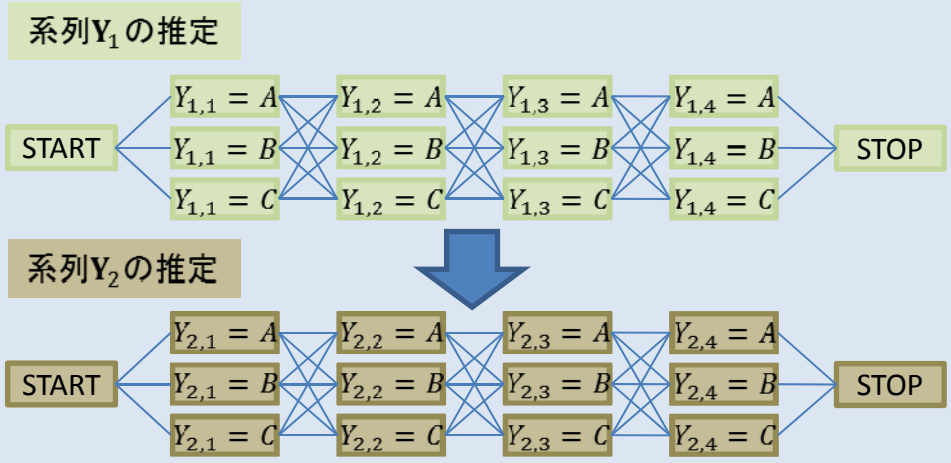
4-16　系列ラベリングの多層化

東　藍　　松本　裕治
奈良先端科学技術大学院大学

研究の背景

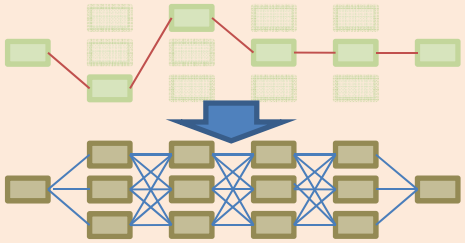
系列ラベリングが様々なタスクで顕著な成果を挙げている
→自然言語処理　　→音声認識　　→遺伝子解析

複数の系列ラベリングの段階適用の必要性
（e.g. 入力自然言語文に対する品詞付与と基本句同定）



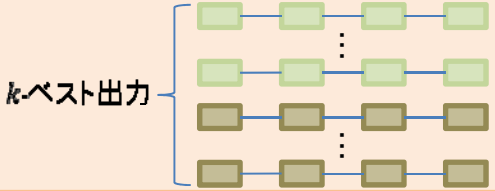
既存手法と問題点

1-ベストパイプライン



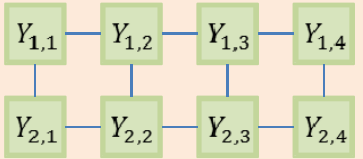
実装コストが低いが
解析誤りが後続で伝搬･拡大

k-ベスト，リランキング



解析誤りの伝搬を減じるが
同時学習は考慮されない

グラフィカルモデルによる同時学習



同時学習可能だが
exact inference が不可能
近似手法が必要

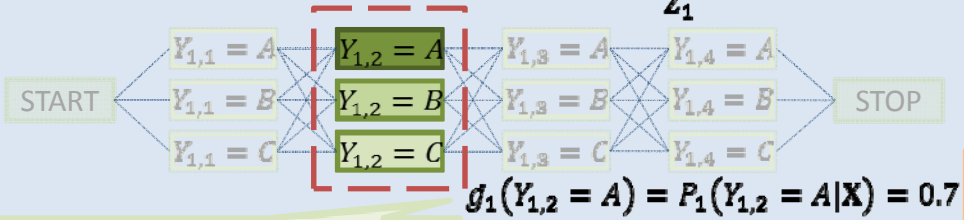
研究の目的

解析誤りの伝搬を防ぎ，かつ動的計画法による exact inference の利点を維持した同時学習手法の提案

採用する既存手法（素性の周辺化）

[Bunescu, 2008] のアイデアを採用
各段階に確率モデルを仮定

$$P_1(\mathbf{Y}_1|\mathbf{X};\boldsymbol{\theta}_1)=\frac{1}{Z_1}\exp(\boldsymbol{\theta}_1\cdot\mathbf{F}_1(\mathbf{X},\mathbf{Y}_1))$$

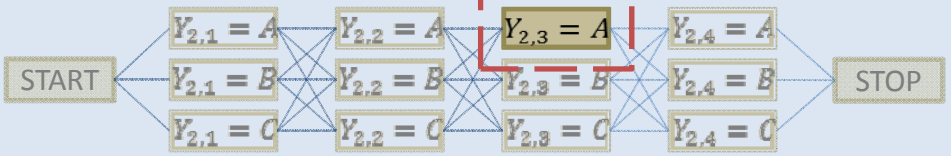


$\mathbf{Y}_1$ の各部分に関する情報を対応する周辺確率で重み付け

$$g_1(Y_{1,2}=A)=P_1(Y_{1,2}=A|\mathbf{X})=0.7$$
$$g_1(Y_{1,2}=B)=P_1(Y_{1,2}=B|\mathbf{X})=0.2$$
$$g_1(Y_{1,2}=C)=P_1(Y_{1,2}=C|\mathbf{X})=0.1$$

$\mathbf{Y}_1$ の各部分の情報に依存する $\mathbf{Y}_2$ の素性も対応して重み付け

$$f_2(Y_{1,2}=A\wedge Y_{2,3}=A)=0.7$$
$$f_2(Y_{1,2}=B\wedge Y_{2,3}=A)=0.2$$
$$f_2(Y_{1,2}=C\wedge Y_{2,3}=A)=0.1$$



$$P_2(\mathbf{Y}_2|\mathbf{X};\boldsymbol{\theta}_2)=\frac{1}{Z_2}\exp(\boldsymbol{\theta}_2\cdot\mathbf{F}_2(\mathbf{X},\bar{g}_1,\mathbf{Y}_2))$$

本研究の手法

パラメタ$\boldsymbol{\theta}_1$により周辺確率が変化
⇒$\bar{g}_1$およびf_2の値も変化　⇒ $\boldsymbol{\theta}_1$に依存した関数とみなす

$$\begin{cases} g_1(Y_{1,2}=A;\boldsymbol{\theta}_1) \\ g_1(Y_{1,2}=B;\boldsymbol{\theta}_1) \\ g_1(Y_{1,2}=C;\boldsymbol{\theta}_1) \\ \vdots \end{cases} \qquad \begin{cases} f_2(Y_{1,2}=A\wedge Y_{2,3}=A;\boldsymbol{\theta}_1) \\ f_2(Y_{1,2}=B\wedge Y_{2,3}=A;\boldsymbol{\theta}_1) \\ f_2(Y_{1,2}=C\wedge Y_{2,3}=A;\boldsymbol{\theta}_1) \\ \vdots \end{cases}$$

後続の段階のモデルも$\boldsymbol{\theta}_1$に間接的に
依存しているとみなして最適化

$$P_2(\mathbf{Y}_2|\mathbf{X};\boldsymbol{\theta}_2,\boldsymbol{\theta}_1)=\frac{1}{Z_2}\exp(\boldsymbol{\theta}_2\cdot\mathbf{F}_2(\mathbf{X},\bar{g}_1(\boldsymbol{\theta}_1),\mathbf{Y}_2))$$

適当な目標関数（e.g. P_1, P_2の対数尤度関数の和）で同時最適化
最適化の際にはパラメタ勾配を
合成関数の微分法で分解して計算

$$\frac{\partial P_2}{\partial \boldsymbol{\theta}_1}=\frac{\partial P_2}{\partial f_2}\cdot\frac{\partial f_2}{\partial \bar{g}_1}\cdot\frac{\partial \bar{g}_1}{\partial \boldsymbol{\theta}_1}$$ ⇨ 誤差逆伝搬法に類似

パラメタ勾配の各因子は動的計画法で exact に計算可能

実験

英文に対する品詞タグ付と基本句同定の2つの系列ラベリングタスクに提案手法を適用し提案手法の有効性を検証

	1-ベストパイプライン	周辺化素性を利用した1-ベストパイプライン（[Bunescu, 2008] 相当）	提案手法（周辺化素性+同時最適化）
品詞タグ付	0.956	(0.956)	**0.958**
基本句同定	0.921	0.927	**0.931**

実験により提案手法の有効性を確認. 今後は既存手法との精緻な比較や他タスクへの適用が課題.