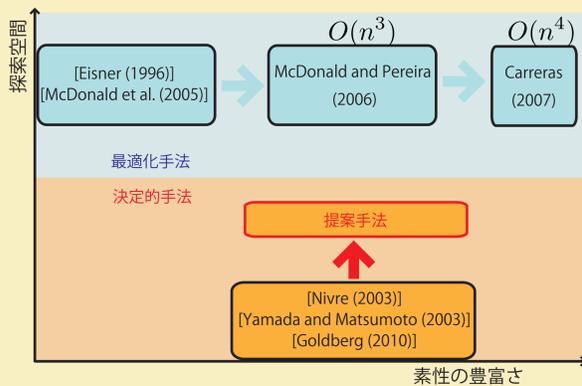


東京大学大学院 情報理工学系研究科  
北川 浩太郎, 田中久美子

## 背景

最適化を行うアプローチ  
— Graph-based parsers  
決定的なアプローチ  
— Transition-based parsers  
複数の Parser の組み合わせ



## 決定的手法

コーパスから学習した分類器によって、解析アクションを反復的に選択  
解析アクションは単語間の関係といった局所的な構造を決定

- Nivre (2003): インクリメンタルに解析
- Yamada, Matsumoto (2003): ボトムアップの制約のもと解析

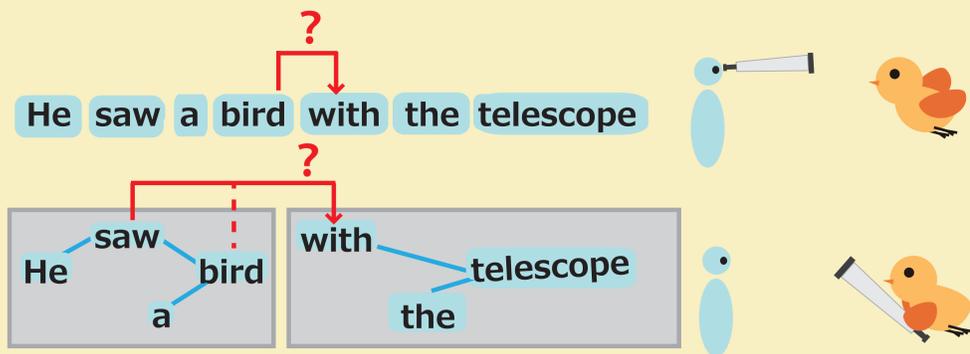
## 問題点

1. 解析アクションの決定には 2 単語間の関係以上に多くの構造を判断する必要がある。しかし、限られた素性から遠い係り関係を判断するのは難しい。
  - Nivre の手法: 語の head/dependent が左側にあるかどうか
  - Yamada の手法: 語の dependent が全て解析済みかどうか
2. 複数の係り関係から最適なものを選ぶことができない。

## Tree-based Model

解析アクションを決定する対象として、語のかわりに木を扱う。  
決定的手法に探索を加え、隣接する二つの木構造間の最適な係り関係を決定する。  
探索は二つの木構造の間に作成され得る係り関係全て

- 解析アクションを決定するために必要な係り関係を探索
- 係り先候補となる複数の語を考慮できる



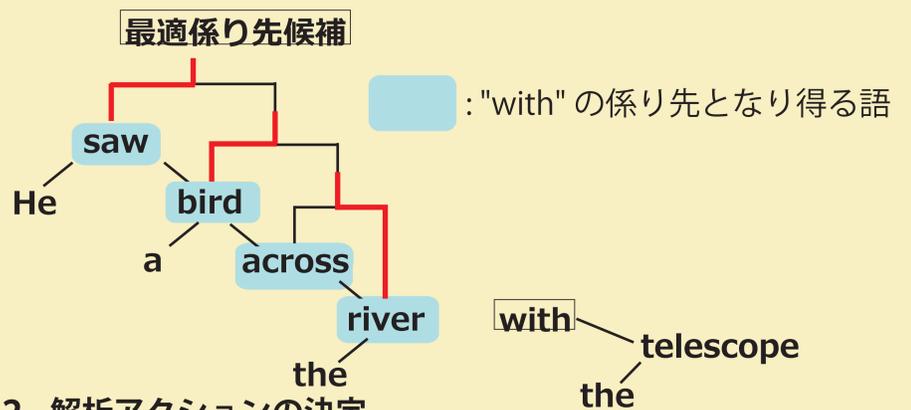
## 手法

隣接する二つの木構造間に作成され得る係り関係は複数考えられる。  
そのため 2つのステップを経て、最適な係り関係を決定する。

### 1. 最適係り先候補の選択

下図のように二つの木構造が隣接しているとき、with の head として最も相応しい語を、学習器を用いて選択する。

- 2 値分類器が、二つの語のどちらの方が係り先らしいかを選択
- より係り先らしい語が勝ち上がるトーナメントにより決定



### 2. 解析アクションの決定

上図における二つの隣接した木の関係は、① with の head が左の木に存在する ② saw の head が右の木に存在する ③二つの木の間に係り関係は存在しない の 3通りである。

①②の場合にどのような係り関係かは先ほどのステップで求めたため、学習器を用いてこの 3 値分類問題を解くことで、係り関係を決定することができる。

## 素性

それぞれ以下の語の品詞タグと表層形を利用した。  
トーナメント: head を探す語 (木の root)、head 候補の 2 語、それぞれの係り関係にある語、後続する 3 語  
解析アクションの決定: 木の root、最適 head 候補、それぞれの係り関係にある語、後続する 3 語



## 実験

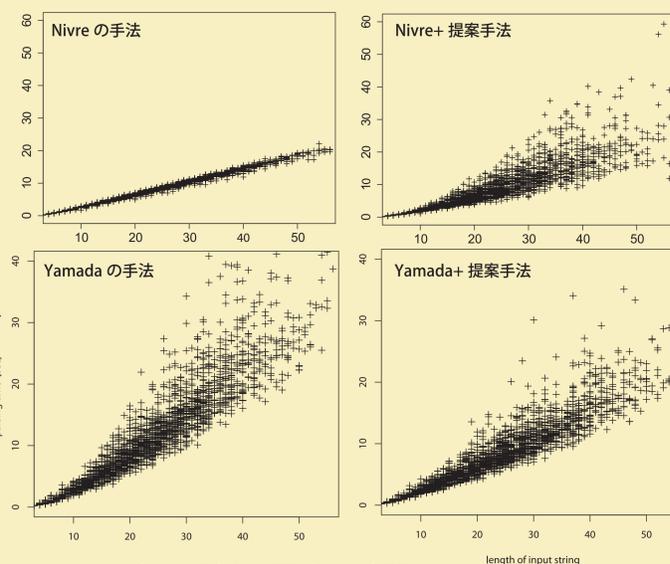
データセット PennTreebank (WSJ)  
Section 2~21 for training data  
Section 23 for test data

### 評価関数

Accuracy: 正しい係り先を決定できた語の割合  
Complete Match: 全ての係り関係を正しく決定できた文の割合

### 学習器

SVM (3 次の多項式カーネル)



一文の解析に要する時間

	AC	CM	計算量	平均解析時間 [s]
Nivre の手法	91.1	39.7	$O(n)$	7.9
Nivre + 提案手法	91.3	41.7	$O(n^2)$	8.9
Yamada の手法	90.2	38.0	$O(n^2)$	11.5
Yamada + 提案手法	90.7	38.6	$O(n^2)$	8.3

## 結果

- 提案手法によって解析精度が若干上昇した。
- Yamada の手法の場合は探索を加えたにも関わらず、速度に改善が見られた。これは、ボトムアップという解析順序の制約を除くことで、解析アクションの回数が減少したためと考えられる。

## 参考文献

Masakazu Iwatake, Masayuki Asahara, and Yuji Matsumoto. 2008. Japanese dependency parsing using a tournament model. In *Proceedings of the 22nd COLING*, pp. 361–368.  
Ryan McDonald and Fernando Pereira. 2006. Online learning of approximate dependency parsing algorithms. In *Proceedings of the 11th EACL*, pp. 81–88.  
Joakim Nivre. 2003. An efficient algorithm for projective dependency parsing. In *Proceedings of the 8th IWPT*, pp. 149–160.  
Hiroyasu Yamada and Yuji Matsumoto. 2003. Statistical dependency analysis with support vector machines. In *Proceedings of the 8th IWPT*, pp. 195–206.  
Yoav Goldberg and Michael Elhadad. 2010. An Efficient Algorithm for Easy-First Non-Directional Dependency Parsing. In *Proceedings of the 11th NAACL*.