Double Graph Based Reasoning for Document-level Relation Extraction

Shuang Zeng, Runxin Xu, Baobao Chang, Lei Li

EMNLP 2020

Task

- Relation extraction in document level

Elias Brown
[1] *Elias Brown* (*May 9, 1793*– *July 7, 1857*) was a ***U.S.*** Representative from ***Maryland***. [2] Born near ***Baltimore***, ***Maryland***, *Brown* attended the common schools. … [7] He died near ***Baltimore***, ***Maryland***, and is interred in a private cemetery near ***Eldersburg***, ***Maryland***.

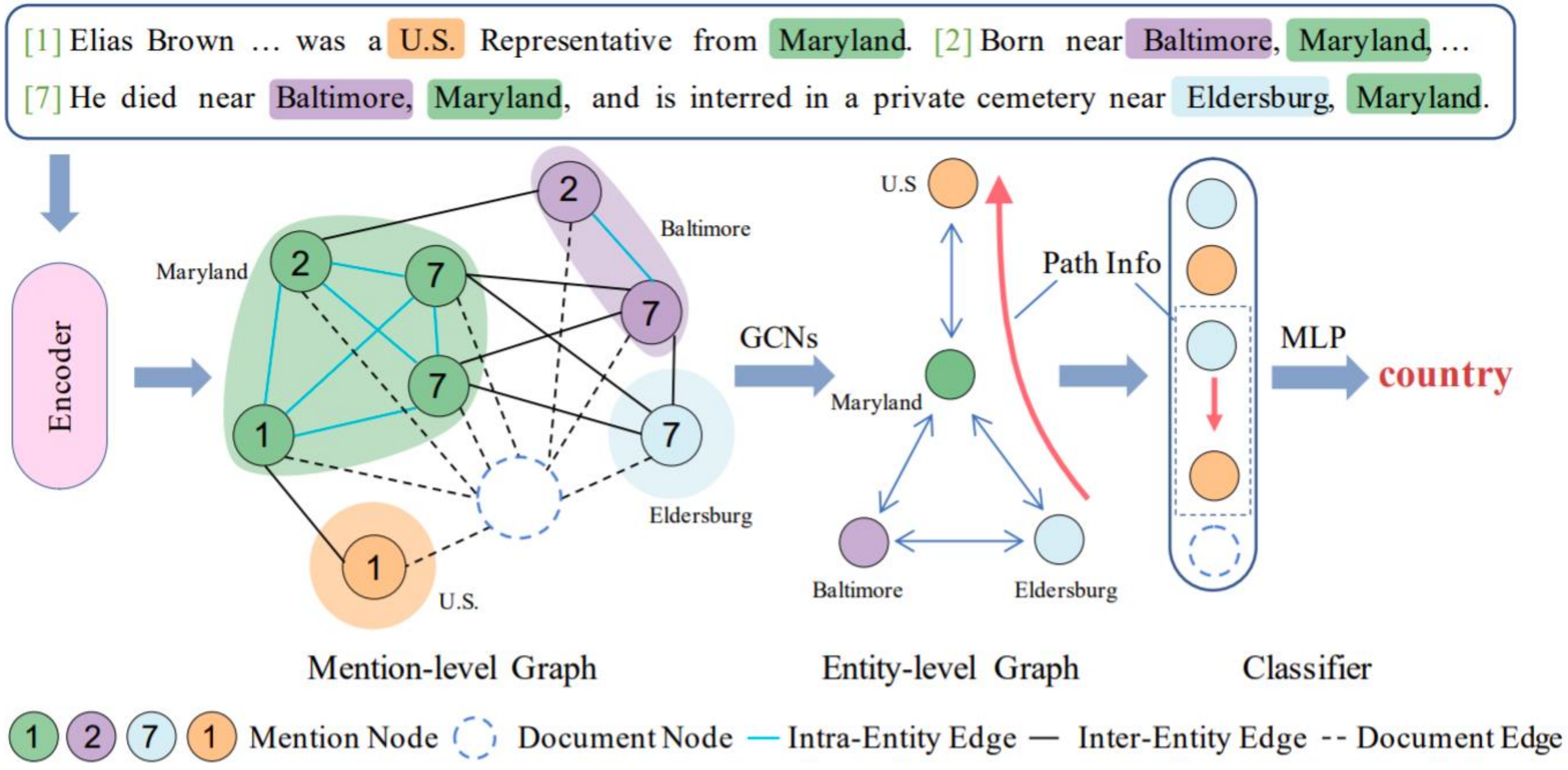
Subject: ***Maryland***
Object: ***U.S.***
relation: **country**

Subject: ***Baltimore***; ***Eldersburg***
Object: ***Maryland***
relation: **located in the administrative territorial entity**

Subject: ***Baltimore***; ***Eldersburg***
Object: ***U.S.***
relation: **country**

Motivation

- Infer document structure
  - Improve mention representation
  - Perform two-hop reasoning on entity graph



Encoder

- Encode words by word embedding, entity type and coreference embedding:

$$x_i = [E_w(w_i); E_t(t_i); E_c(c_i)]$$

- Encode the sequential order of the words:

$$[g_1, g_2, \ldots, g_n] = Encoder([x_1, x_2, \ldots, x_n])$$

Mention-level Graph Aggregation Module

- Nodes:
  - Entity mentions
  - Document node
- Edges:
  - Intra-entity edge: All entity mentions are connected to each other
  - Inter-entity edge: Entity mentions of two entities that cooccur in a sentence are connected to each other
  - Document Edge: The document node is connected to all edges
- GCN is applied over the constructed graph

$$\mathbf{m}_u = [h_u^{(0)}; h_u^{(1)}; \ldots; h_u^{(N)}]$$

Entity-level Graph Inference Module

- Use entity mentions to represent the entity nodes:

$$\mathbf{e}_i = \frac{1}{N}\sum_n \mathbf{m}_n$$

- Represent edges from entities:

$$\mathbf{e}_{ij} = \sigma\left(W_q[\mathbf{e}_i;\mathbf{e}_j] + b_q\right)$$

- Path reasoning:

$$\mathbf{p}^i_{h,t} = [\mathbf{e}_{ho};\mathbf{e}_{ot};\mathbf{e}_{to};\mathbf{e}_{oh}]$$

$$s_i = \sigma([\mathbf{e}_h;\mathbf{e}_t]\cdot W_l \cdot \mathbf{p}^i_{h,t})$$

$$\alpha_i = \frac{e^{s_i}}{\sum_j e^{s_j}}$$

$$\mathbf{p}_{h,t} = \sum_i \alpha_i \mathbf{p}^i_{h,t}$$

Classification

- Prediction based on the vector:

$$I_{h,t} = [\mathbf{e}_h; \mathbf{e}_t; |\mathbf{e}_h - \mathbf{e}_t|; \mathbf{e}_h \odot \mathbf{e}_t; \mathbf{m}_{doc}; \mathbf{p}_{h,t}]$$

- Prediction:

$$P(r|\mathbf{e}_h, \mathbf{e}_t) = sigmoid\left(W_b \sigma(W_a I_{h,t} + b_a) + b_b\right)$$

- Loss:

$$\mathcal{L} = -\sum_{\mathcal{D} \in \mathcal{S}} \sum_{h \neq t} \sum_{r_i \in \mathcal{R}} \mathbb{I}\left(r_i = 1\right) \log P\left(r_i|\mathbf{e}_h, \mathbf{e}_t\right)$$
$$+ \mathbb{I}\left(r_i = 0\right) \log\left(1 - P\left(r_i|\mathbf{e}_h, \mathbf{e}_t\right)\right)$$

Results

Model	Dev				Test	
	Ign F1	Ign AUC	F1	AUC	Ign F1	F1
CNN* (Yao et al., 2019)	41.58	36.85	43.45	39.39	40.33	42.26
LSTM* (Yao et al., 2019)	48.44	46.62	50.68	49.48	47.71	50.07
BiLSTM* (Yao et al., 2019)	48.87	47.61	50.94	50.26	48.78	51.06
Context-Aware* (Yao et al., 2019)	48.94	47.22	51.09	50.17	48.40	50.70
HIN-GloVe* (Tang et al., 2020)	51.06	-	52.95	-	51.15	53.30
GAT‡ (Velickovic et al., 2017)	45.17	-	51.44	-	47.36	49.51
GCNN‡ (Sahu et al., 2019)	46.22	-	51.52	-	49.59	51.62
EoG‡ (Christopoulou et al., 2019)	45.94	-	52.15	-	49.48	51.82
AGGCN‡ (Guo et al., 2019)	46.29	-	52.47	-	48.89	51.45
LSR-GloVe* (Nan et al., 2020)	48.82	-	55.17	-	52.15	54.18
GAIN-GloVe	**53.05**	**52.57**	**55.29**	**55.44**	**52.66**	**55.08**
BERT-RE$^{*}_{base}$ (Wang et al., 2019a)	-	-	54.16	-	-	53.20
RoBERTa-RE$^{\dagger}_{base}$	53.85	48.27	56.05	51.35	53.52	55.77
BERT-Two-Step$^{*}_{base}$ (Wang et al., 2019a)	-	-	54.42	-	-	53.92
HIN-BERT$^{*}_{base}$ (Tang et al., 2020)	54.29	-	56.31	-	53.70	55.60
CorefBERT-RE$^{*}_{base}$ (Ye et al., 2020)	55.32	-	57.51	-	54.54	56.96
LSR-BERT$^{*}_{base}$ (Nan et al., 2020)	52.43	-	59.00	-	56.97	59.05
GAIN-BERT$_{base}$	**59.14**	**57.76**	**61.22**	**60.96**	**59.00**	**61.24**
BERT-RE$^{*}_{large}$ (Ye et al., 2020)	56.67	-	58.83	-	56.47	58.69
CorefBERT-RE$^{*}_{large}$ (Ye et al., 2020)	56.73	-	58.88	-	56.48	58.70
RoBERTa-RE$^{*}_{large}$ (Ye et al., 2020)	57.14	-	59.22	-	57.51	59.62
CorefRoBERTa-RE$^{*}_{large}$ (Ye et al., 2020)	57.84	-	59.93	-	57.68	59.91
GAIN-BERT$_{large}$	**60.87**	**61.79**	**63.09**	**64.75**	**60.31**	**62.76**

Ablation Study

Model	Dev				Test	
	Ign F1	Ign AUC	F1	AUC	Ign F1	F1
GAIN-GloVe	**53.05**	**52.57**	**55.29**	**55.44**	**52.66**	**55.08**
- *hMG*	50.97	48.84	53.10	51.73	50.76	53.06
- *Inference Module*	50.84	48.68	53.02	51.58	50.32	52.66
- *Document Node*	50.86	48.68	53.01	52.46	50.32	52.67
GAIN-BERT$_{base}$	**59.14**	**57.76**	**61.22**	**60.96**	**59.00**	**61.24**
- *hMG*	57.12	51.54	59.17	54.61	57.31	59.56
- *Inference Module*	56.97	54.29	59.28	57.25	57.01	59.34
- *Document Node*	57.26	52.07	59.62	55.51	57.01	59.63

Inter and Intra-relations

Model	Intra-F1	Inter-F1
CNN*	51.87	37.58
LSTM*	56.57	41.47
BiLSTM*	57.05	43.49
Context-Aware*	56.74	42.26
LSR-GloVe*	60.83	48.35
GAIN-GloVe	**61.67**	**48.77**
- *hMG*	59.72	46.49
BERT-RE$^{*}_{base}$	61.61	47.15
RoBERTa-RE$_{base}$	65.65	50.09
BERT-Two-Step$^{*}_{base}$	61.80	47.28
LSR-BERT$^{*}_{base}$	65.26	52.05
GAIN-BERT$_{base}$	**67.10**	**53.90**
- *hMG*	66.15	51.42

Inference Relations

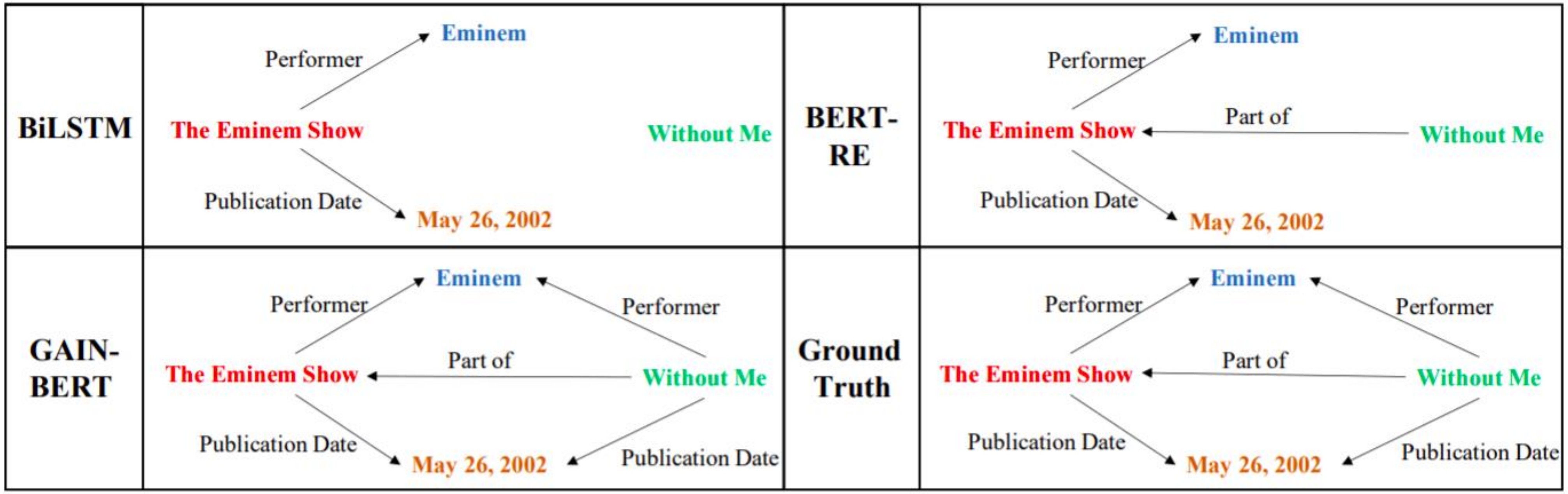
- Relations that need inference

$$e_h \xrightarrow{r_1} e_o \xrightarrow{r_2} e_t \qquad e_h \xrightarrow{r_3} e_t$$

Model	Infer-F1	P	R
CNN	37.11	32.81	42.72
LSTM	39.03	33.16	47.44
BiLSTM	38.73	31.60	50.01
Context-Aware	39.73	**33.97**	47.85
GAIN-GloVe	**40.82**	32.76	**54.14**
- *Inference Module*	39.76	32.26	51.80
BERT-RE$_{base}$	39.62	34.12	47.23
RoBERTa-RE$_{base}$	41.78	37.97	46.45
GAIN-BERT$_{base}$	**46.89**	**38.71**	**59.45**
- *Inference Module*	45.11	36.91	57.99

Case Study

[1] ***The Eminem Show*** is the fourth studio album by American rapper ***Eminem***, released on ***May 26, 2002*** by Aftermath Entertainment, Shady Records, and Interscope Records.
[2] ***The Eminem Show*** includes the commercially successful singles "***Without Me***", "Cleanin' Out My Closet", "Superman", and "Sing for the Moment"....



Thanks