Recursive Networks for NER

2018/09/11

Peng-Hsuan Li

Outline

- Named Entity Recognition
  - Task
  - Features
  - Related Work
- Leveraging Linguistic Structures for NER

Named Entities

- CoNLL-2003
  - PER, LOC, ORG, MISC

$$\textit{The } \underbrace{\textit{defense}}_{\textit{ORG}} \textit{ secretary } \underbrace{\textit{Donald Rumsfeld}}_{\textit{PER}}$$

- OntoNotes 5.0
  - person, NORP, facility, organization, GPE, location, product, event, work-of-art, law, language
  - date, time, percent, money, quantity, ordinal, cardinal

Gazetteer Features

- Senna
  - PER
  - LOC
  - ORG
  - MISC

Word Features

- Embedding
  - English -> 840B (common crawl)
  - Chinese -> CNA (gigaword) + ASBC (sinica)

- Uppercase
- Upper-initial
- Lowercase
- Mixed

Character Features

- Embedding
  - English -> Random initialization
  - Chinese -> CNA (gigaword) + ASBC (sinica)

- Uppercase
- Lowercase
- Digit

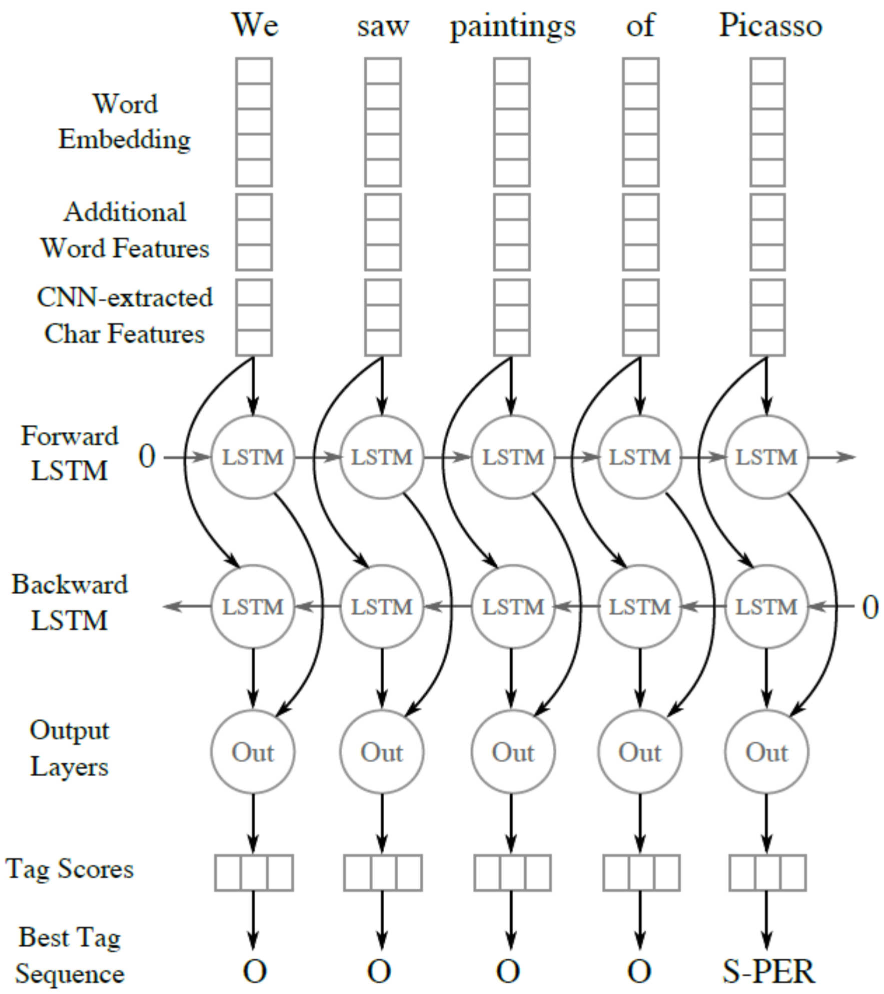
Sequence Tagging

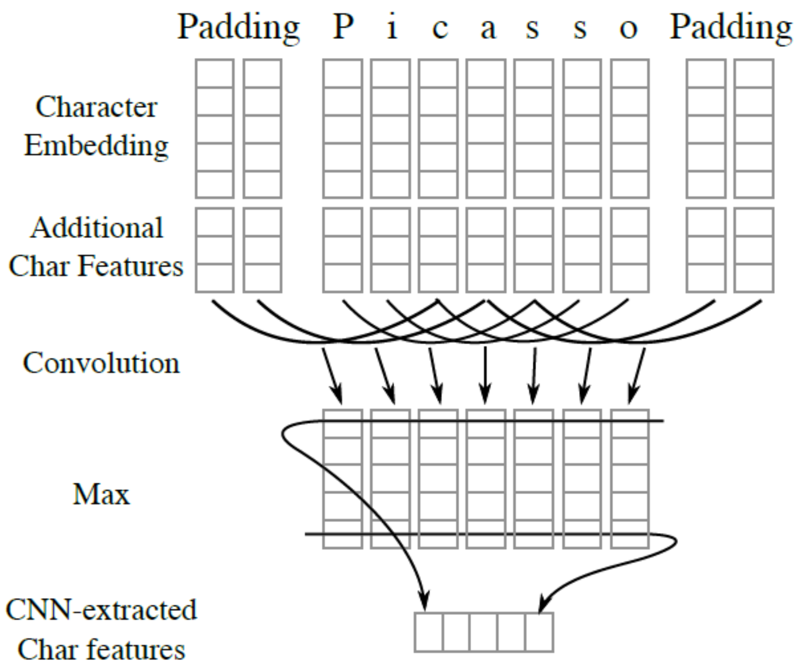
$$\textit{The } \underbrace{\textit{defense}}_{\textit{ORG}} \textit{ secretary } \underbrace{\textit{Donald Rumsfeld}}_{\textit{PER}}$$

$$\begin{tikzcd}
\textit{The} \arrow[r] \arrow[d] & \textit{defense} \arrow[r] \arrow[d] & \textit{secretary} \arrow[r] \arrow[d] & \textit{Donald} \arrow[r] \arrow[d] & \textit{Rumsfeld} \arrow[d] \\
\textit{O} & \textit{S-ORG} & \textit{O} & \textit{B-PER} & \textit{E-PER}
\end{tikzcd}$$

Chunk Labels	
B	*(begin)*
I	*(inside)*
O	*(outside)*
E	*(end)*
S	*(single)*

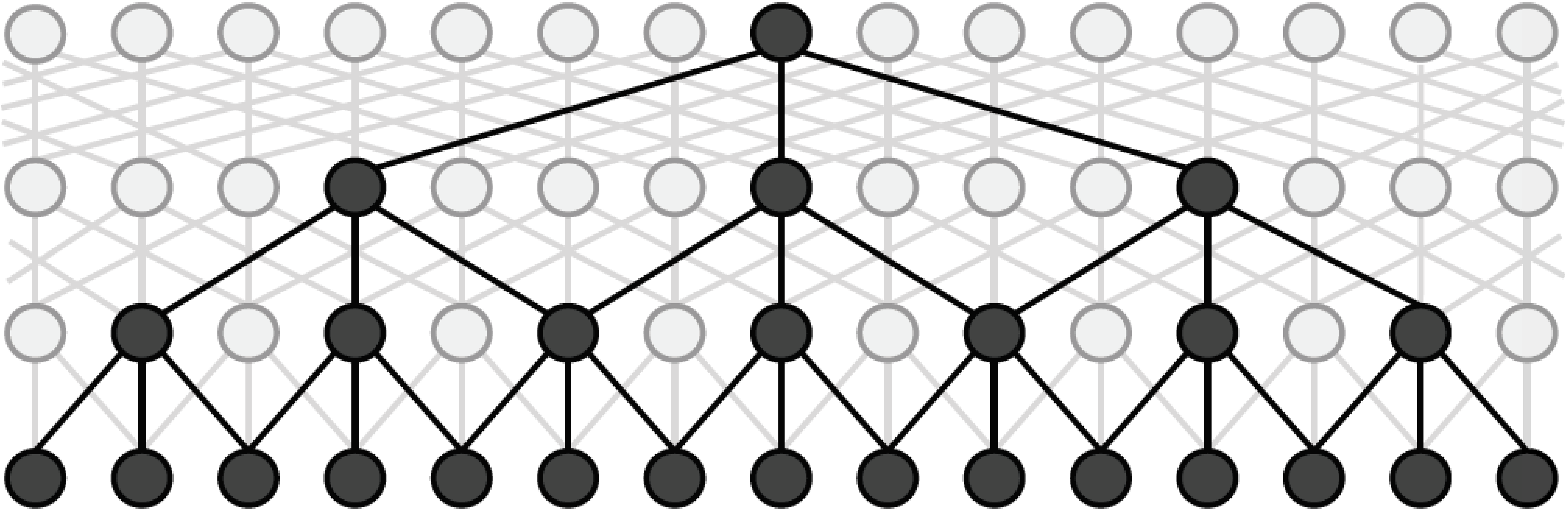
Bi-LSTM for Sequence Tagging





Jason Chiu and Eric Nichols. 2016. Named Entity Recognition with Bidirectional LSTM-CNNs. Transactions of ACL.

Dilated CNN for Sequence Tagging



Emma Strubell, Patrick Verga, David Belanger, and Andrew McCallum. 2017. Fast and Accurate Entity Recognition with Iterated Dilated Convolutions. In Proceedings of EMNLP.

Results of Sequence Tagging

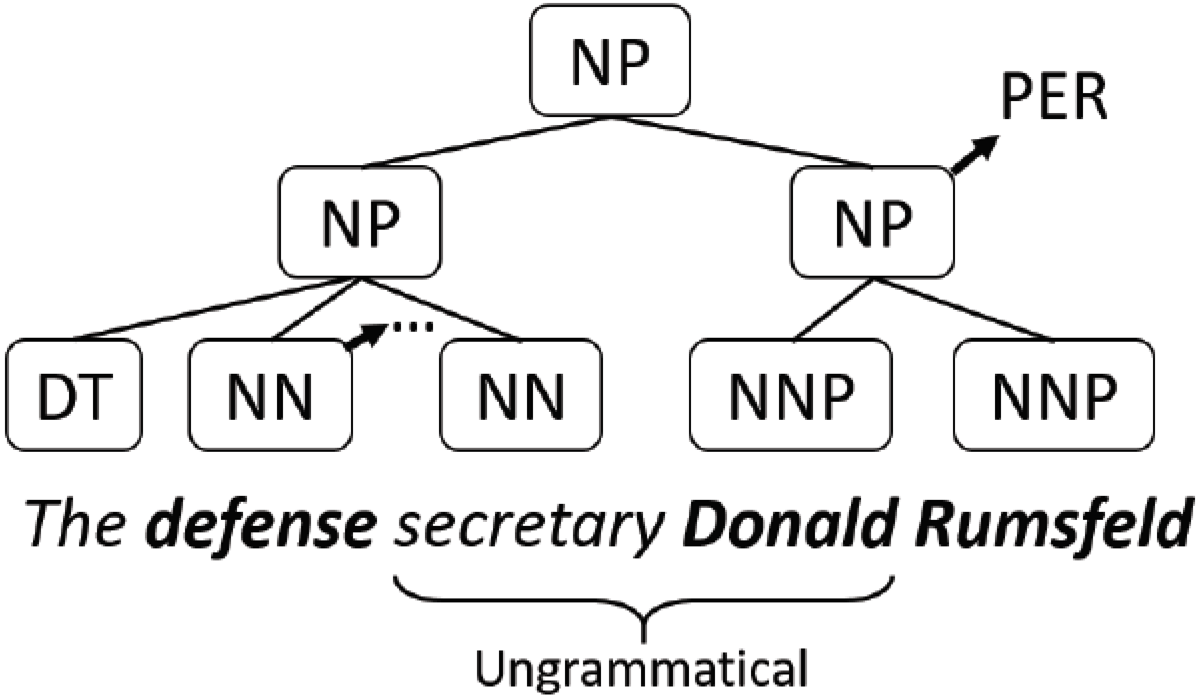
Model	Sources	CoNLL-2003	OntoNotes 5.0
BLSTM	Huang et al., 2015 Chiu and Nichols, 2016 Ma and Hovy, 2016 Lample et al., 2016 Strubell et al., 2017	90.67	83.76
BLSTM-CRF		90.94	**86.99**
BLSTM-CNN		90.98	-
BLSTM-CNN-CRF		**91.21**	-
Deep BLSTM		-	86.19
Deep-BLSTM-CNN		-	86.41
ID-CNN-CRF	Strubell et al., 2017	90.65	86.84

Outline

- Named Entity Recognition

- Leveraging Linguistic Structures for NER
  - Joint parsing and NER
  - Tree-LSTM for NER
  - Mitigating inconsistencies between parsing and NER

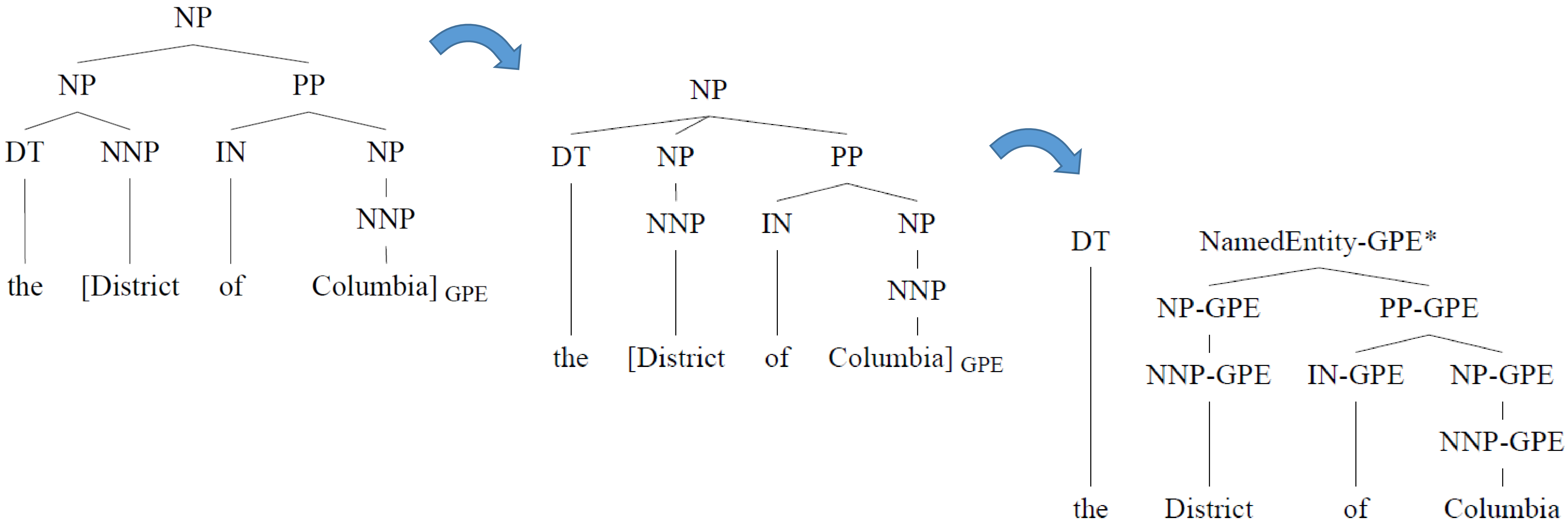
Constituent Prediction



Constituent → A plausible NE candidate

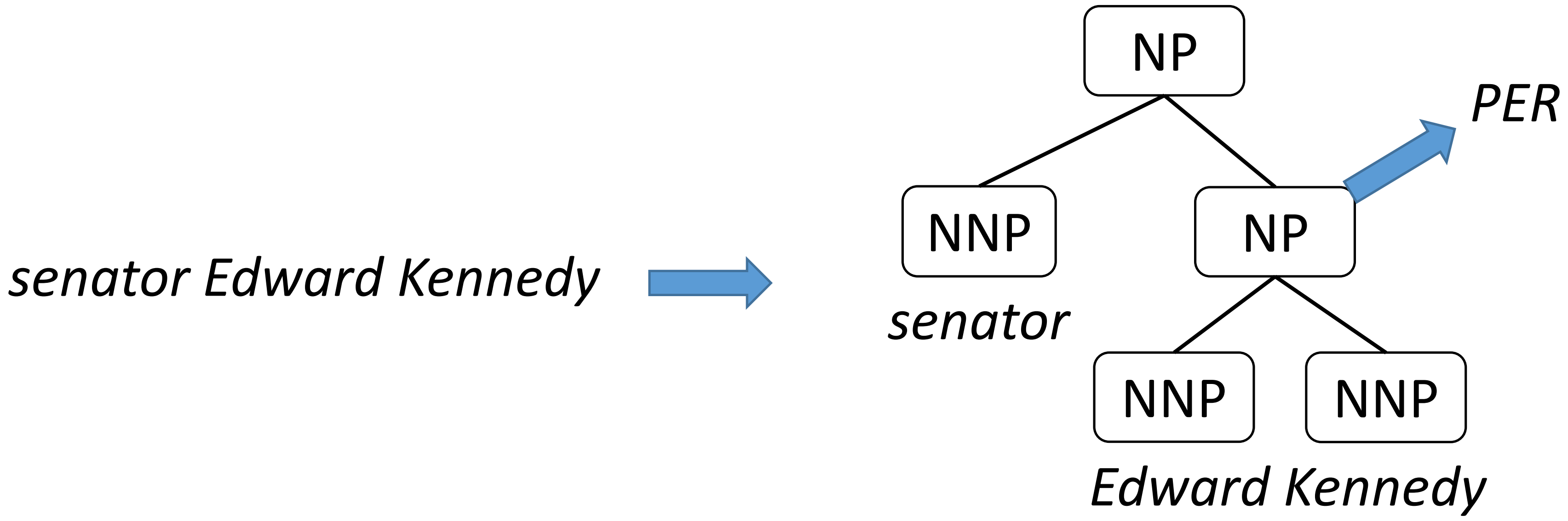
Ungrammatical → Unlikely an NE

CRF-CFG for Constituent Prediction



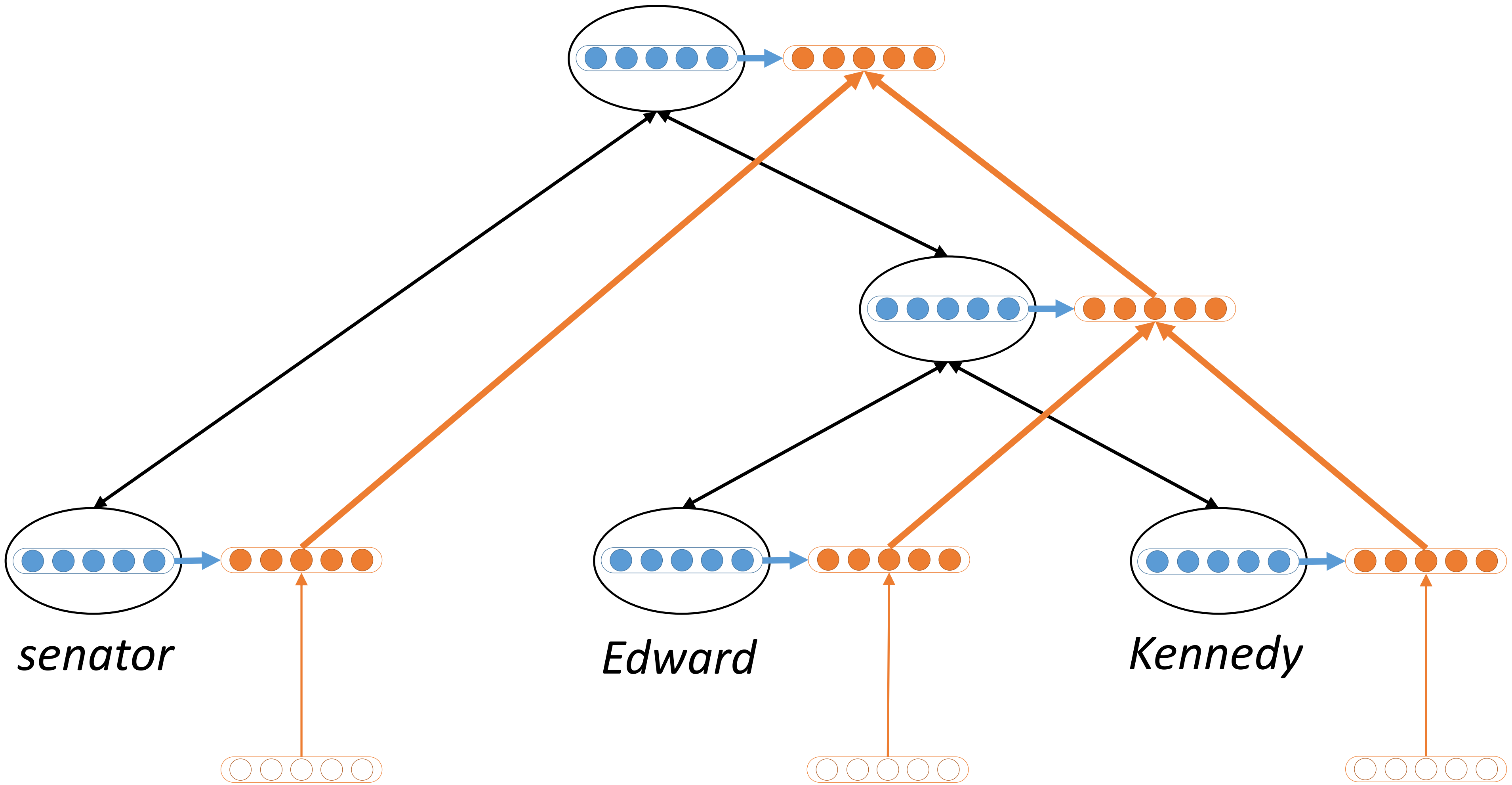
Jenny Rose Finkel and Christopher D. Manning. 2009.
Joint Parsing and Named Entity Recognition. In
Proceedings of HLT-NAACL.

Bi-Tree-LSTM for Constituent Prediction

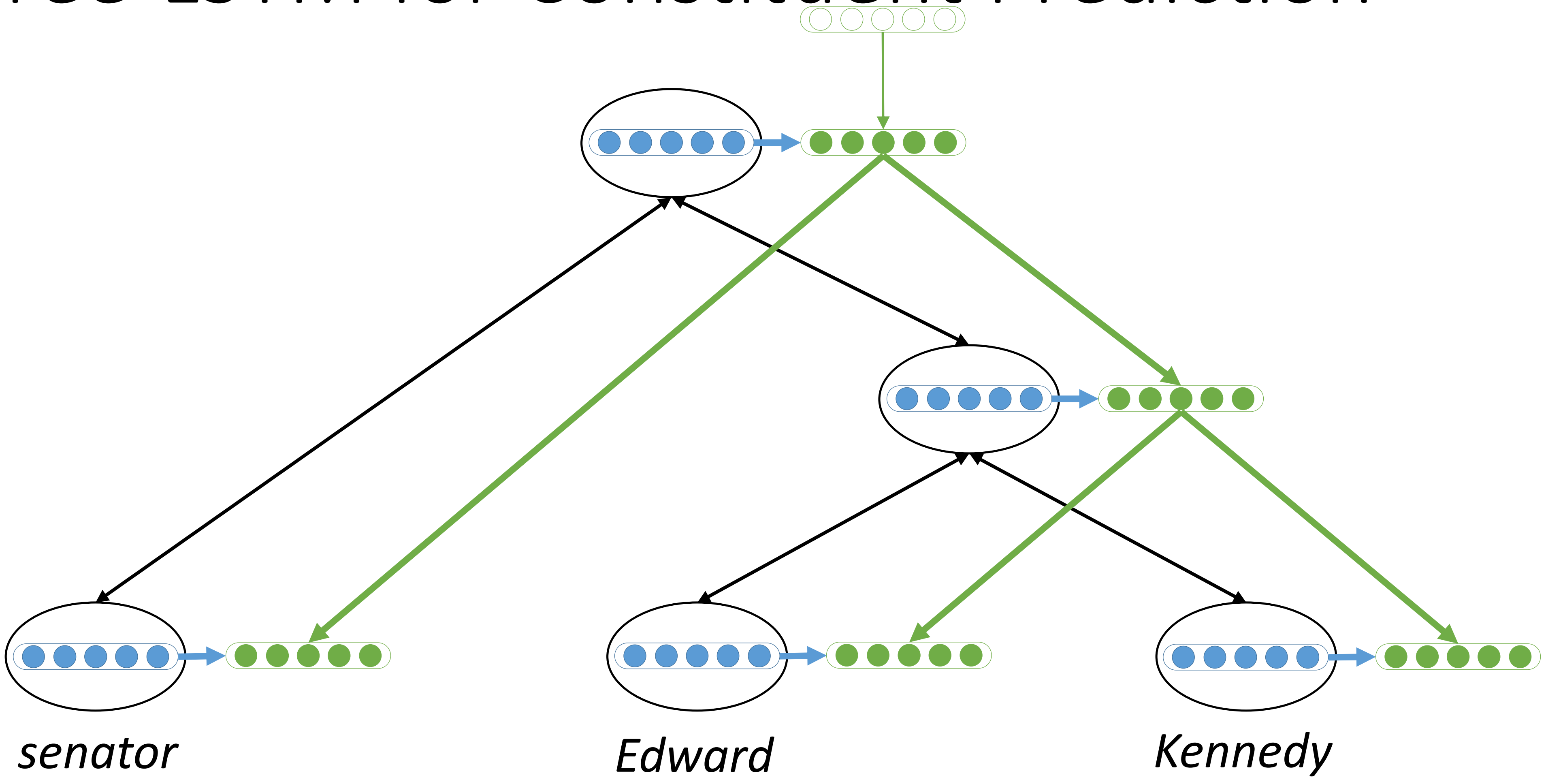


Peng-Hsuan Li, Ruo-Ping Dong, Yu-Siang Wang, Ju-Chieh Chou, and Wei-Yun Ma. 2017. Leveraging Linguistic Structures for Named Entity Recognition with Bidirectional Recursive Neural Networks. In Proceedings of EMNLP.

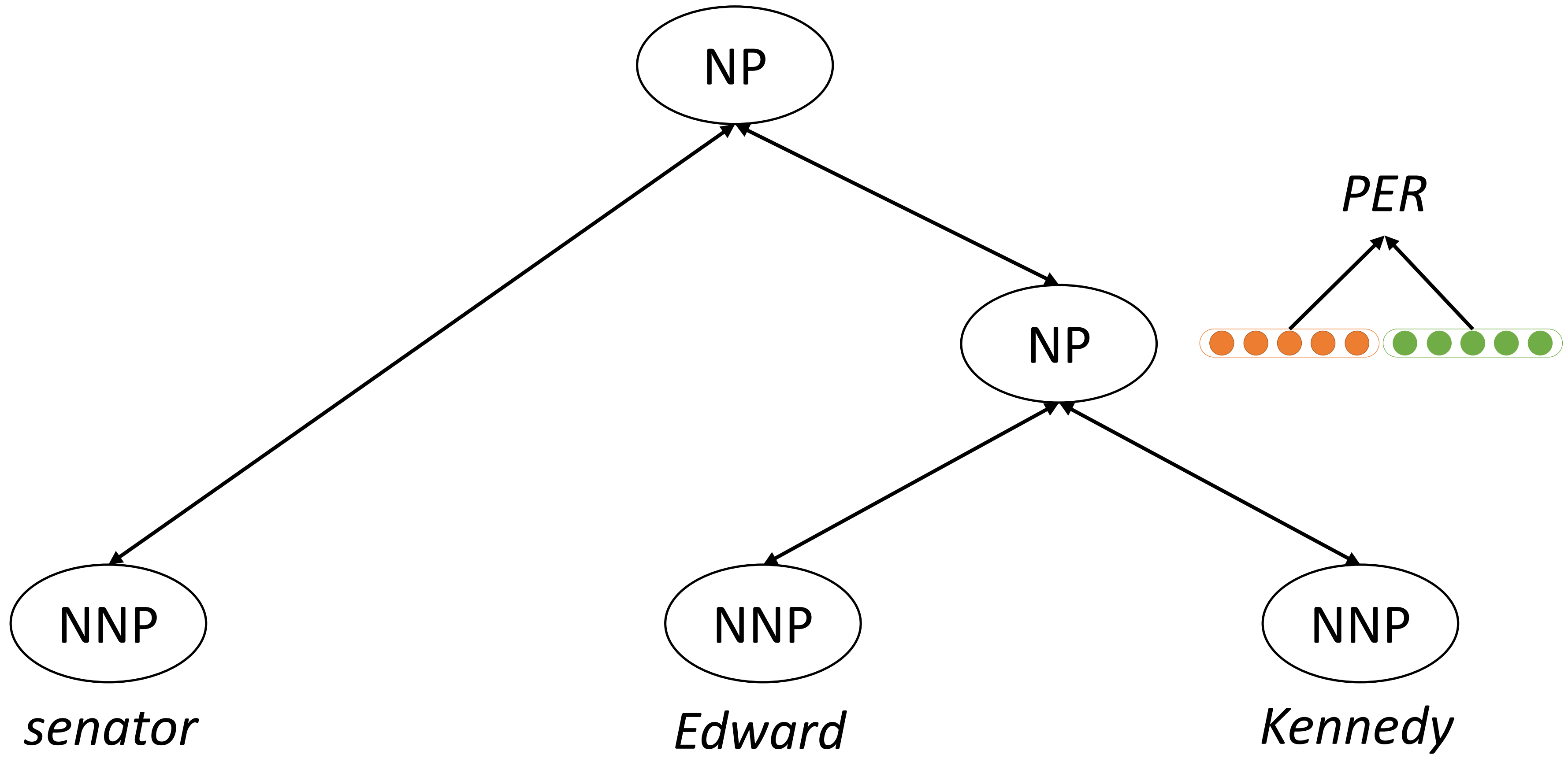
Bi-Tree-LSTM for Constituent Prediction



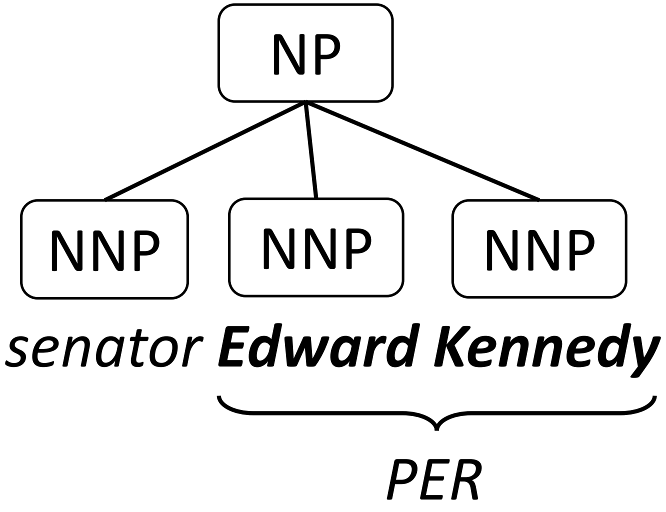
Bi-Tree-LSTM for Constituent Prediction



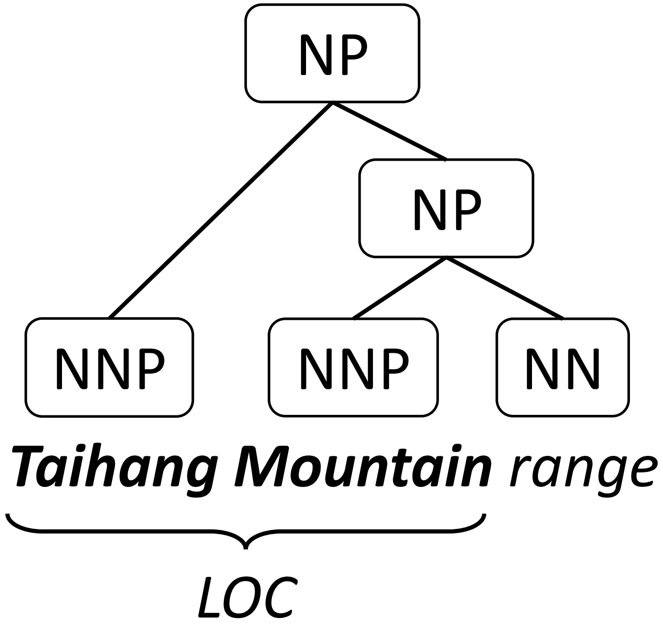
Bi-Tree-LSTM for Constituent Prediction



Inconsistencies between Parse and NER



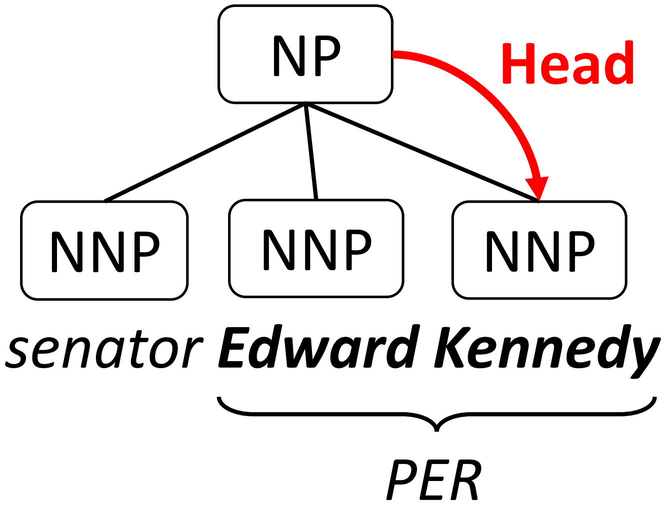
Type-1
Cross Siblings



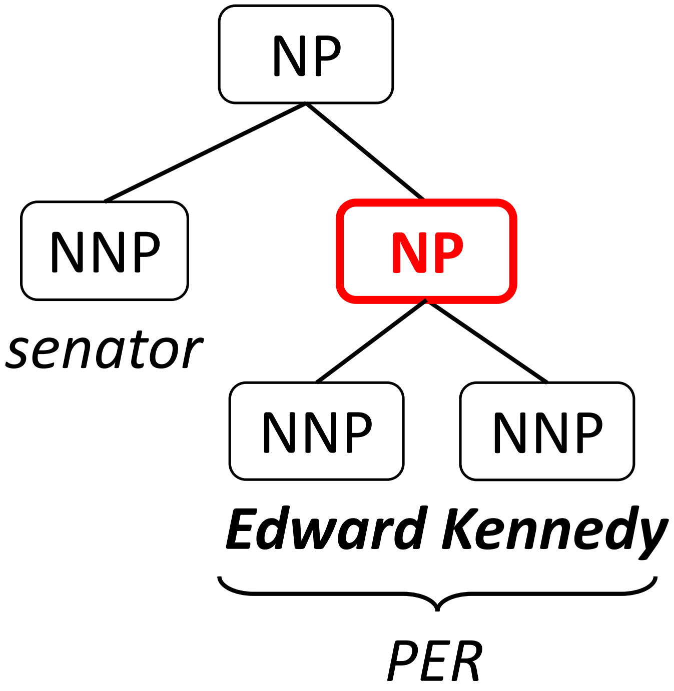
Type-2
Cross Branches

Peng-Hsuan Li. 2017. Leveraging Linguistic Structures for Named Entity Recognition with Bidirectional Recursive Neural Networks. Master's Thesis, Department of Computer Science and Information Engineering, National Taiwan University.

Eliminate Type-1: Constituency Tree Binarization

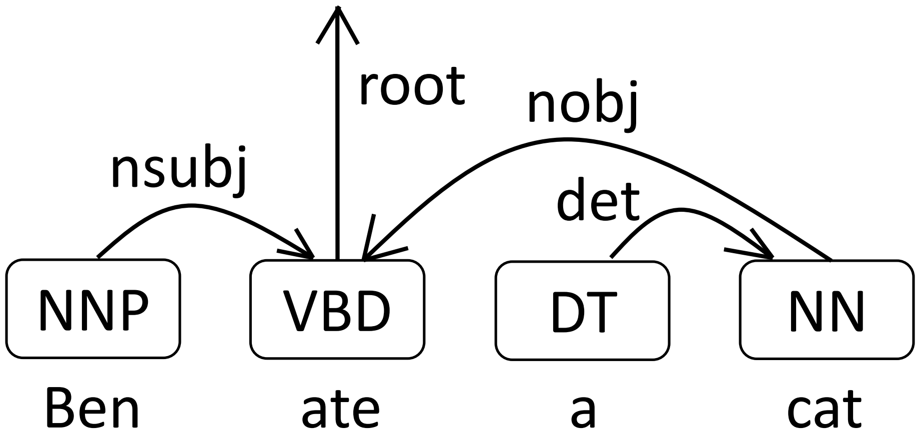


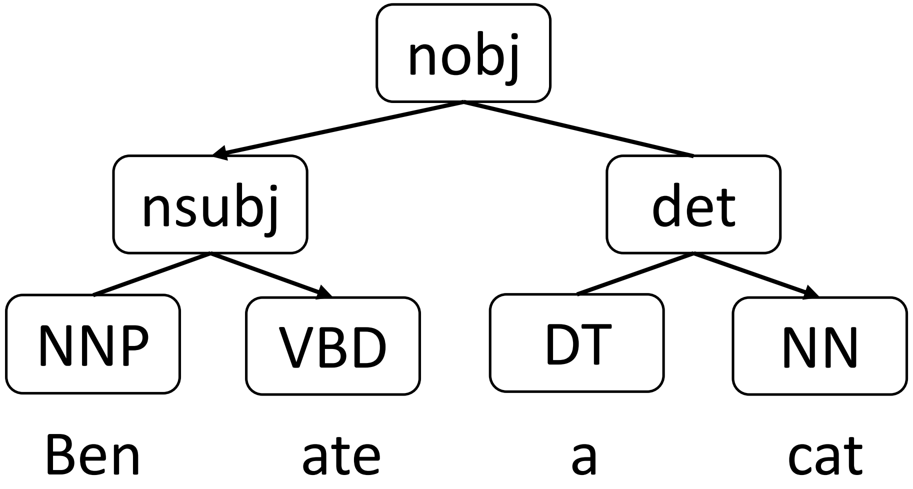
Type-1
Cross Siblings



Consistent

Eliminate Type-1: Dependency Transformation

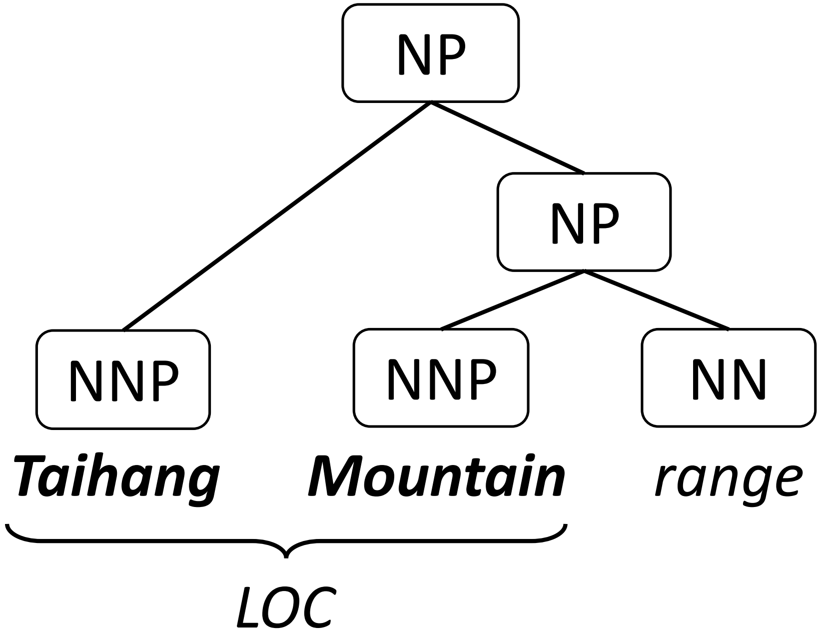


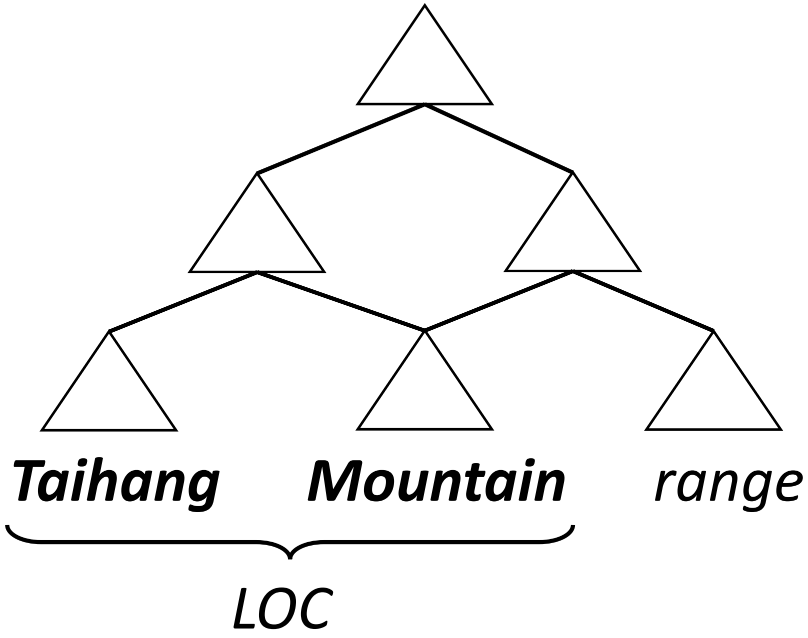


No Constituents

No Type-1 Inconsistencies

Eliminate Type-2: Pyramid Construction



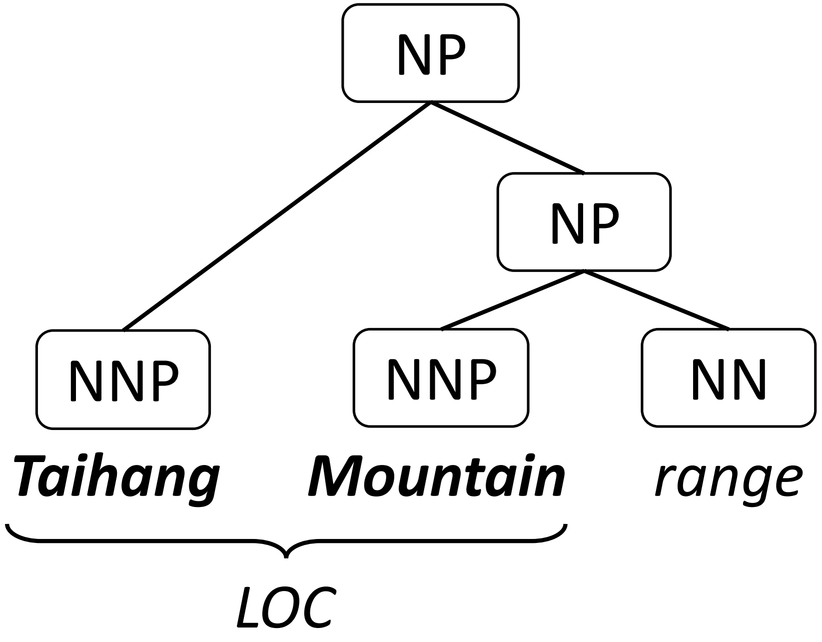


Type-2

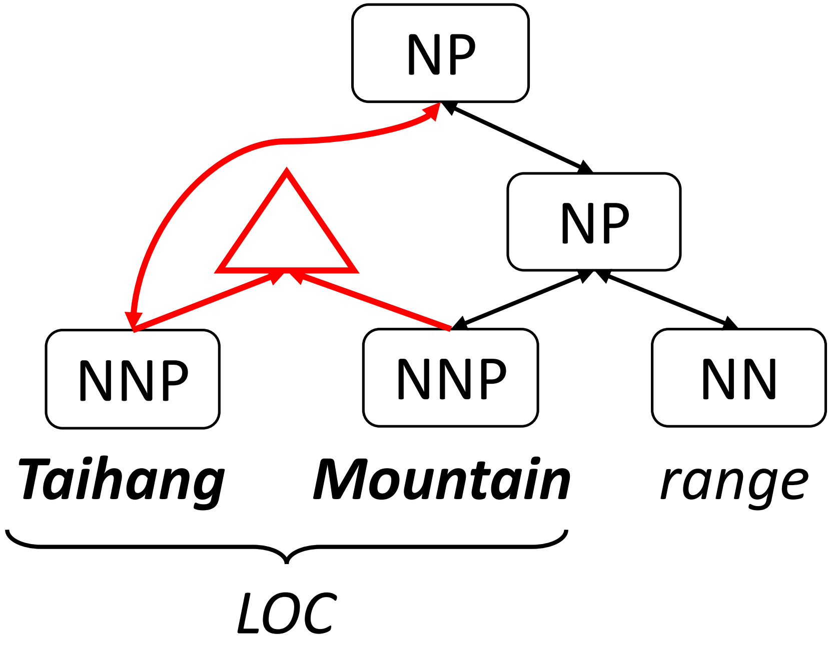
Cross Branches

No Liguistic Structures

Eliminate Type-2: Pyramid Construction



Type-2
Cross Branches



No Inconsistencies

Results of Constituent Prediction

Method	Model	Sources	CoNLL-2003	OntoNotes 5.0
Sequence Tagging	BLSTM	Huang et al., 2015 Chiu and Nichols, 2016 Ma and Hovy, 2016 Lample et al., 2016 Strubell et al., 2017	90.67	83.76
	BLSTM-CRF		90.94	86.99
	BLSTM-CNN		90.98	-
	BLSTM-CNN-CRF		**91.21**	-
	Deep BLSTM		-	86.19
	Deep BLSTM-CNN		-	86.41
	ID-CNN-CRF	Strubell et al., 2017	90.65	86.84
Constituent Prediction	CRF-CFG	Finkel and Manning, 2009	-	82.42
	Bi-Tree-RNN-CNN	Li et al., 2017	88.91	**87.21**

Analyses of Constituent Prediction

- Sequence Tagging vs. Constituent Prediction

Method	CoNLL-2003	OntoNotes 5.0
Sequence Tagging	**91.21**	86.99
Constituent Prediction	88.91	**87.21/88.92**

93% Consistency **97%/100% Consistency**

Analyses of Constituent Prediction

- Sequence vs Tree

*the first couple moves out of the **White House** on January 20th .*

			OntoNotes 5.0		
Model	Const-Only	Prediction	Precision	Recall	F1
Bi-RNN	X	***the White***	85.7	**86.5**	86.10
Bi-RNN	O	-	87.2	85.1	86.14
Bi-Tree-RNN	O	***White House***	**88.0**	86.2	**87.10**

Ablation Study: Constituency Tree Binarization

		OntoNotes 5.0			
Model	Binarize	Consistency	Precision	Recall	F1
BRNN	X	93%	87.3	83.0	85.11
BRNN	O	**97%**	**88.0**	**86.2**	**87.10**

Ablation Study: Dependency Transformation

		CoNLL 2003		
Model	Parser	Precision	Recall	F1
BRNN	StanfordRNN	88.9	86.9	87.91
BRNN	SyntaxNet	**90.2**	**87.7**	**88.91**

Ablation Study: Pyramid Construction

		CoNLL 2003		
Model	Pyramid	Precision	Recall	F1
BRNN	X	89.1	82.9	85.89
BRNN	O	**90.2**	**87.7**	**88.91**

Ablation Study: Bidirectional

		OntoNotes 5.0		
Model	***Koran***	Precision	Recall	F1
Top-Down	-	79.2	69.3	73.93
Bottom-Up	***PERSON***	86.6	86.2	86.41
BRNN	***WORK OF ART***	**88.0**	**86.2**	**87.10**

```
|       |--PP
|           |--IN by
|           |--S
|               |--VP
|                   |--VBG repeating
|                   |--NP
|                       |--NP
|                       |   |--NP
|                       |   |   |--NP
|                       |   |   |   |--NNS verses
|                       |   |   |
|                       |   |   |--PP
|                       |   |       |--IN from
|                       |   |       |--NP
|                       |   |           |--DT the
|                       |   |           |--NP
|                       |   |               |--JJ noble
|                       |   |               |--NNP Koran
|                       |   |
|                       |   |--CC and
|                       |
|                       |--NP
|                           |--DT the
|                           |--NP
|                               |--CD two
|                               |--NNS testimonies
|
|--. .
```

He confirmed it by repeating the verses from the noble Koran and the two testimonies.