Building Neural NER Models with Structural Prior

2018/09/11

Peng-Hsuan Li

Outline

- Named Entity Recognition
  - Task
  - Features
  - Related Work
- Leveraging Linguistic Structures for NER
- Constructing Deep Cross-BLSTM with Self-Attention for NER
- CKIP 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)

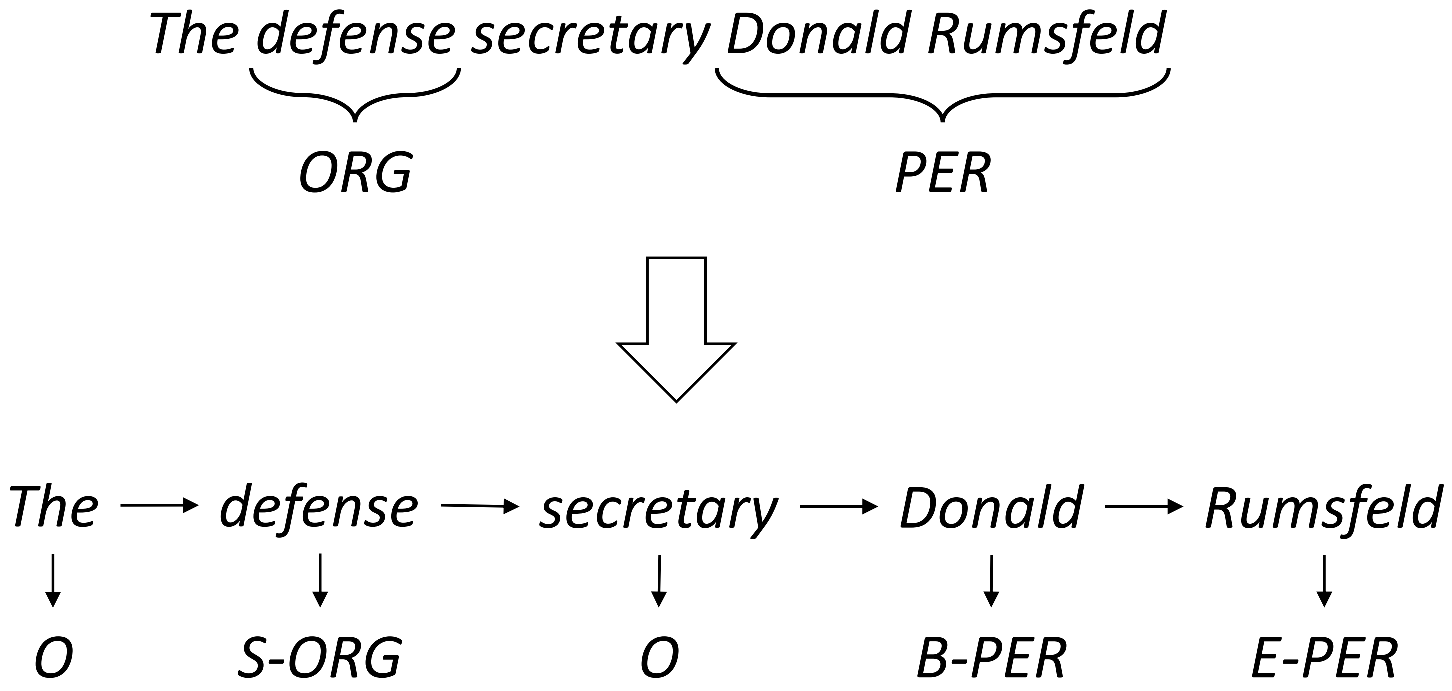
- Uppercase
- Upper-initial
- Lowercase
- Mixed

Character Features

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

- Uppercase
- Lowercase
- Digit

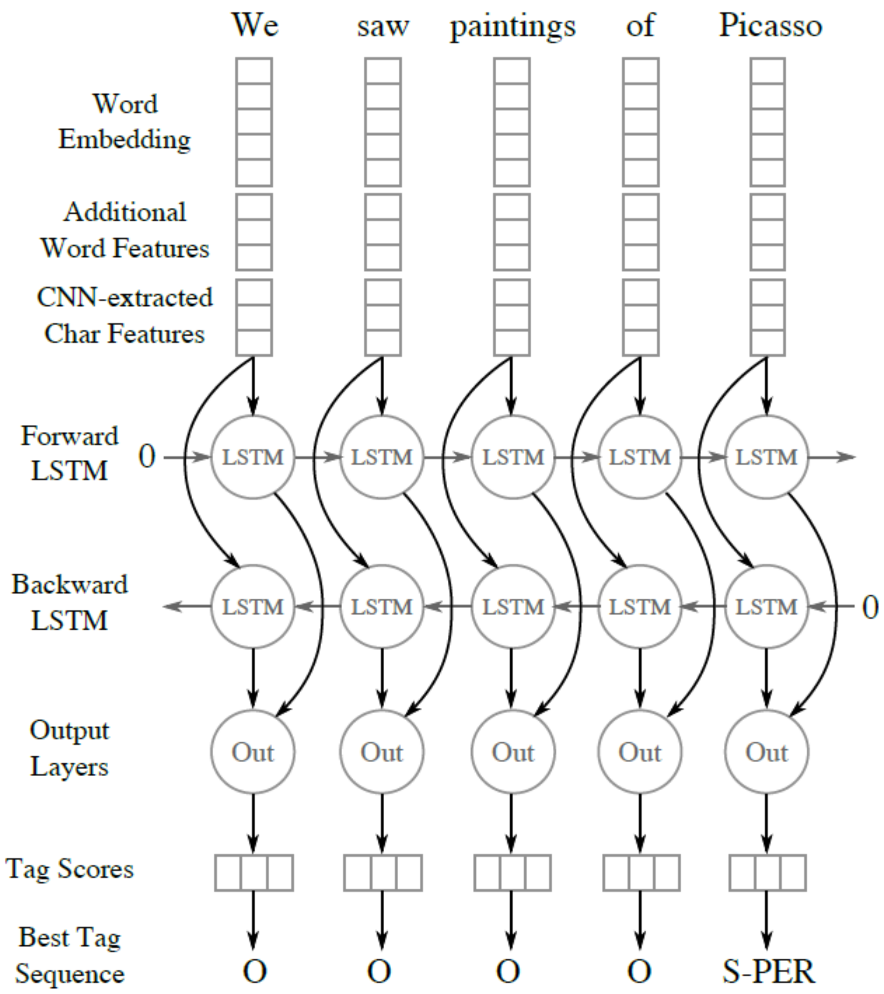
Sequence Tagging

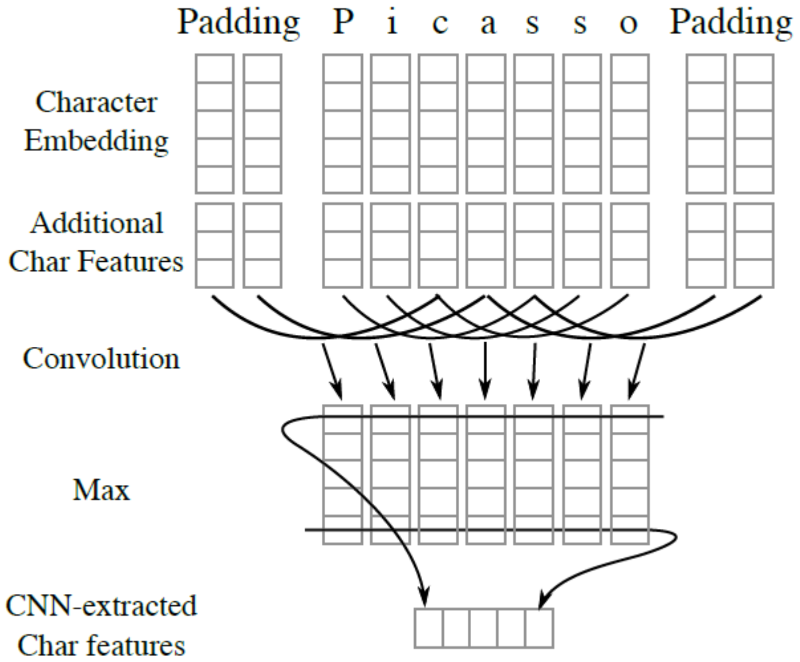


Chunk Labels

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

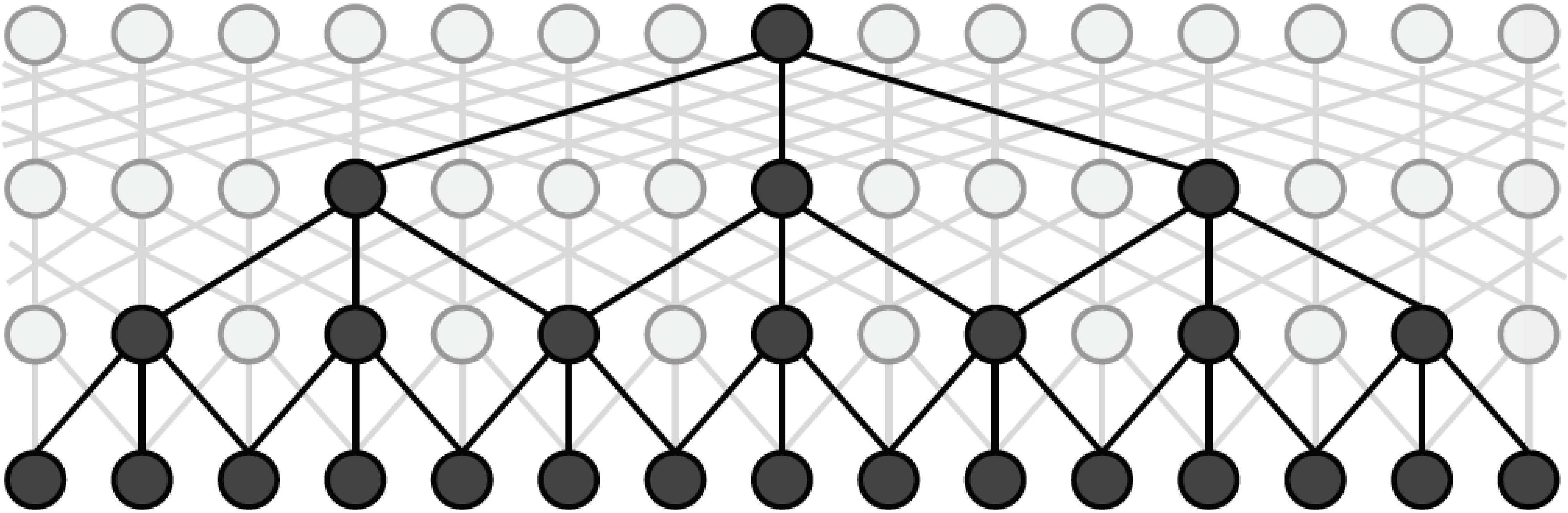
Bi-LSTM for Sequence Tagging





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.

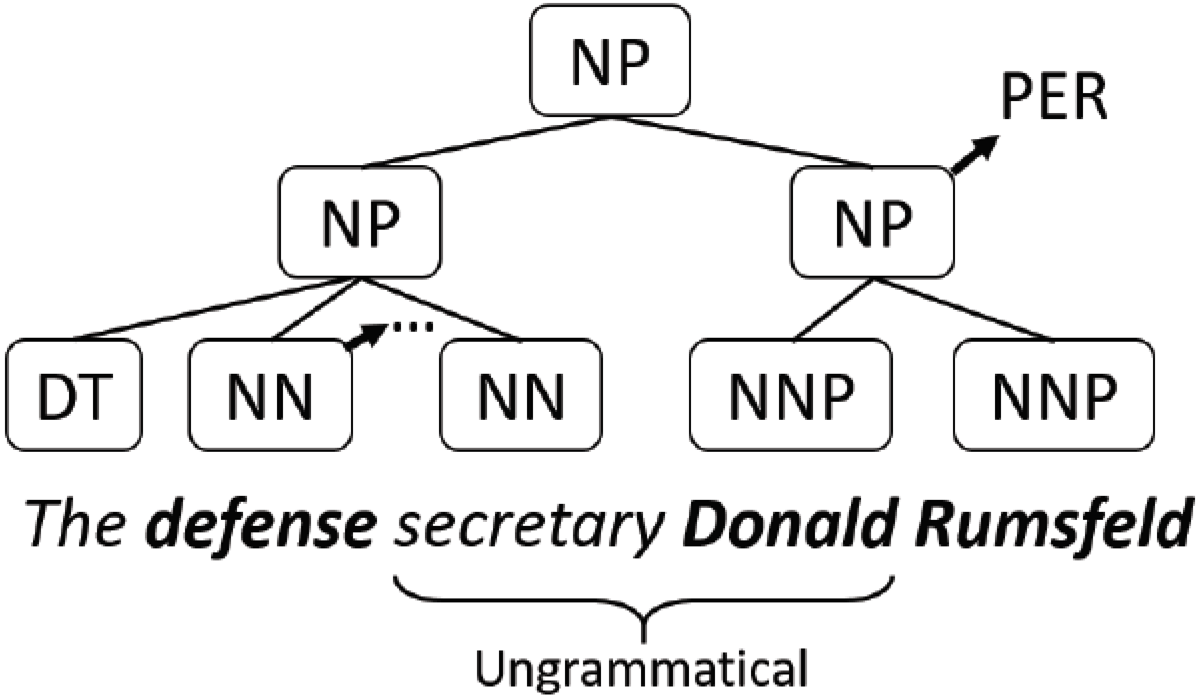
Results of Sequence Tagging

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
- Constructing Deep Cross-BLSTM with Self-Attention for NER
- CKIP NER

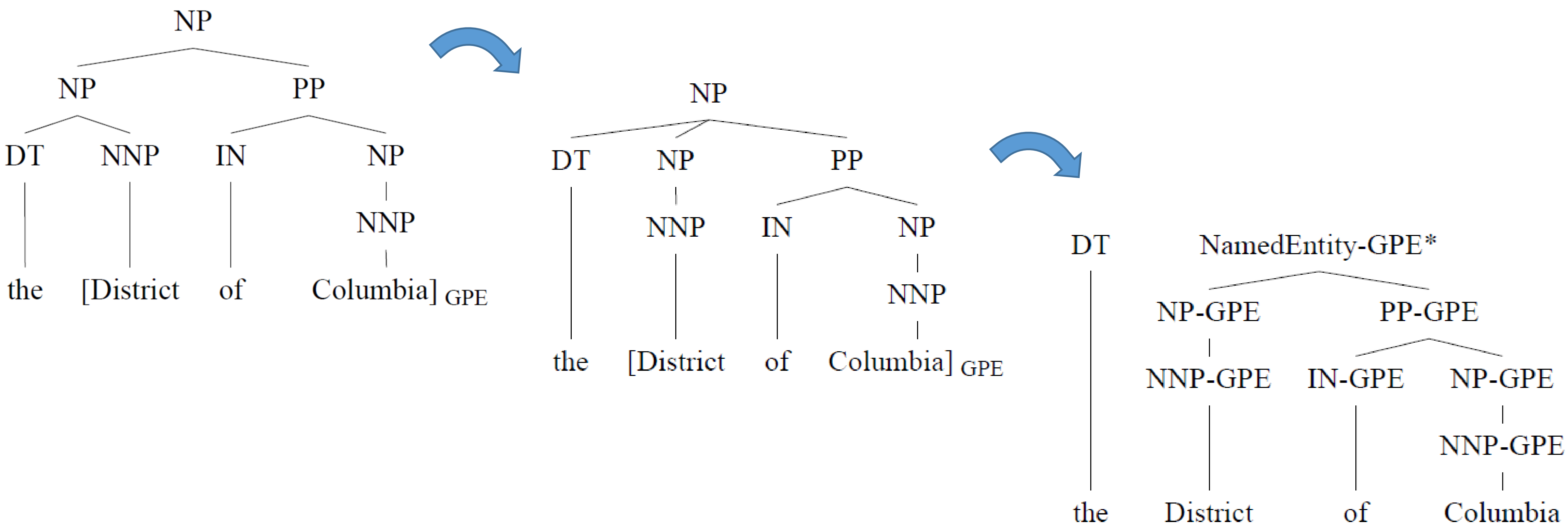
Constituent Prediction



Constituent → A plausible NE candidate

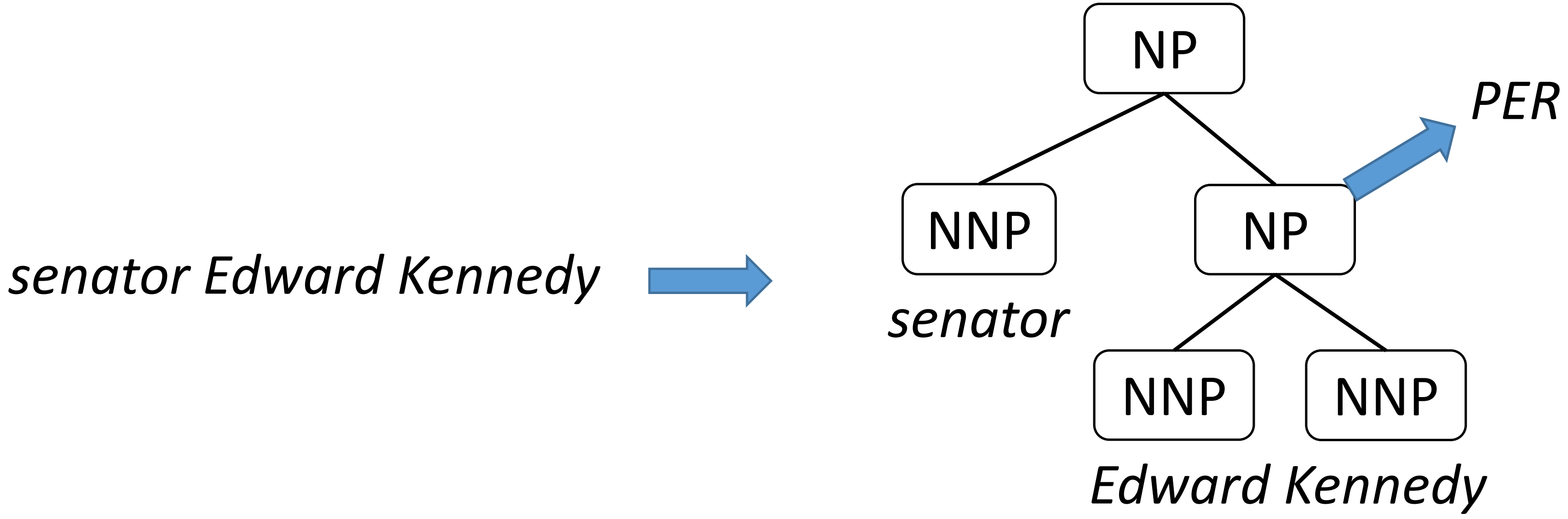
Ungrammatical → Unlikely an NE

CRF-CFG for Constituent Prediction



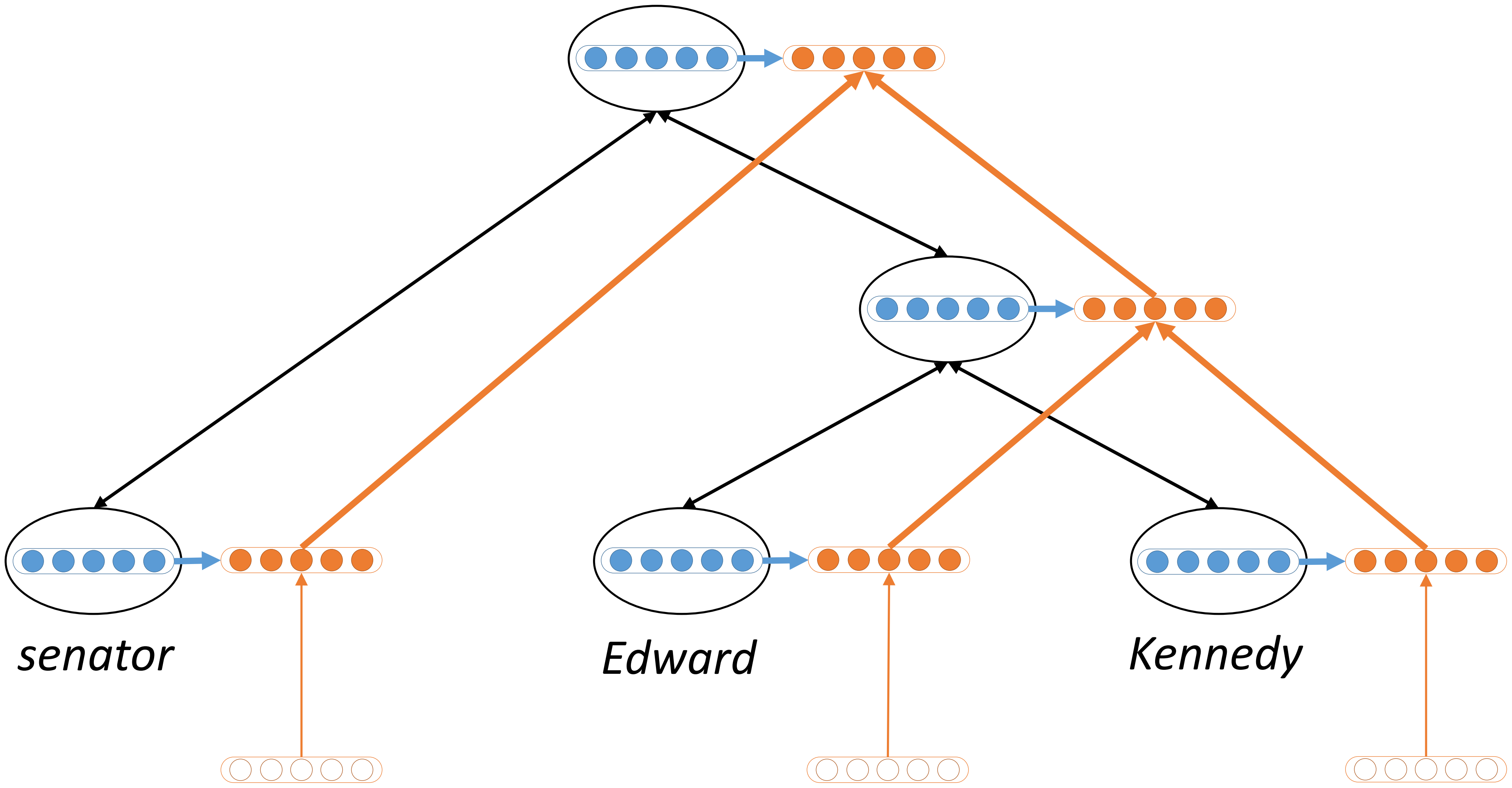
Jenny Rose Finkel and Christopher D. Manning. 2009.
Joint Parsing and Named Entity Recognition. In
Proceedings of HLT-NAACL.

Bi-Tree-LSTM for Constituent Prediction

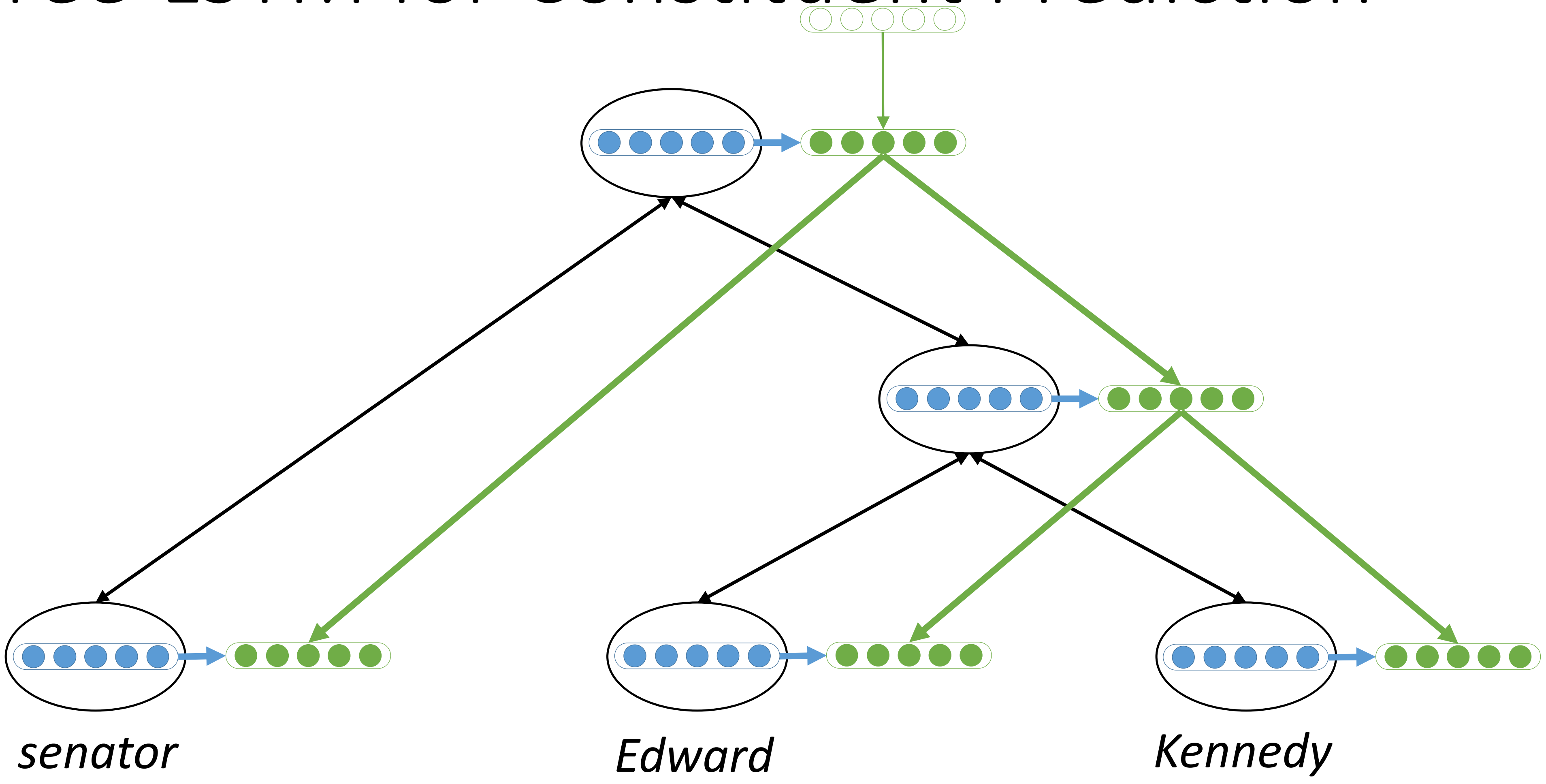


Peng-Hsuan Li, Ruo-Ping Dong, Yu-Siang Wang, Ju-Chieh Chou, and Wei-Yun Ma. 2017. Leveraging Linguistic Structures for Named Entity Recognition with Bidirectional Recursive Neural Networks. In Proceedings of EMNLP.

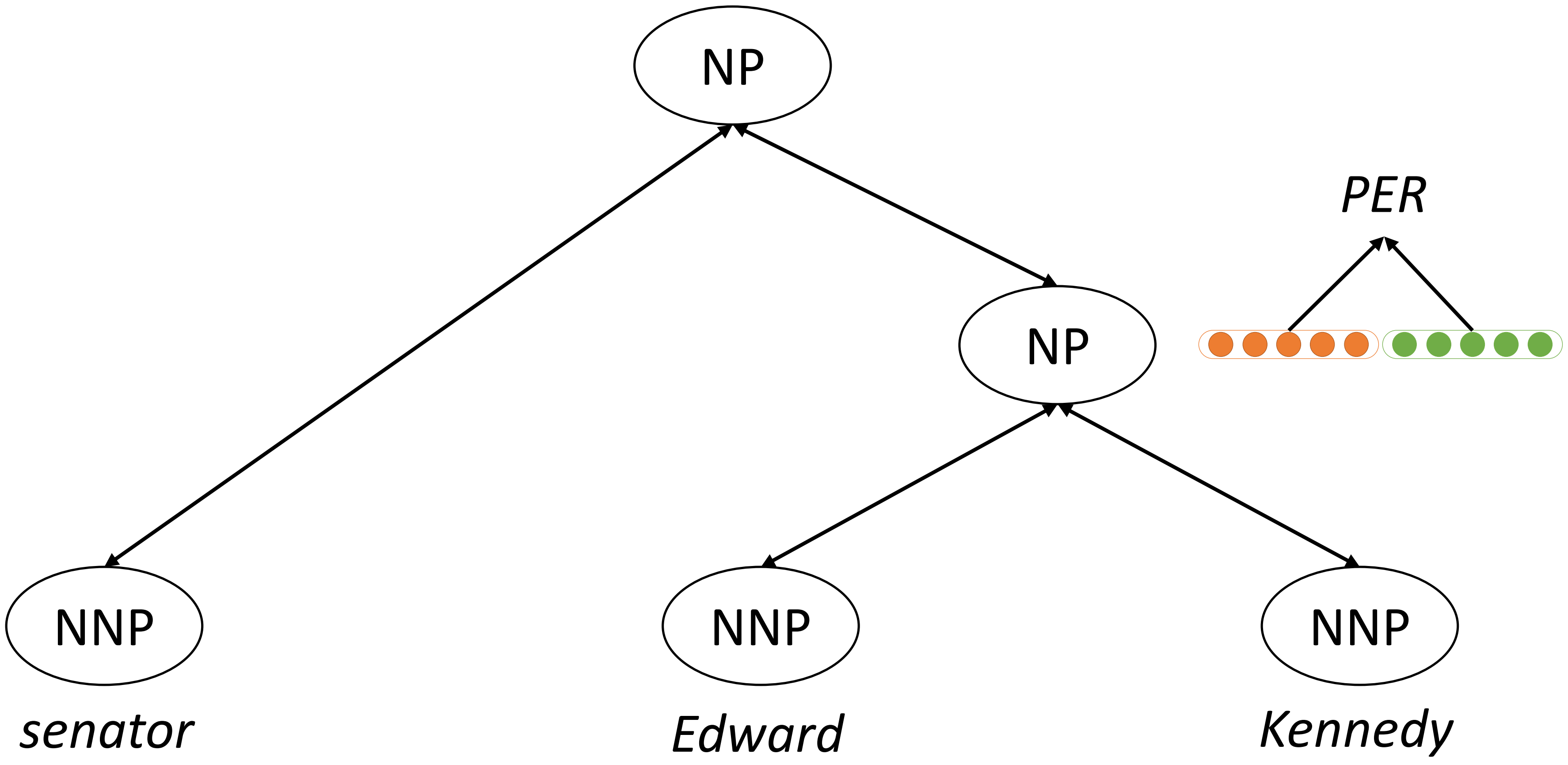
Bi-Tree-LSTM for Constituent Prediction



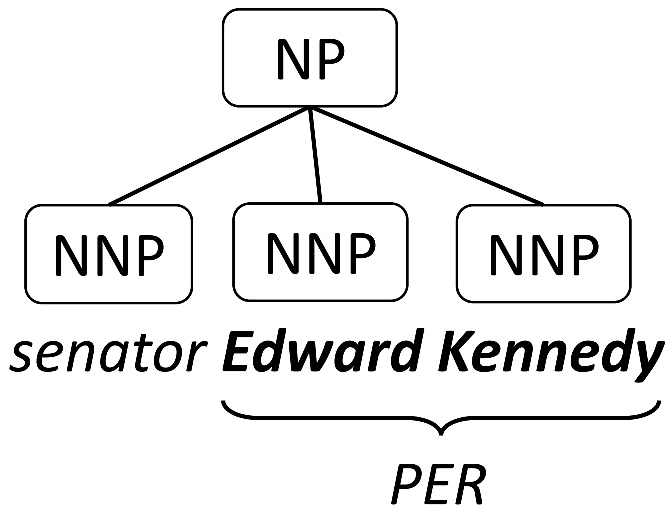
Bi-Tree-LSTM for Constituent Prediction

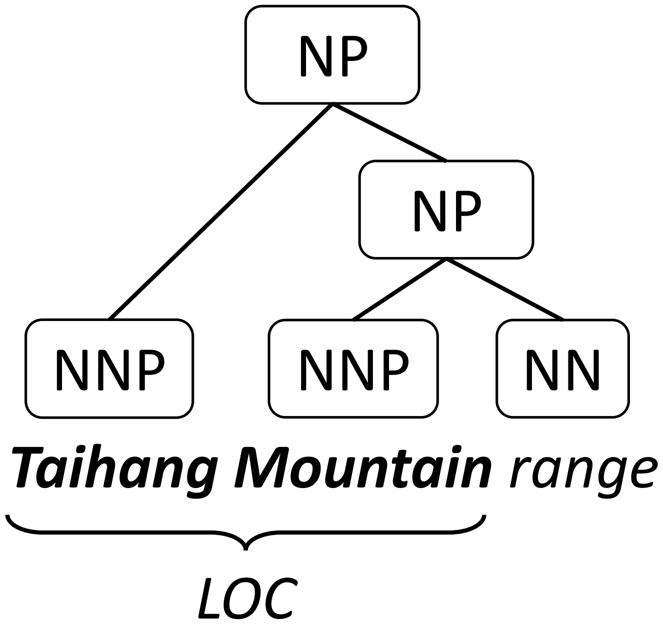


Bi-Tree-LSTM for Constituent Prediction



Inconsistencies between Parse and NER



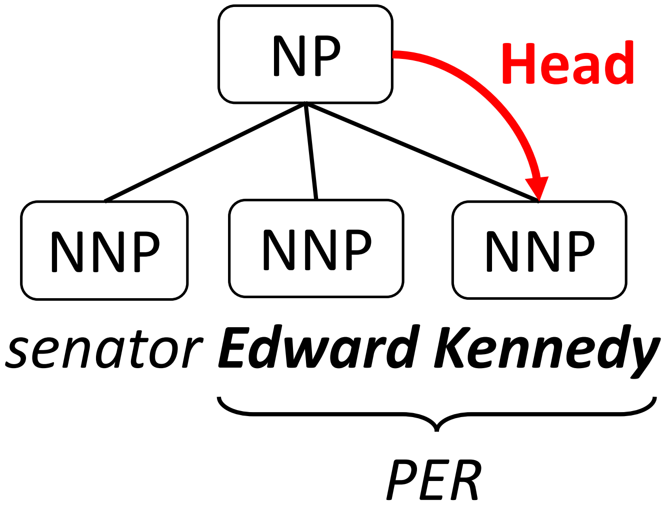


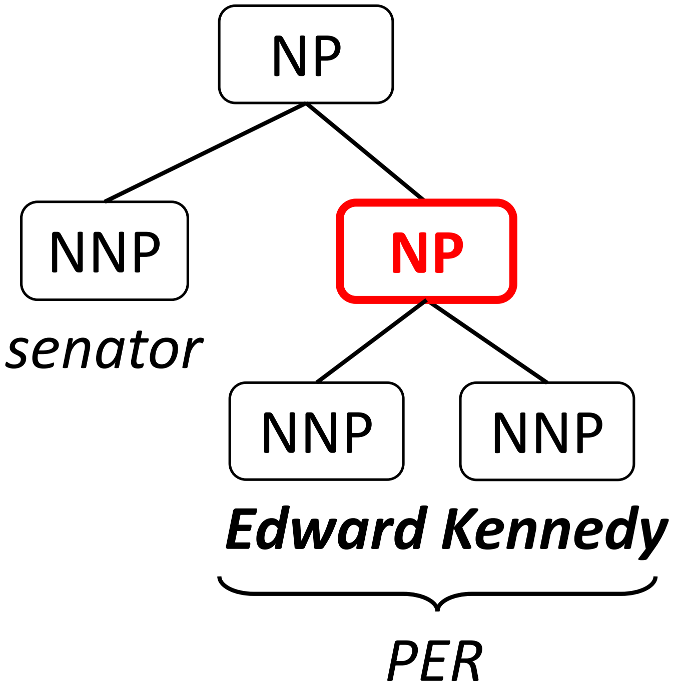
Type-1
Cross Siblings

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

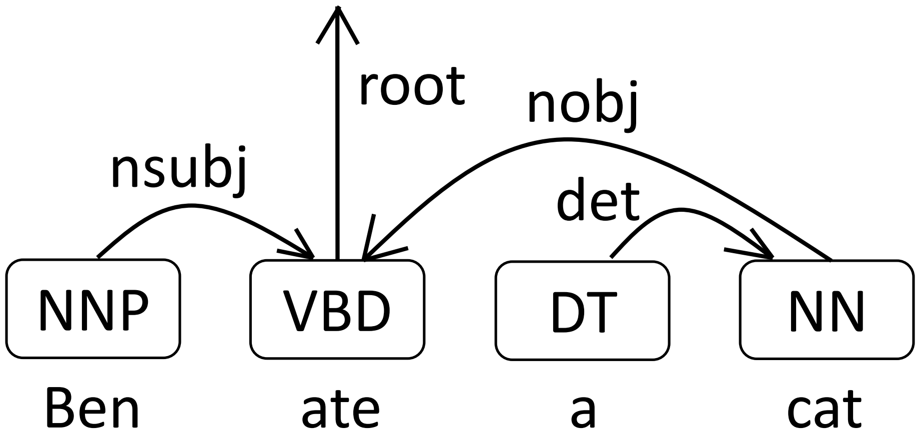


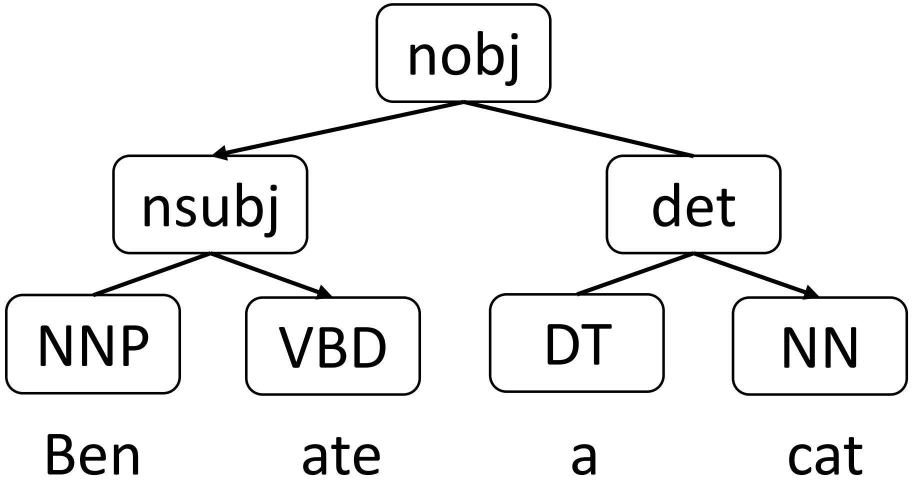


Type-1
Cross Siblings

Consistent

Eliminate Type-1: Dependency Transformation

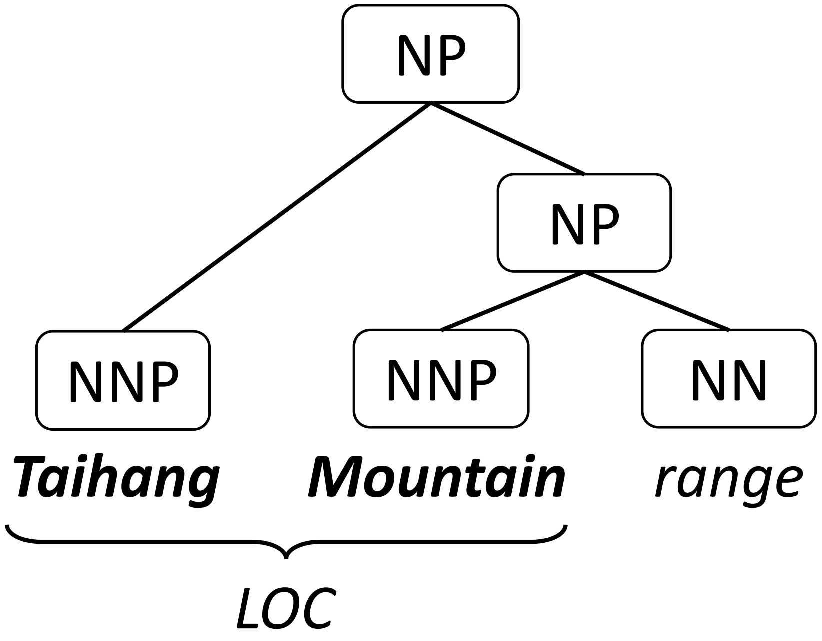




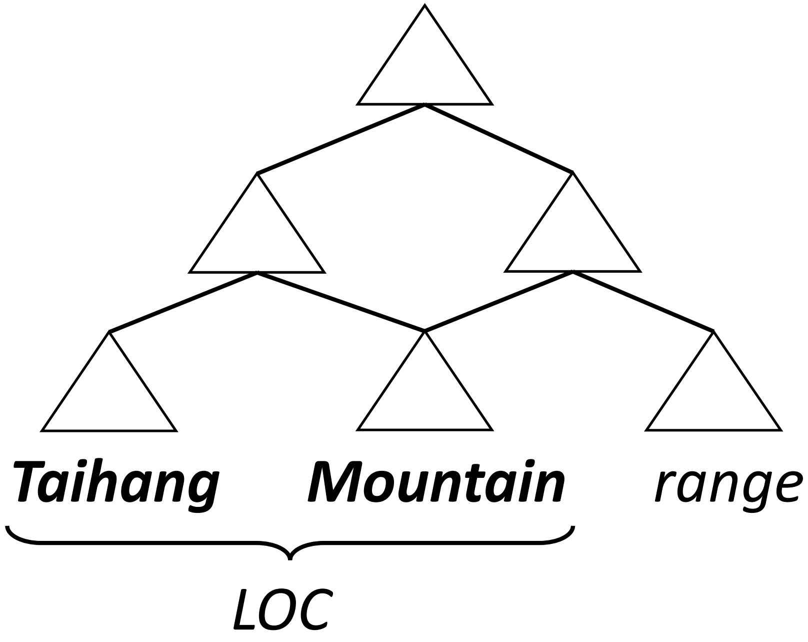
No Constituents

No Type-1 Inconsistencies

Eliminate Type-2: Pyramid Construction

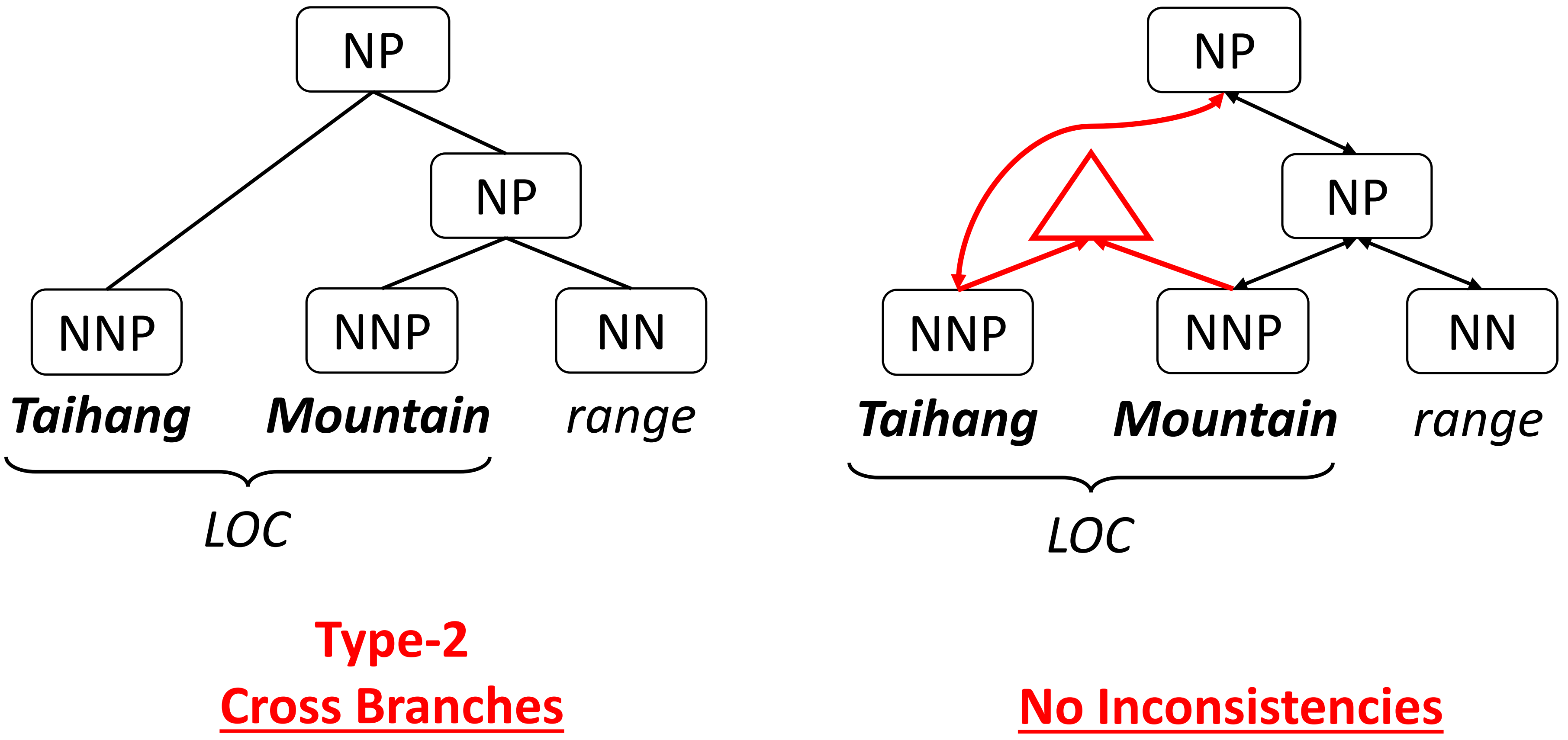


Type-2
Cross Branches



No Liguistic Structures

Eliminate Type-2: Pyramid Construction



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.

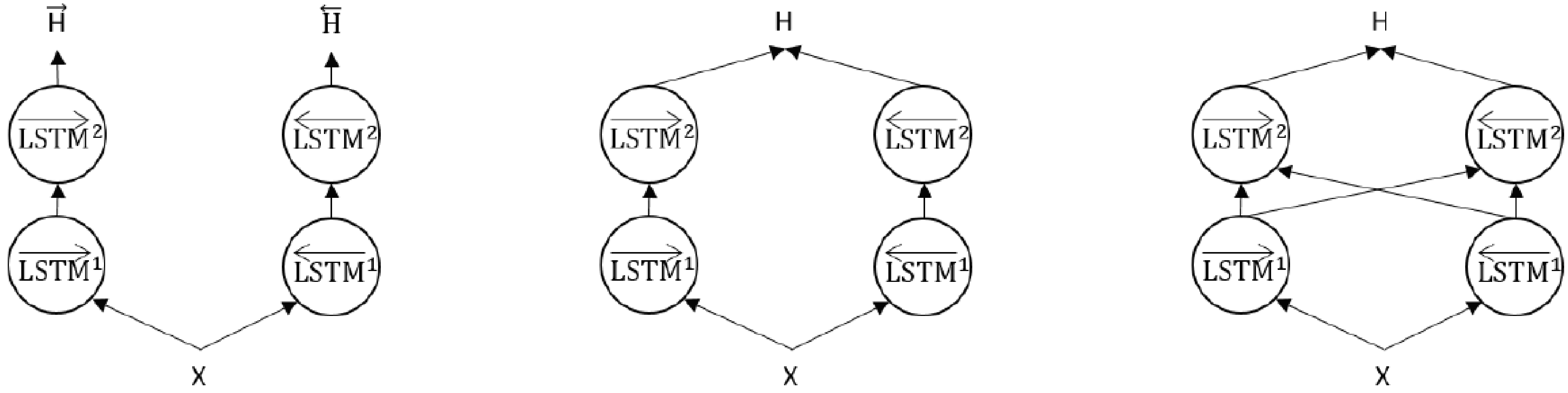
Outline

- Named Entity Recognition
- Leveraging Linguistic Structures for NER
- Constructing Deep Cross Bi-LSTM with Self-Attention for NER
  - Deep Cross Bi-LSTM
  - Multi-head self-attention
- CKIP NER

A Pattern Across Past and Future

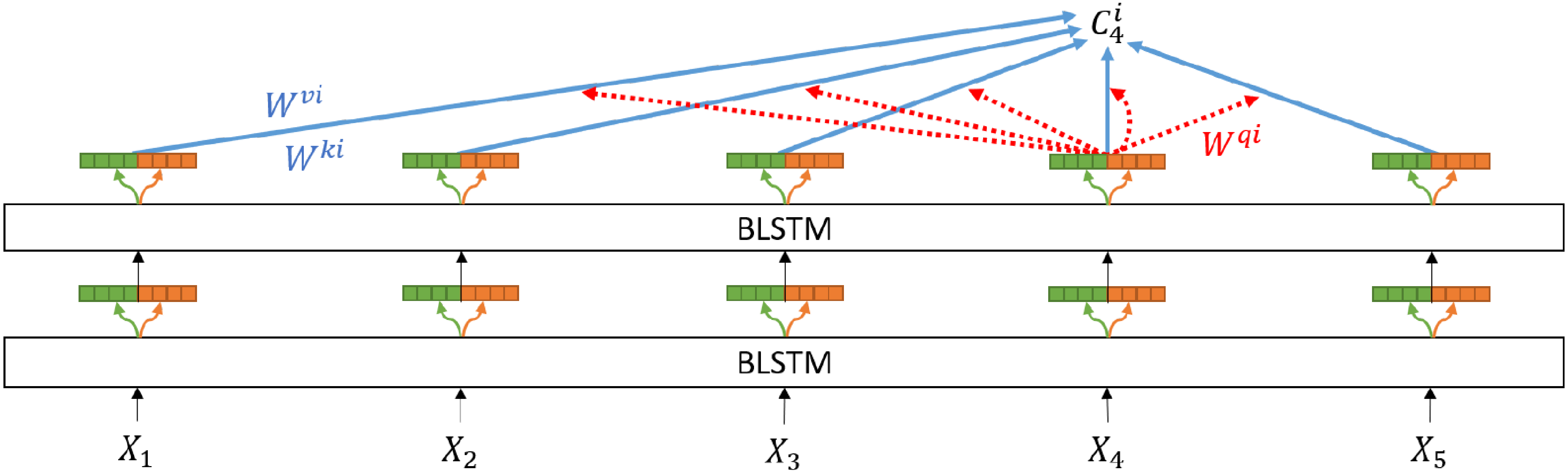
- $(w_{l1}, w_c, w_{r1}) \rightarrow (O, O, O)$
- $(w_{l1}, w_c, w_{r2}) \rightarrow (O, O, O)$
- $(w_{l1}, w_c, w_{r3}) \rightarrow (B, I, E)$-ORG
- $(w_{l2}, w_c, w_{r3}) \rightarrow (O, O, O)$
- $(w_{l3}, w_c, w_{r3}) \rightarrow (O, O, O)$

Deep Bi-LSTM



(a) C&N (b) Parallel-BLSTM (c) Cross-BLSTM

Self-Attention



$$\alpha^i = \sigma\left(\frac{HW^{qi}(HW^{ki})^T}{\sqrt{d'_h}}\right) \qquad C^i = \alpha^i HW^{vi} \qquad C = \left[C^1\ C^2\ ..\ C^m\right]W^c$$

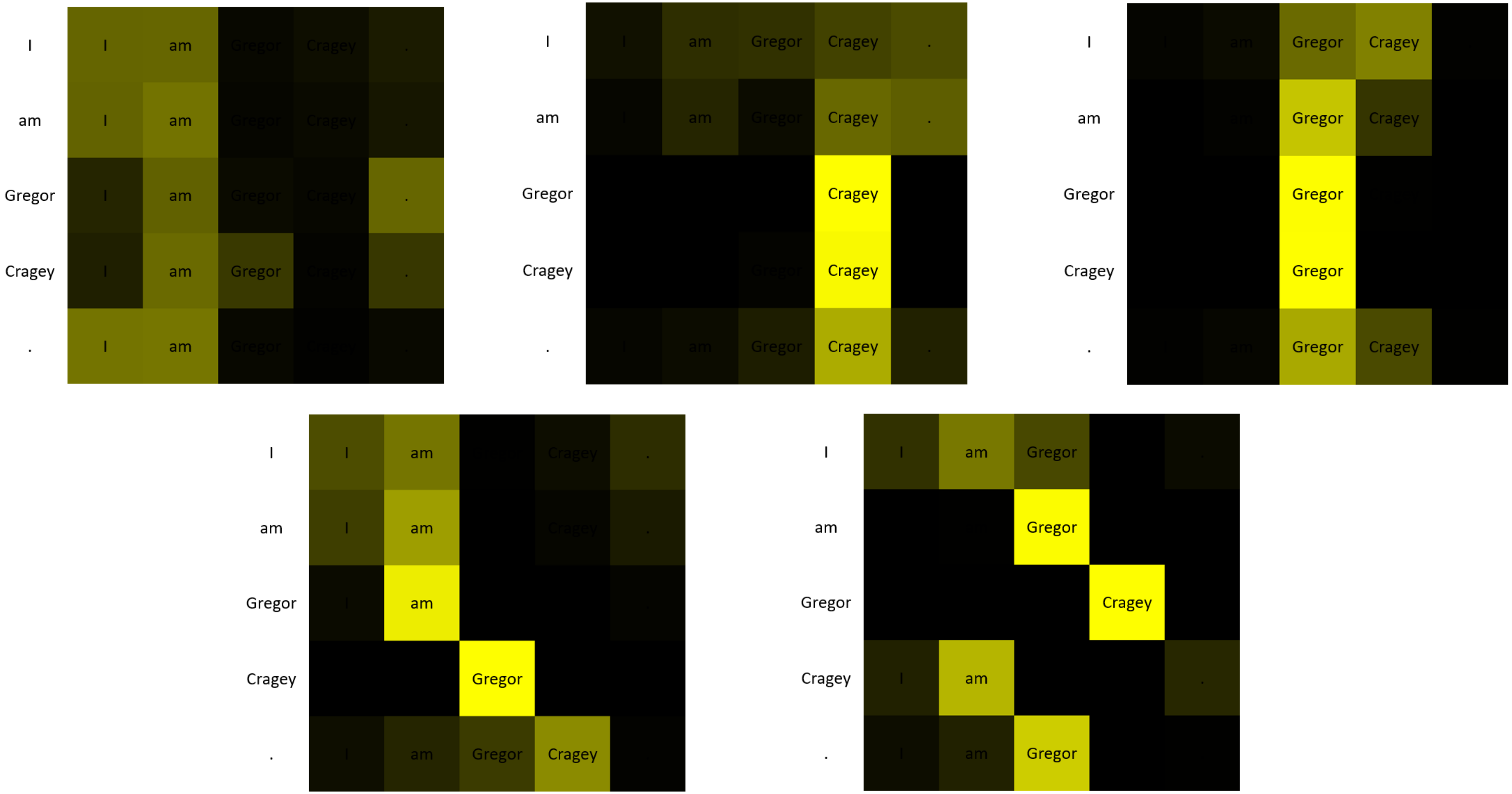
Peng-Hsuan Li and Wei-Yun Ma. Constructing Deep BLSTM-CNN with Self-Attention for Sequence Labeling NER. Under review.

Results

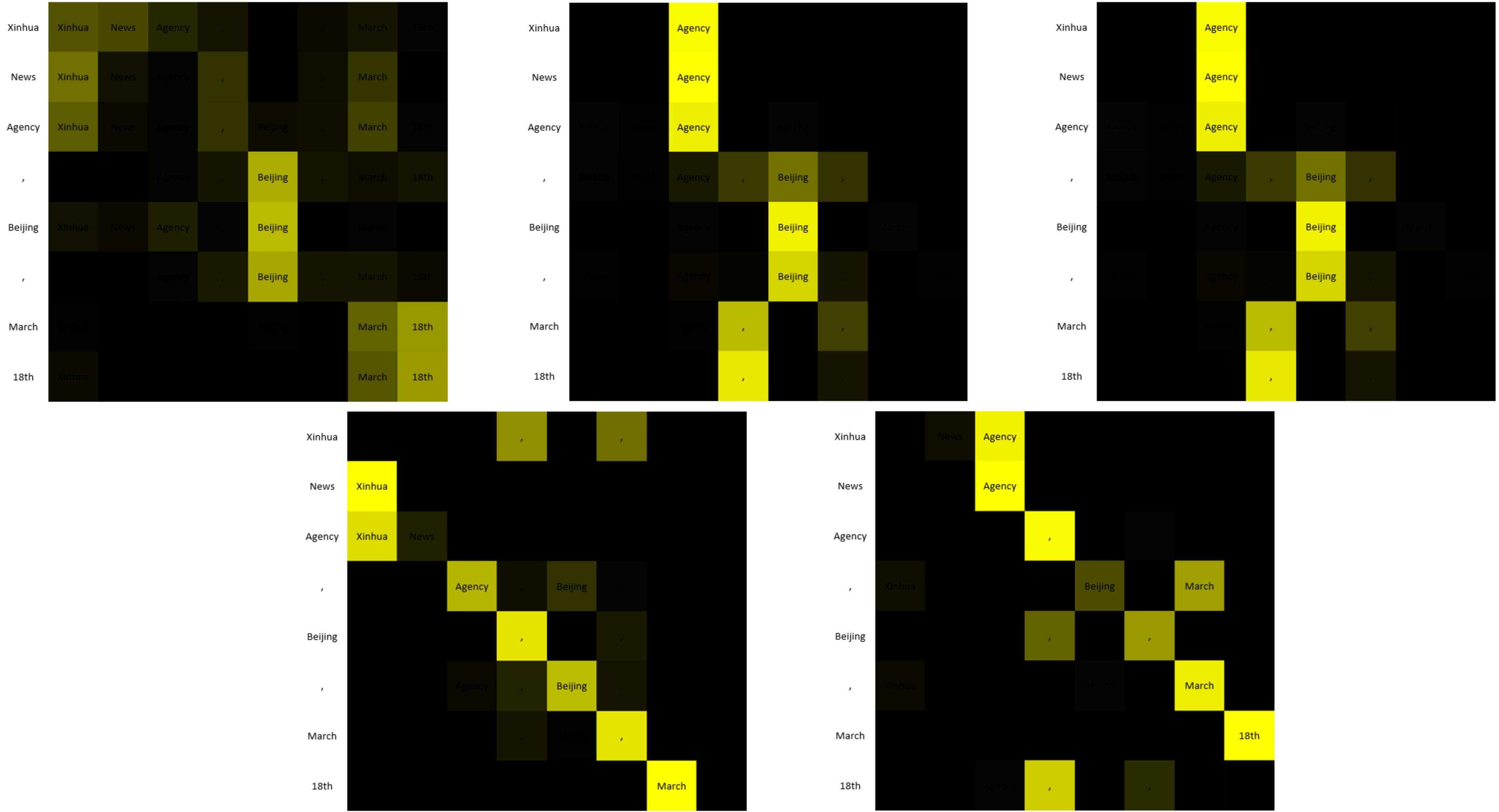
Method	Model	Sources	CoNLL-2003	OntoNotes 5.0
Sequence Tagging	BLSTM	Huang et al., 2015 Chiu and Nichols, 2016 Ma and Hovy, 2016 Lample et al., 2016 Strubell et al., 2017	90.67	83.76
	BLSTM-CRF		90.94	86.99
	BLSTM-CNN		90.98	-
	BLSTM-CNN-CRF		91.21	-
	Deep BLSTM		-	86.19
	Deep BLSTM-CNN		-	86.41
	ID-CNN-CRF	Strubell et al., 2017	90.65	86.84
	Deep Parallel BLSTM-CNN		**91.44**	87.69
	Deep Parallel BLSTM-CNN-Attend (5-head)		91.37	88.13
	Deep Cross BLSTM-CNN		91.24	**88.39**
	Deep Cross BLSTM-CNN-Attend (5-head)		91.14	88.35
Constituent Prediction	CRF-CFG	Finkel and Manning, 2009	-	82.42
	Bi-Tree-RNN-CNN	Li et al., 2017	88.91	87.21

oh, New York was really fun

I am Gregor Cragey.



Xinhua News Agency, Beijing, March 18th



Outline

- Named Entity Recognition
- Leveraging Linguistic Structures for NER
- Constructing Deep Cross Bi-LSTM with Self-Attention for NER
- CKIP NER
  - Chinese NER
  - Ancient Chinese Document NER

CKIP Chinese NER

Paste the text you want to process here(Chinese only):

比爾·弗雷利克是人權觀察組織的難民政策主任，他指出，在過去美國經常援助那些因支持它而遭受迫害的人。
自從越南戰爭以來，有一百萬越南難民在美國定居，包括數萬名南越退伍軍人。
但是布什政府"已經放棄了那個義務，"弗雷利克說。
"出逃的人是那些政府曾賴以在伊拉克建立民主的人；
它寧可忽視他們也不願意承認它的倡議已經失敗。"

◉ CKIP WordSeg (soft boundary) ○ CKIP WordSeg (hard boundary) ○ CKIP Parser (hard boundary)

Clear Submit

比爾·弗雷利克PERSON是**人權觀察組織**ORG的難民政策主任，他指出，在過去**美國**GPE經常援助那些因支持它而遭受迫害的人。

自從**越南戰爭**EVENT以來，有**一百萬**CARDINAL**越南**NORP難民在**美國**GPE定居，包括**數萬**CARDINAL名**南越**NORP退伍軍人。

但是**布什**PERSON政府"已經放棄了那個義務，"**弗雷利克**PERSON說。

"出逃的人是那些政府曾賴以在**伊拉克**GPE建立民主的人；

它寧可忽視他們也不願意承認它的倡議已經失敗。"

- http://deep.iis.sinica.edu.tw:9001/

CKIP Ancient Chinese NER

漢籍詞彙分類信心指數

一行一筆資料，欄位以 tab 分隔，由左到右為label、關鍵詞、前文、後文

```
_	沈演	鄧思啟為雲南右參政○陞禮部郎中	為福建右參政兼僉事○陞兵部郎中
_	沈演	孫如游掌南京翰林院印○福建參政	乞致仕許之
_	沈演	給副都御史陳禹謨林欲廈 參政魏說	郎中倫肇修秦繼宗等各誥?明神宗
_	沈演	萬章 以病乞歸許之 ○陞江西右參政	為按察使肇慶知府陳謨平陽知府傅
_	沈演	府沈自彰為關西道副使山西按察使	為福建右布政四川副使彭自新為雲
_	沈演	○甲寅陞福建布政使司右布政	為陝西布政使司左布政戶部郎中熊
```

◉ 人名 ○ 職官 ○ 地名 ○ 機關 ☑ 使用已標註資料 0211002-明實錄 010000016 載入資料

Clear Submit

Score	Label	關鍵字	前文	後文
81.38%	_	沈演	鄧思啟為雲南右參政○陞禮部郎中	為福建右參政兼僉事○陞兵部郎中
93.50%	_	沈演	孫如游掌南京翰林院印○福建參政	乞致仕許之
4.67%	_	沈演	給副都御史陳禹謨林欲廈 參政魏說	郎中倫肇修秦繼宗等各誥?明神宗
80.87%	_	沈演	萬章 以病乞歸許之 ○陞江西右參政	為按察使肇慶知府陳謨平陽知府傅
92.36%	_	沈演	府沈自彰為關西道副使山西按察使	為福建右布政四川副使彭自新為雲
95.45%	_	沈演	○甲寅陞福建布政使司右布政	為陝西布政使司左布政戶部郎中熊
89.20%	_	沈演	兵馬世龍尚方劍○陞陝西左布政使	為順天府尹

- http://sky.iis.sinica.edu.tw:9003/

Ancient Chinese NER

N	020100001	北京	八月己巳，以應天為南京，開封為	。庚午，徐達入元都，封府庫圖籍	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
N	020100001	京師	暴，遣將巡古北口諸隘。壬申，以	火，四方水旱，詔中書省集議便民	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
Y	020100001	北京	在詔內者，有司具以聞。壬午，幸	。改大都路曰北平府。徵元故臣。	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
Y	020100001	北京	懷慶，澤、潞相繼下。丁丑，至自	。戊寅，以元都平，詔天下。十一	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
N	020100001	京師	元主曰順帝。癸酉，買的里八剌至	，隣臣請獻俘。帝曰：「武王伐殷	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
N	020100001	京師	德下成都，四川平。乙丑，明昇至	，封歸義侯。八月甲午，免中都、	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
N	020100001	京師	。戊辰，詔百官奔父母喪不俟報。	地震。丁丑，免應天、太平、寧國	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
N	020100001	京師	丁丑，有事於圜丘。十二月戊子，	地震。甲寅，遣使振蘇州、湖州、	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
N	020100001	京師	年租賦，悉免之。」夏四月庚戌，	自去年八月不雨，是日始雨。五月	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
N	020100001	京師	南，□賜死。徵天下博學老成之士至	。是年，占城、爪哇、暹羅、日本	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
N	020100001	京師	子，詔求明經老成之士，有司禮送	。庚辰，河決原武、祥符、中牟。	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
N	020100001	京師	二月丙申，初命天下學校歲貢士於	。三月甲辰，召征南師還，沐英留	史／正史／明史／本紀 凡二十四卷／卷三 本紀第三／太祖 朱元璋
N	020100001	京師	王橚有罪，遷雲南，尋罷徙，留居	。定遠侯王弼等練兵山西、河南、	史／正史／明史／本紀 凡二十四卷／卷三 本紀第三／太祖 朱元璋
N	020100001	京師	久旱錄囚。秋七月庚子，徙富民實	。辛丑，免畿內官田租之半。八月	史／正史／明史／本紀 凡二十四卷／卷三 本紀第三／太祖 朱元璋
N	020100001	京師	半。八月乙卯，秦王樉有罪，召還	。乙丑，皇太子巡撫陝西。乙亥，	史／正史／明史／本紀 凡二十四卷／卷三 本紀第三／太祖 朱元璋
N	020100001	京師	僉事茅鼎討平之。庚戌，皇太子還	，晉王棡來朝。辛亥，振河南水災	史／正史／明史／本紀 凡二十四卷／卷三 本紀第三／太祖 朱元璋
N	020100001	京師	。二月戊午，召曹國公李景隆等還	。靖寧侯葉昇等練兵於河南及臨、	史／正史／明史／本紀 凡二十四卷／卷三 本紀第三／太祖 朱元璋
N	020100001	京師	服，毋妨嫁娶。諸王臨國中，毋至	。諸不在令中者，推此令從事。」	史／正史／明史／本紀 凡二十四卷／卷三 本紀第三／太祖 朱元璋
N	020100001	京師	四人充採訪使，□分巡天下。甲午，	地震，求直言。夏四月，湘王柏自	史／正史／明史／本紀 凡二十四卷／卷四 本紀第四／恭閔帝 朱允
N	020100001	京師	等叛降燕。壬辰，谷王橞自宣府奔	。長興侯耿炳文為征虜大將軍，駙	史／正史／明史／本紀 凡二十四卷／卷四 本紀第四／恭閔帝 朱允
N	020100001	京師	不克，引去。召遼王植、寧王權歸	，權不至，詔削護隣。丁卯，曹國	史／正史／明史／本紀 凡二十四卷／卷四 本紀第四／恭閔帝 朱允
N	020100001	京師	朝。帝為罷齊泰、黃子澄官，仍留	。	史／正史／明史／本紀 凡二十四卷／卷四 本紀第四／恭閔帝 朱允炆／建文元年 02020240001
N	020100001	京師	月甲申，召故周王橚於蒙化，居之	。燕兵連陷東阿、東平、汶上、兗	史／正史／明史／本紀 凡二十四卷／卷四 本紀第四／恭閔帝 朱允
N	020100001	京師	孫即位，遺詔諸王臨國中，毋得至	。王自北平入奔喪，聞詔乃止。時	史／正史／明史／本紀 凡二十四卷／卷五 本紀第五／成祖 朱棣
N	020100001	京師	已。無何，中官被黜者來奔，具言	空虛可取狀。王乃慨然曰：「頻年	史／正史／明史／本紀 凡二十四卷／卷五 本紀第五／成祖 朱棣
N	020100001	京師	真，則淮、鳳自震。我耀兵江上，	孤危，必有內變。」諸將皆曰善。	史／正史／明史／本紀 凡二十四卷／卷五 本紀第五／成祖 朱棣
Y	020100001	北京	、代王桂、岷王楩舊封。以北平為	。癸巳，保定侯孟善鎮遼東。丁酉	史／正史／明史／本紀 凡二十四卷／卷六 本紀第六／成祖 朱棣
Y	020100001	北京	平羌將軍，鎮甘肅。二月庚戌，設	留守行後軍都督府、行部、國子監	史／正史／明史／本紀 凡二十四卷／卷六 本紀第六／成祖 朱棣

Ancient Chinese NER

Entity ID Prefix

2	N	020100001	北京	八月己巳，以應天為南京，開封為	。庚午，徐達入元都，封府庫圖籍	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
3	N	020100001	京師	暴，遣將巡古北口諸隘。壬申，以	火，四方水旱，詔中書省集議便民	史／正史／明史／ ／卷二 本紀第二／太祖 朱元璋
4	Y	020100001	北京	在詔內者，有司具以聞。壬午，幸	。改大都路曰北平府。徵元故臣。	史／正史／明史／ ／卷二 本紀第二／太祖 朱元璋
5	Y	020100001	北京	懷慶，澤、潞相繼下。丁丑，至自	。戊寅，以元都平，詔天下。十一	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
6	N	020100001	京師	元主曰順帝。癸酉，買的里八剌至	，隣臣請獻俘。帝曰：「武王伐殷	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
7	N	020100001	京師	德下成都，四川平。乙丑，明昇至	，封歸義侯。八月甲午，免中都、	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
8	N	020100001	京師	，戊辰，詔百官奔父母喪不俟報。	地震。丁丑，免應天、太平、寧國	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
9	N	020100001	京師	丁丑，有事於圜丘。十二月戊子，	地震。甲寅，遣使振蘇州、湖州、	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
10	N	020100001	京師	年租賦，悉免之。」夏四月庚戌，	自去年八月不雨，是日始雨。五月	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
11	N	020100001	京師	南，▯賜死。徵天下博學老成之士至	。 是年，占城、爪哇、暹羅、日本	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
12	N	020100001	京師	子，詔求明經老成之士，有司禮送	。庚辰，河決原武、祥符、中牟。	史／正史／明史／本紀 凡二十四卷／卷二 本紀第二／太祖 朱元璋
13	N	020100001	京師	二月丙申，初命天下學校歲貢士於	。三月甲辰，召征南師還，沐英留	史／正史／明史／本紀 凡二十四卷／卷三 本紀第三／太祖 朱元璋

Sample

Label Mention Suffix

Ancient Chinese NER

- Entity Types
  - Person, officer, location, organization
  - Can be decided by entity ID
- Goal
  - Sample typing

Ancient Chinese NER

- Extract usable data

  1. Include all samples

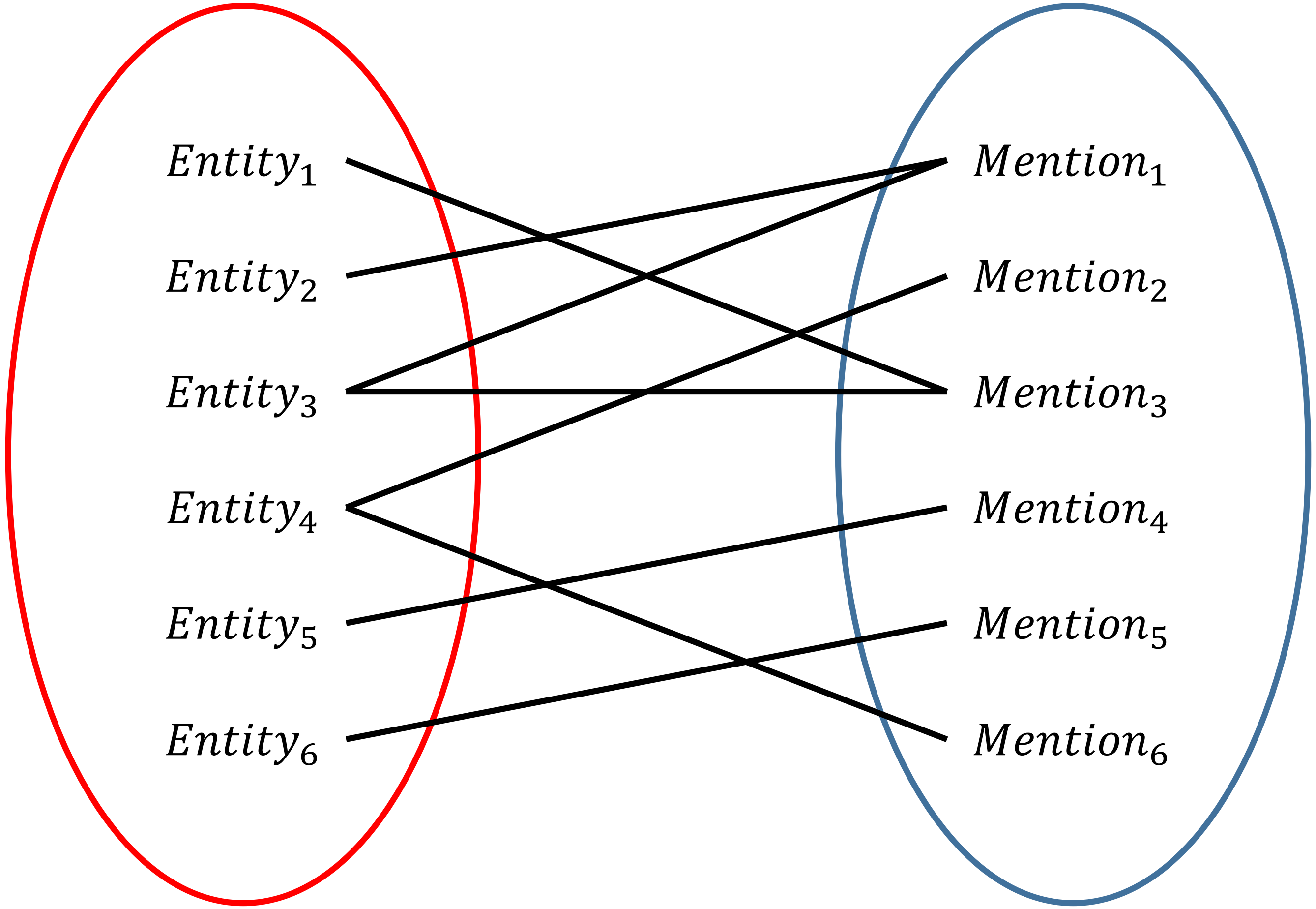
  2. Compute entity-mention bipartite map

  3. Remove the samples of which labels are not Y or N

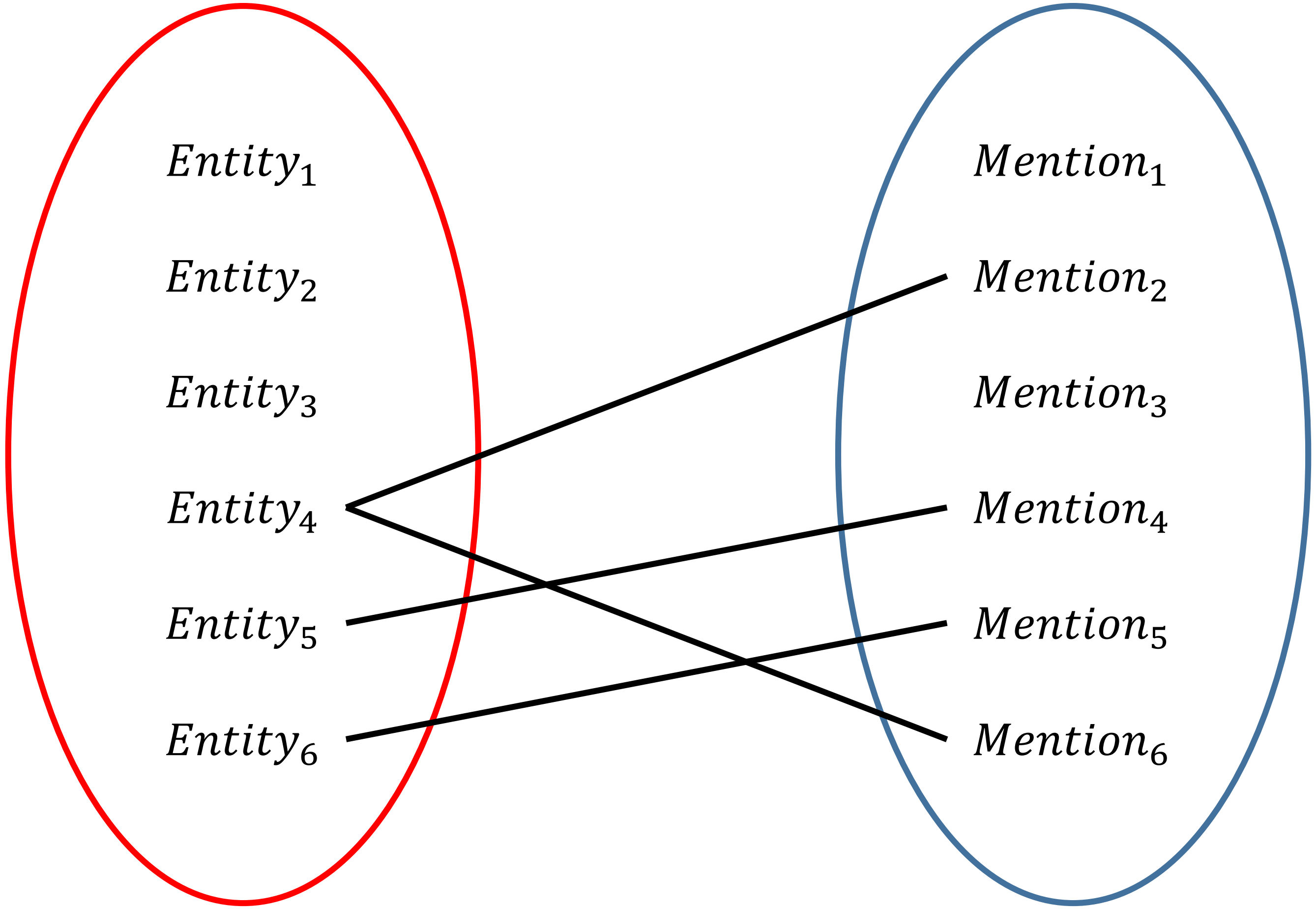
  4. Exclude the mentions that do not map to one single entity

  5. Exclude the entities of which labels are not human-verified

Ancient Chinese NER



Ancient Chinese NER

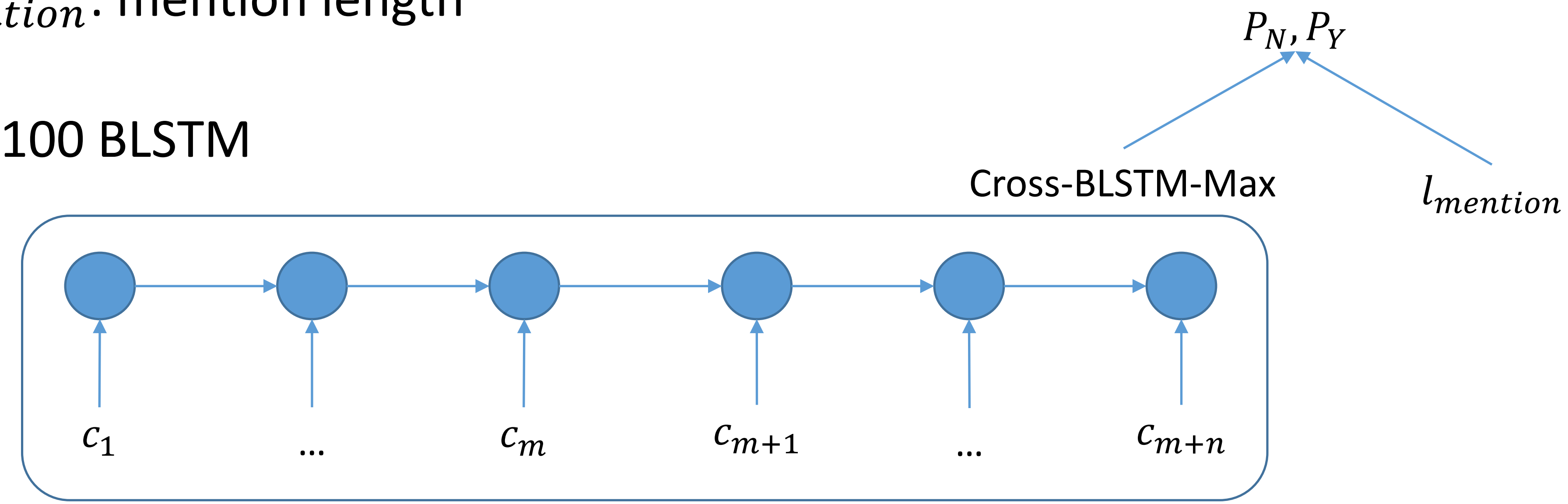


Ancient Chinese NER

Dataset		Unique Entities	Unique Mentions	Samples	Characters	Y(%)
Person	Train	5,238	6,106	473,766	15,160,443	75.60
	Validate	1,559	1,627	156,396	5,006,014	72.91
	Test	2,047	2,205	157,241	5,047,811	71.32
Officer	Train	12	44	5,600	186,210	92.96
	Validate	8	9	1,861	62,306	98.93
	Test	6	6	213	6,796	97.65
Location	Train	62	94	129,059	4,063,204	83.20
	Validate	40	48	36,369	1,131,185	84.52
	Test	52	68	38,927	1,218,336	93.25
Organization	Train	215	258	117,813	3,668,503	97.76
	Validate	47	49	40,958	1,325,299	97.80
	Test	63	68	35,047	1,124,819	98.29

Ancient Chinese NER

- m prefix characters, n suffix characters
- c_i : character + 2 features indicating prefix/suffix
- $l_{mention}$: mention length
- 100-100 BLSTM



Ancient Chinese NER

- Character embedding
  - Random initialization
  - Alphabet: training set occurrences≥30

Corpus	Training Set Characters	Alphabet Size
Person	15,160,443	4,311
Officer	186,210	763
Location	4,063,204	2,836
Organization	3,668,503	2,415

Ancient Chinese NER

Dataset		Unique Entities	Unique Mentions	Samples	Characters	Y (%)	Accuracy (%)
Person	Train	5,238	6,106	473,766	15,160,443	75.60	-
	Validate	1,559	1,627	156,396	5,006,014	72.91	**86.11**
	Test	2,047	2,205	157,241	5,047,811	71.32	**87.91**
Officer	Train	12	44	5,600	186,210	92.96	-
	Validate	8	9	1,861	62,306	**98.93**	**98.93**
	Test	6	6	213	6,796	**97.65**	**97.65**
Location	Train	62	94	129,059	4,063,204	83.20	-
	Validate	40	48	36,369	1,131,185	84.52	**85.59**
	Test	52	68	38,927	1,218,336	**93.25**	83.91
Organization	Train	215	258	117,813	3,668,503	97.76	-
	Validate	47	49	40,958	1,325,299	**97.80**	**97.80**
	Test	63	68	35,047	1,124,819	**98.29**	**98.29**

Outline

- Named Entity Recognition
    - Task
    - Features
    - Related Work
- Leveraging Linguistic Structures for NER
    - Joint parsing and NER
    - Tree-LSTM for NER
    - Mitigating inconsistencies between parsing and NER
- Constructing Deep Cross Bi-LSTM with Self-Attention for NER
    - Deep Cross Bi-LSTM
    - Multi-head self-attention
- CKIP NER
    - Chinese NER
    - Ancient Chinese Document NER