AAAI 2019

Jointly Extracting Multiple Triplets with Multilayer Translation Constraints

Zhen Tan[1], Xiang Zhao[1,3], WeiWang[2] and Weidong Xiao[1,3]

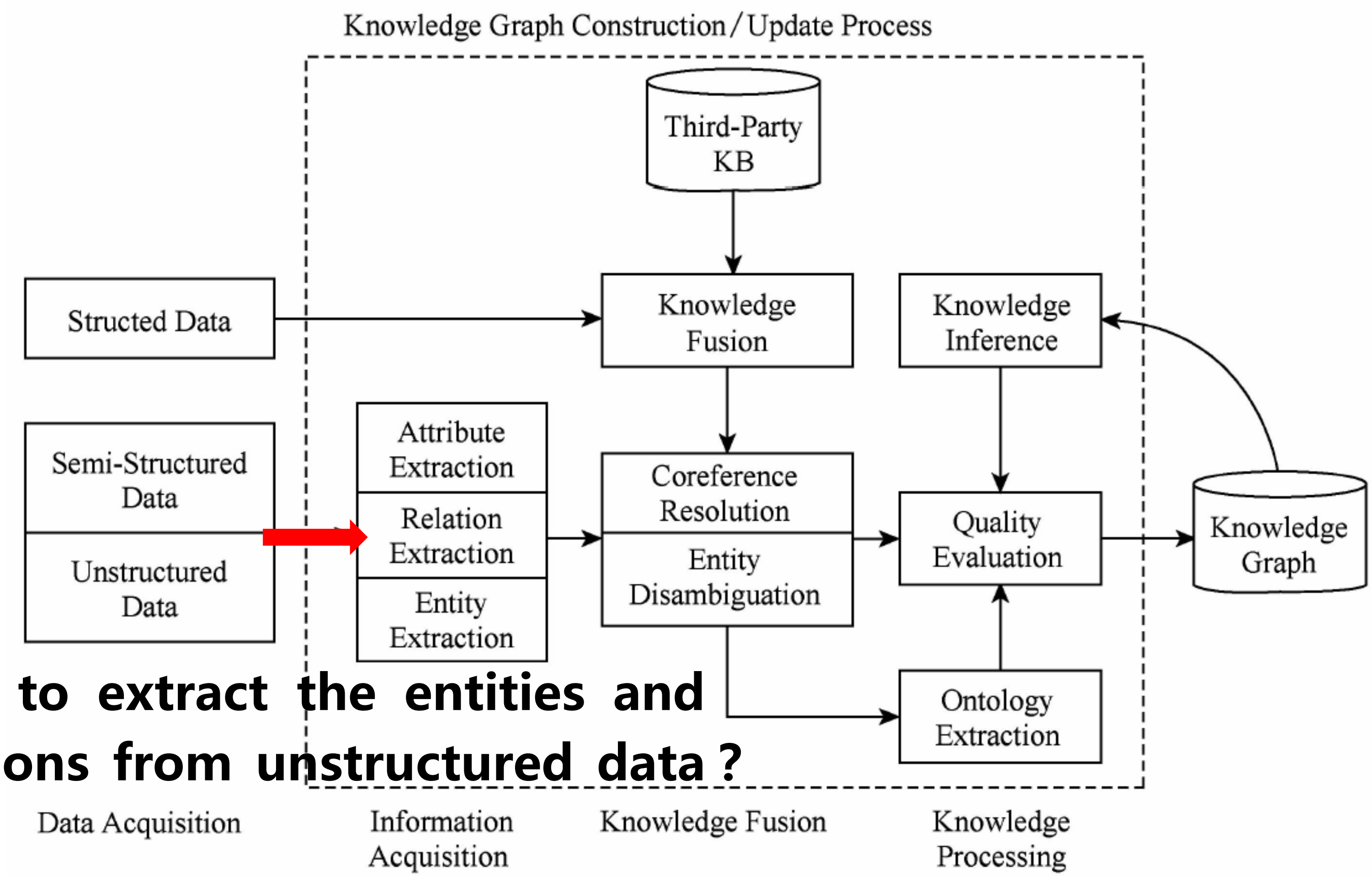
1 National University of Defense Technology, China.

2 The University of New SouthWales, Australia.

3 Collaborative Innovation Center of Geospatial Technology, China.

Jointly extraction

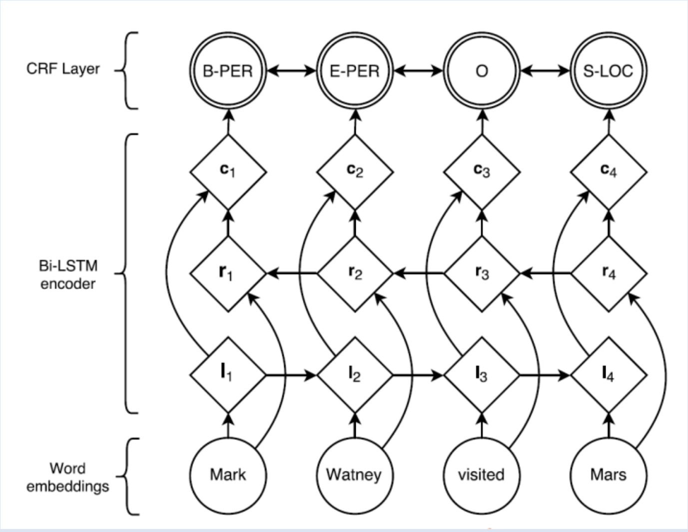
Framework



How to extract the entities and relations from unstructured data ?

1.entity

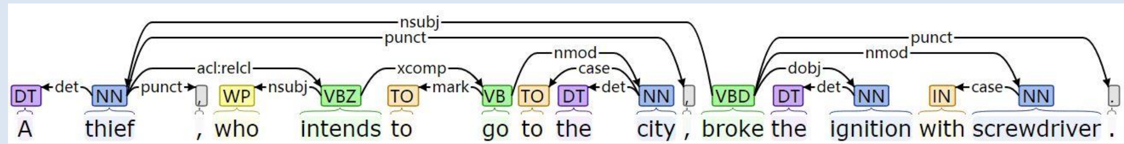
The rule , DNN and CRF are leveraged to extract the entities or attributes in the unstructured data.



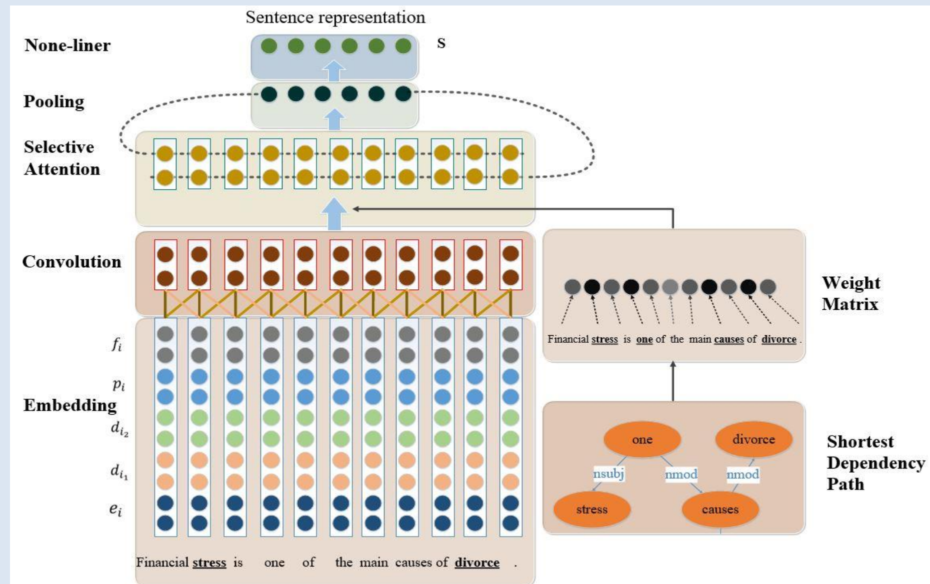
Bi-LSTM+CRF

Neural Architectures for Named Entity Recognition, NAACL2016

2.relation



- CNN or RNN
- +Dependency path
- +Attention Mechanism
- +variety Features



Error propagation!

Jointly extraction

NER+RC

- LSTM
- CNN
- Encode-Decode

too many candidate entities

low joint extraction accuracy

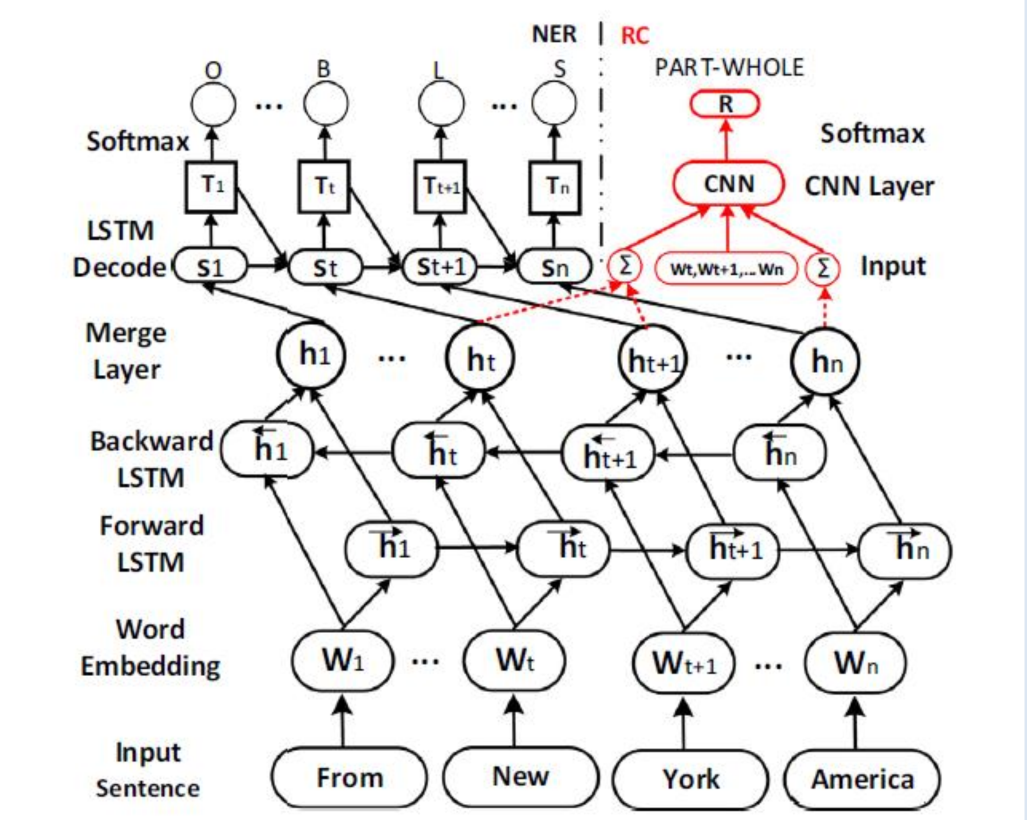
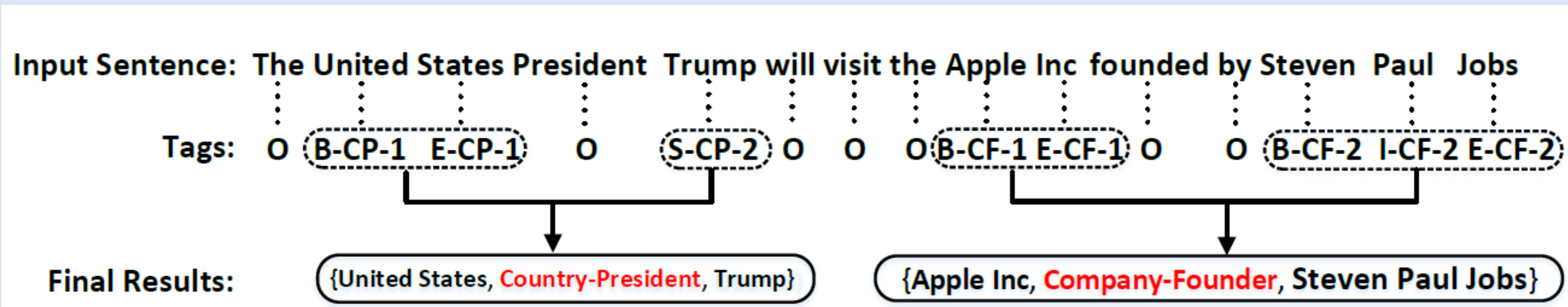


Fig. 1. The framework of the hybrid neural network for jointly extracting entities and relations.

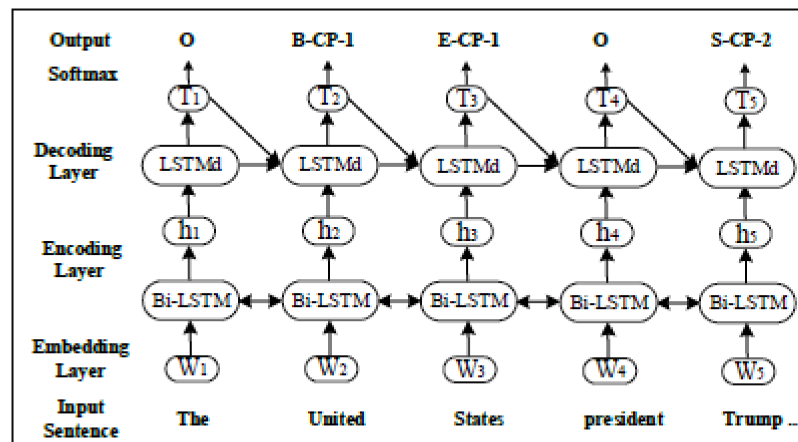
Joint entity and relation extraction based on a hybrid neural network, Neurocomputing

Jointly extraction

NTS



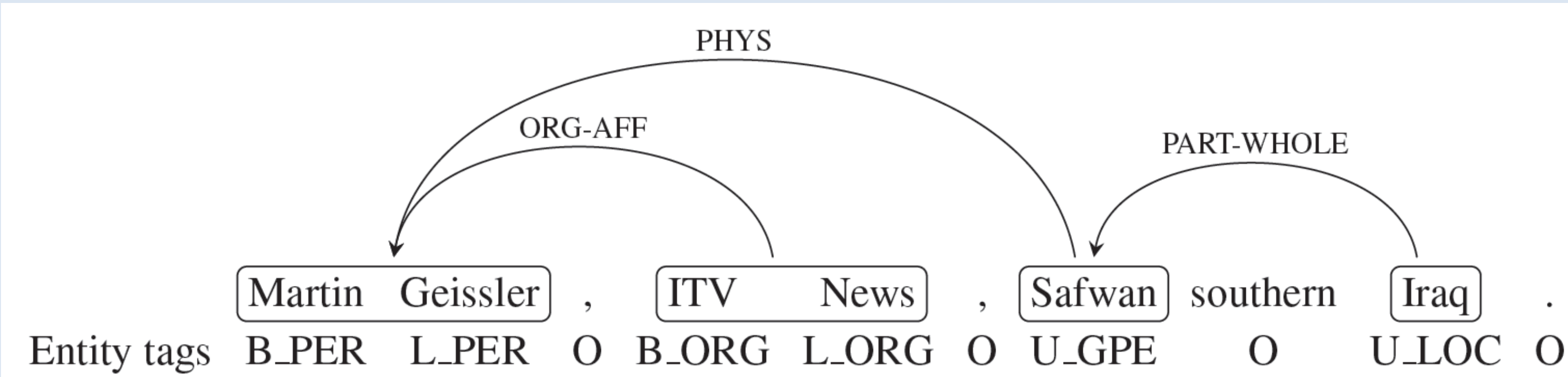
- Encode-Decode
- Tagging Scheme

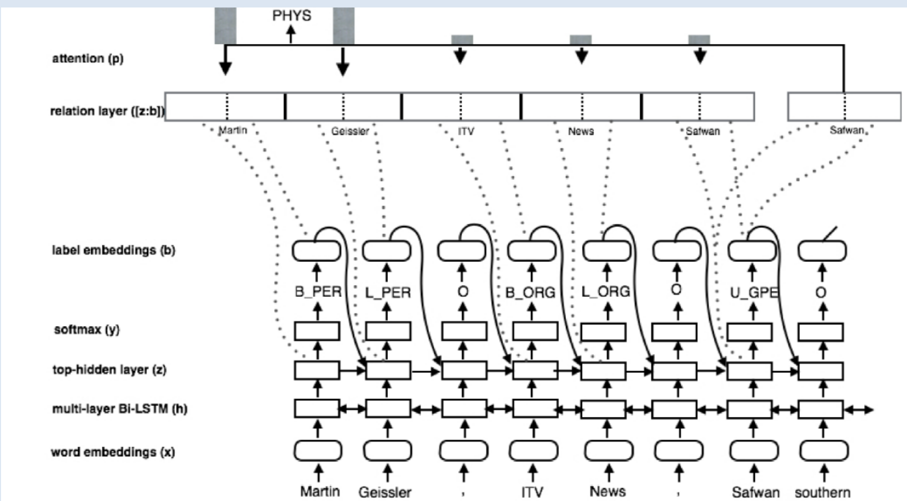


(a) The End-to-End Model

Each entity pair can only have a relationship

Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme, ACL 2017





- Self-attention
- Tagging Scheme

Each entity can only be related to the previous entity

Going out on a limb: Joint Extraction of Entity Mentions and Relations without Dependency Trees, ACL 2017

shortcoming

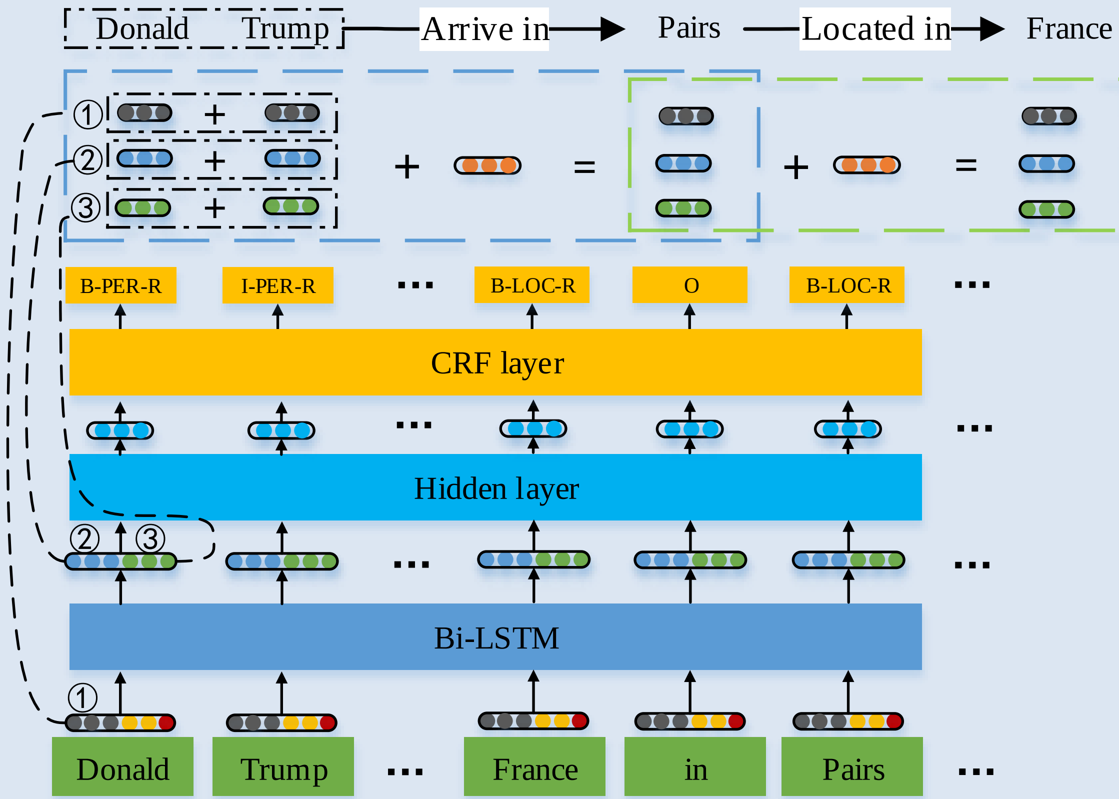
How to extract multiple triplets (entities and relations) from unstructured data ?

- Pipeline extraction leads to error propagation and low accuracy
- Existing joint extraction models are not suited for extracting multiple triples in sentences

Our work

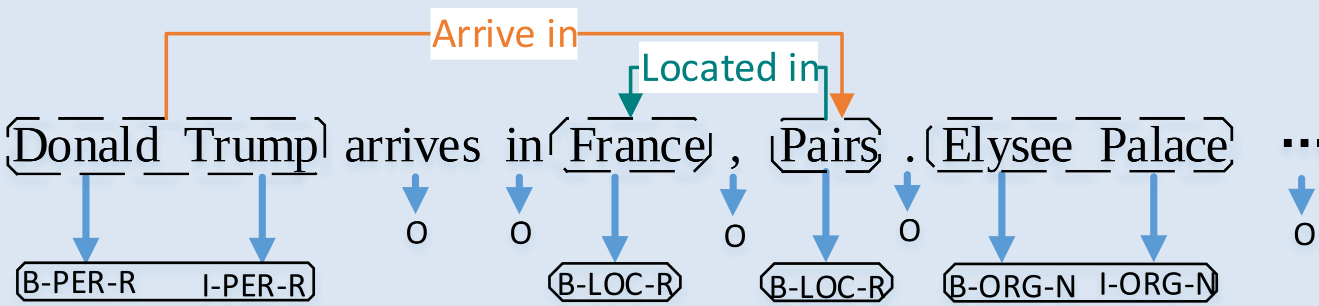
joint extraction model for multi-triplets TME

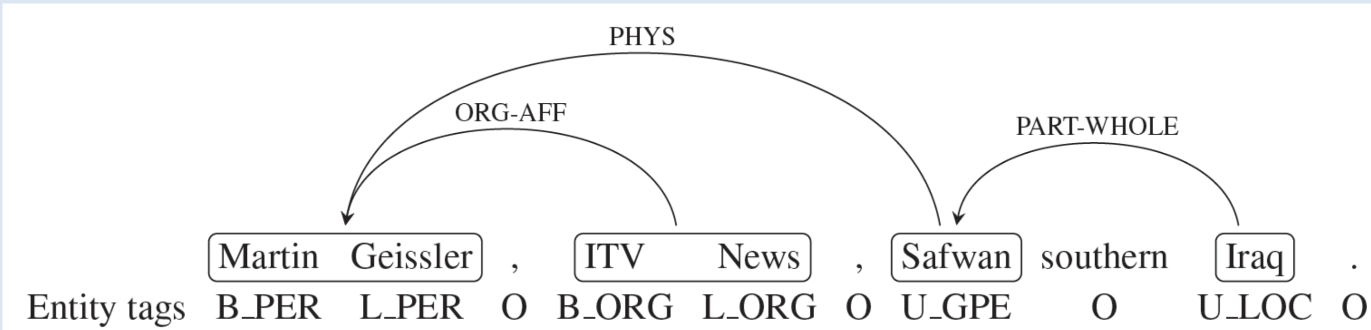
- Bi-LSTM
- TTS
- Multi-layer translation-based constrains

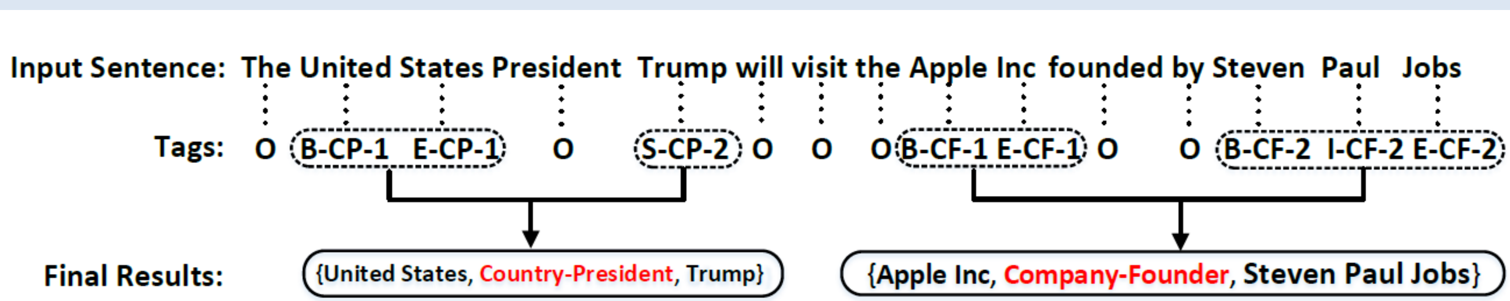


TTS

- Position Part
- Type Part
- Relation Part





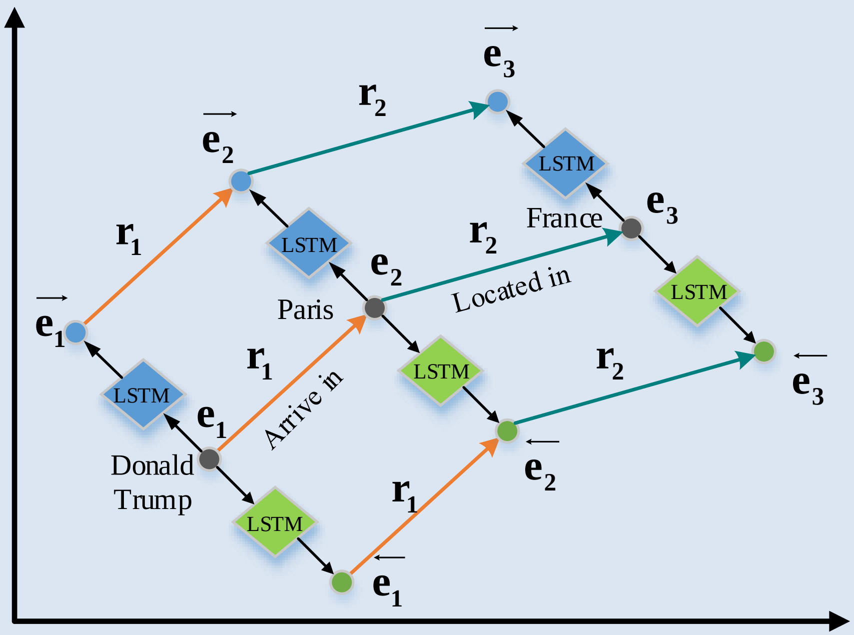


Jointly extraction

$$\mathbf{e}=\sum_{k=i}^{i+e_l}\mathbf{w}_k,\ \overrightarrow{\mathbf{e}}=\sum_{k=i}^{i+e_l}\overrightarrow{\mathbf{h}}_k,\ \overleftarrow{\mathbf{e}}=\sum_{k=i}^{i+e_l}\overleftarrow{\mathbf{h}}_k,$$

$$f(\tau)=-\|\mathbf{e}_1+\mathbf{r}-\mathbf{e}_2\|_2^2.$$

$$\overrightarrow{f}(\tau)=-\|\overrightarrow{\mathbf{e}_1}+\mathbf{r}-\overrightarrow{\mathbf{e}_2}\|_2^2,$$
$$\overleftarrow{f}(\tau)=-\|\overleftarrow{\mathbf{e}_1}+\mathbf{r}-\overleftarrow{\mathbf{e}_2}\|_2^2.$$



Jointly extraction

Training

- Entity loss function

$$\mathcal{L}_e = \log(p(\mathbf{y}|\mathbf{X})) = f(\mathbf{X}, \mathbf{y}) - \log(\sum_{y \in Y} e^{f(\mathbf{X}, \tilde{\mathbf{y}})}).$$

- Relation loss function

$$\mathcal{L}_{em} = \sum_{\tau \in \mathcal{T}} \sum_{\tau' \in \mathcal{T}'} \mathtt{ReLu}(f(\tau) + \gamma - f(\tau')),$$

$$\overrightarrow{\mathcal{L}} = \sum_{\tau \in \mathcal{T}} \sum_{\tau' \in \mathcal{T}'} \mathtt{ReLu}(\overrightarrow{f}(\tau) + \gamma - \overrightarrow{f}(\tau')),$$

$$\overleftarrow{\mathcal{L}} = \sum_{\tau \in \mathcal{T}} \sum_{\tau' \in \mathcal{T}'} \mathtt{ReLu}(\overleftarrow{f}(\tau) + \gamma - \overleftarrow{f}(\tau')).$$

$$\mathcal{L}_r = \mathcal{L}_{em} + \overrightarrow{\mathcal{L}} + \overleftarrow{\mathcal{L}}.$$

- Final loss function

$$\mathcal{L} = \mathcal{L}_e + \lambda \mathcal{L}_r,$$

Jointly extraction

Extraction

- Get the label of each word
- Generate candidate entities
- Generate candidate entity pairs
- Calculate the score for each candidate entity pair
- Select one triple for each entity pair
- Sort the candidate triples and select the triplet with the highest score as the correct triple.

$$\hat{y} = \arg\max_{\tilde{y} \in Y} f(\mathbf{X}, \tilde{y}).$$

$$\hat{\mathcal{E}} = \{\hat{e}_1, \ldots, \hat{e}_i, \ldots, \hat{e}_m\}.$$

$$\tilde{\mathcal{T}} = \{\tilde{\tau} = (\hat{e}_i, \hat{e}_j, r) | r \in \mathcal{R}\},$$

$$f_c(\tilde{\tau}) = f(\tilde{\tau}) + \overrightarrow{f}(\tilde{\tau}) + \overleftarrow{f}(\tilde{\tau}),$$

$$\hat{\tau} = \arg\max_{\tilde{\tau} \in \tilde{\mathcal{T}}} f_c(\tilde{\tau}).$$

Jointly extraction

results

表 4.1 数据集统计信息

数据集	#Train	#Test	#Triplet	#Ent	#Rel
NYT-single	235,983	395	17,663	67,148	24
NYT-multi	63,602	1,000	17,494	25,894	24

表 4.2 在数据集 **NYT-single** 上的三元组抽取效果

Methods	Prec	Rec	F1
FCM	0.553	0.154	0.240
DS+logistic	0.258	0.393	0.311
LINE	0.335	0.329	0.332
MultiR	0.338	0.327	0.333
DS-Joint	0.574	0.256	0.354
CoType	0.423	0.511	0.463
NTS-Joint	**0.615**	0.414	0.495
TME (Top-1)	0.583	0.485	**0.530**
TME (Top-2)	0.515	0.508	0.511
TME (Top-3)	0.458	**0.522**	0.489

表 4.3 在数据集 **NYT-multi** 上的三元组抽取效果

Methods	Prec	Rec	F1
CoType	0.385	0.340	0.361
NTS-Joint	0.533	0.336	0.412
TME-MR	0.638	0.421	0.507
TME-RR	0.423	0.452	0.437
TME-NS	0.558	0.496	0.525
TME (Top-1)	**0.749**	0.436	0.551
TME (Top-2)	0.696	0.478	**0.567**
TME (Top-3)	0.631	**0.500**	0.558

Ablation

表 4.4 NYT-multi 数据集上 TME 模型的成分分析

Model	Top-1			Top-2			Top-3		
	Prec	Rec	F1	Prec	Rec	F1	Prec	Rec	F1
TME-MR	0.692	0.385	0.495	0.638	0.421	0.507	0.575	0.438	0.498
TME-RR	0.478	0.417	0.445	0.423	0.452	0.437	0.365	0.462	0.408
TME-NS	0.687	0.419	0.520	0.558	0.496	0.525	0.448	0.523	0.483
TME	**0.749**	**0.436**	**0.551**	**0.696**	0.478	**0.567**	**0.631**	0.500	**0.558**
-TTS (-TP)	0.741	0.436	0.549	0.680	0.478	0.561	0.610	0.498	0.548
-TTS (-RP)	0.610	0.376	0.465	0.488	0.484	0.486	0.400	**0.547**	0.462
-TTS (-TP-RP)	0.575	0.353	0.438	0.474	0.468	0.470	0.391	0.531	0.450
-Character	0.723	0.428	0.538	0.663	0.472	0.552	0.597	0.497	0.542
-CRF	0.690	0.414	0.517	0.608	0.470	0.530	0.522	0.495	0.509
-$\overrightarrow{f}$ - $\overleftarrow{f}$	0.552	0.310	0.398	0.521	0.368	0.431	0.468	0.399	0.431
-f	0.569	0.332	0.419	0.518	0.372	0.433	0.465	0.395	0.428
-Dropout	0.723	0.424	0.535	0.666	0.478	0.556	0.593	0.503	0.544
-Pretrain	0.686	0.411	0.514	0.613	0.466	0.530	0.539	0.495	0.516

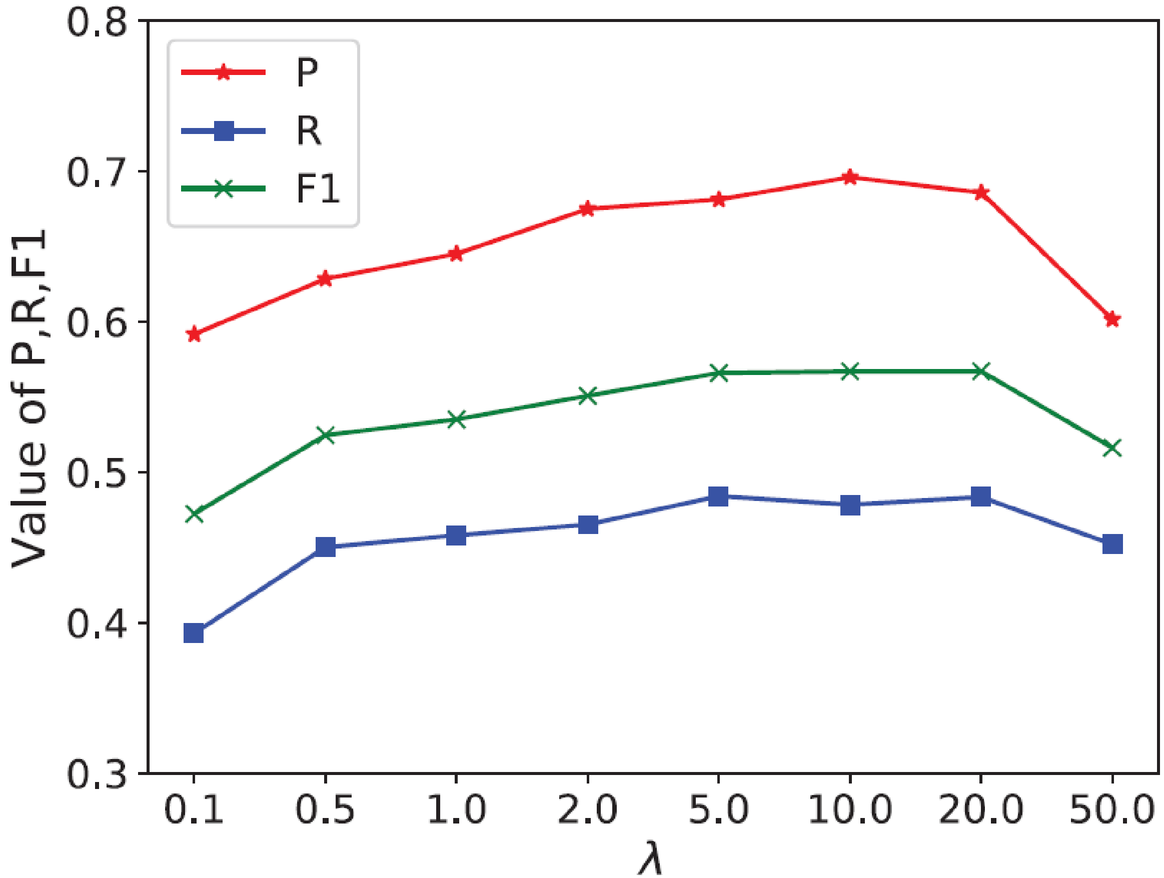
Jointly extraction

Case Study

表 4.5　TME（Top-3）模型在 NYT-multi 数据集上的样例分析

Sentence I	...President Jacques Chirac$_{[PER]}$ of France$_{[LOC]}$ and Chancellor Angela Merkel$_{[PER]}$ of Germany$_{[LOC]}$ to press for agreement on a Security Council resolution demanding that *Iran*$_{[LOC]}$ stop ...
(Jacques Chirac, nationality, France) (Angela Merkel, nationality, Germany)	(Jacques Chirac, nationality, France) (Angela Merkel, nationality, Germany) (Jacques Chirac, nationality, Germany)
Sentence II	... grasping the critical need for the *United States*$_{[LOC]}$ to get Afghanistan$_{[LOC]}$ right, she moved to Kandahar$_{[LOC]}$ to help... Afghans for Civil Society, founded by the brother of Hamid Karzai$_{[PER]}$...
(Afghanistan, contains, Kandahar) (Hamid Karzai, place_of_birth, Kandahar) (Hamid Karzai, nationality, Afghanistan)	(Kandahar, contains, Hamid Karzai) (Afghanistan, contains, Kandahar) (Hamid Karzai, nationality, Afghanistan)
Sentence III	... Across Iraq$_{[LOC]}$, from Mosul$_{[LOC]}$ and Ramadi$_{[LOC]}$ to Basra$_{[LOC]}$ and Kirkuk$_{[LOC]}$, the lines of votes hummed with excitement, and with the hope that a permanent Iraqi government...
(Iraq, contains, Mosul) (Iraq, contains, Ramadi) (Iraq, contains, Basra) (Iraq, contains, Kirkuk)	(Iraq, contains, Mosul) (Iraq, contains, Basra) (Iraq, contains, Ramadi)

H-Parameter



$$\mathcal{L} = \mathcal{L}_e + \lambda \mathcal{L}_r,$$

Future

- Unable to resolve the same entity pair containing two different relationships (Located in and Capital)
- Unable to automatically recognize the number of entities in a sentence

Thanks！

Q&A！