Connecting the Dots: Document-level Neural Relation Extraction with Edge-oriented Graphs

Fenia Christopoulou, Makoto Miwa, Sophia Ananiadou
efstathia.christopoulou@manchester.ac.uk

Introduction

Task: Document-level, intra- & inter-sentence Relation Extraction between concepts

Motivation: (i) Model a document as a graph structure (Quirk and Poon, 2017)
(ii) Relations depend on different contexts: unique edge representations
(iii) Edges: modelled via heuristics (natural document-level associations)

Idea: Model both intra- & inter-sentence relations between concept entities using inference via Edge-oriented Graphs

Example

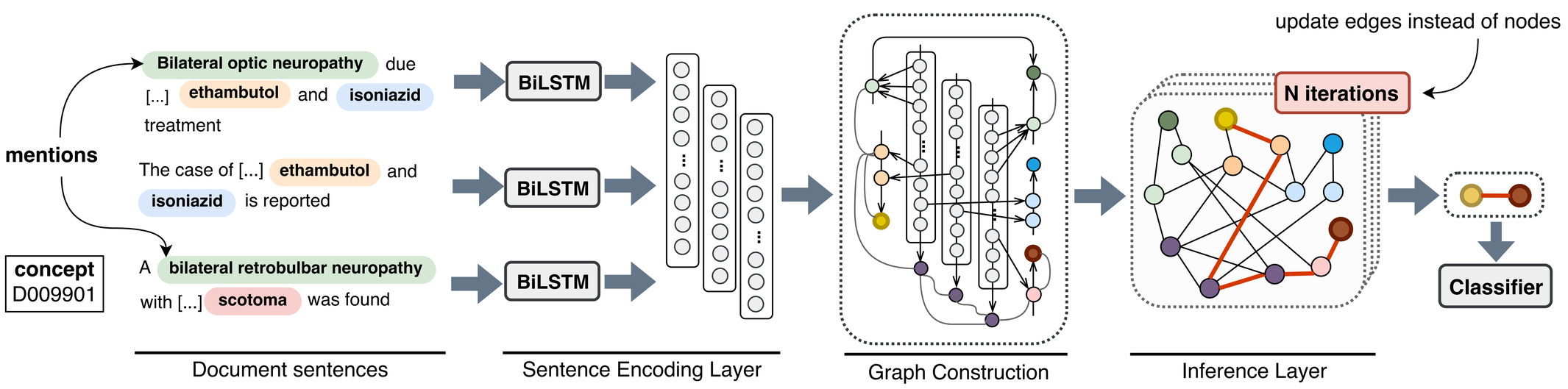
Bilateral optic neuropathy due to combined **ethambutol** and **isoniazid** treatment . The case of a 40 - year - old patient who underwent an unsuccessful cadaver kidney transplantation and was treated with **ethambutol** and **isoniazid** is reported . A **bilateral retrobulbar neuropathy** with an unusual central bitemporal hemianopic **scotoma** was found .

··· inter-sentence, — intra-sentence

EoG code

Walks code

Model Architecture



Nodes & Edges

Node mentions ($\mathbf{n}_m$) | entities ($\mathbf{n}_e$) | sentences ($\mathbf{n}_s$) = average over tokens | mentions | words

Initial edge representations:

- Mention-Mention: $\mathbf{x}_{\text{MM}} = [\mathbf{n}_{m_i}; \mathbf{n}_{m_j}; \mathbf{c}_{m_i,m_j}; \mathbf{d}_{m_i,m_j}]$ → Occur in the same sentence
- Mention-Sentence: $\mathbf{x}_{\text{MS}} = [\mathbf{n}_m; \mathbf{n}_s]$ → Mention occurs in sentence
- Mention-Entity: $\mathbf{x}_{\text{ME}} = [\mathbf{n}_m; \mathbf{n}_e]$ → Mention associated with Entity (pre-defined)
- Sentence-Sentence: $\mathbf{x}_{\text{SS}} = [\mathbf{n}_{s_i}; \mathbf{n}_{s_j}; \mathbf{d}_{s_i,s_j}]$ → Direct: distance = 1, Indirect: distance > 1
- Entity-Sentence: $\mathbf{x}_{\text{ES}} = [\mathbf{n}_e; \mathbf{n}_s]$ → At least one mention occurs in sentence

$\mathbf{d}$: distance embedding, $\mathbf{c}$: context embedding

Inference & Classification

- Generate/Update edge representations via walks in the graph (Christopoulou et al., 2018)

$$f\left(\mathbf{e}_{ik}^{(l)}, \mathbf{e}_{kj}^{(l)}\right) = \sigma\left(\mathbf{e}_{ik}^{(l)} \odot \left(\mathbf{W}\, \mathbf{e}_{kj}^{(l)}\right)\right),\ l \text{ edge length}$$

$$\mathbf{e}_{ij}^{(2l)} = \beta\, \mathbf{e}_{ij}^{(l)} + (1-\beta) \sum_{k \neq i,j} f\left(\mathbf{e}_{ik}^{(l)}, \mathbf{e}_{kj}^{(l)}\right),\ \beta \in [0,1]$$

- Classify Entity-Entity (EE) generated edge representations

$$\mathbf{y} = \text{softmax}\left(\mathbf{W}\, \mathbf{e}_{\text{EE}}^{(L)} + \mathbf{b}\right),\ L = 2^N,\ N \text{ iterations}$$

Results

	Method [CDR]	F1 (%) Overall	Intra	Inter
Independent	Gu et al. (2017)	61.3	57.2	11.7
	Verga et al. (2018)	62.1	-	-
	Nguyen and Verspoor (2018)	62.3	-	-
	EoG	**63.6**	**68.2**	**50.9**
	EoG (`Full`)	57.6	66.5	39.4
	EoG (`NoInf`)	49.2	60.2	30.6
	EoG (`Sent`)	55.2	65.2	-
+ Data/Tools	Zhou et al. (2016)	61.3	-	-
	Peng et al. (2016)	63.1	-	-
	Li et al. (2016b)	67.7	58.9	-
	Panyam et al. (2018)	60.3	65.1	45.7
	Zheng et al. (2018)	61.5	-	-

`Full`: Fully-connected graph
`NoInf`: No inference mechanism (walks)
`Sent`: Train on single sentences

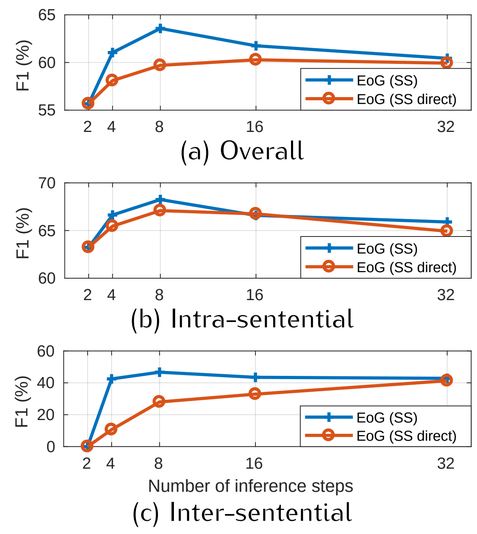
EoG outperforms even models with extra data/tools for intra- & inter-sentence pairs

CDR (Li et al., 2016a): Manually annotated, 1,500 PubMed abstracts, Chemical-Disease associations, binary

GDA (Wu et al., 2019): Distantly Supervised, 30,192 MEDLINE abstracts, Gene-Disease associations, binary

Effect of Edge Types

Edge Types	F1 (%) Overall	Intra	Inter
`EE`	55.14	61.31	40.34
EoG	63.57	68.25	46.68
−`MM`	62.77	67.93	46.65
−`ME`	61.57	66.39	45.40
−`MS`	62.92	67.55	44.74
−`ES`	61.41	66.44	43.04
−$\text{SS}_{\text{indirect}}$	59.70	67.09	28.00
−`SS`	57.41	65.45	1.59
−`MM,ME,MS`	60.46	66.07	39.56
−`ES,MS,SS`	56.86	64.63	0.00



- EE edges fail on intra-sentence pair detection
- Usage of direct sentence edges ($\text{SS}_{\text{direct}}$) requires more inference steps
- **Document-level associations are crucial for both intra- and inter- RE**
- Removal of M nodes results in low inter-sentence performance
- Removal of S nodes disables inter-sentence pair identification

Error Analysis

entities joined with conjunctions	Following short exposure to oral **prednisone** [...]. Both presented in the emergency room with profound **coma**, **hypotension**, severe **hyperglycemia**, and acidosis.
missing co-reference connections	The etiology of **pyeloureteritis cystica** has long been [...] The **disease** occurred subsequent to the initiation of **heparin** therapy [...]
incomplete entity linking	Time trends in **warfarin**-associated **hemorrhage**. [...] The proportion of patients with major and **intracranial bleeding** increased [...]

Sentence-to-sentence edges (SS) potentially simulate co-reference links, by encoding the co-referring entities into the sentence representation

Conclusions

- Document-level graphs can be created with heuristics
- **Intra- pairs can be supported by inter- associations (SS edges)**
- Longer inference steps are weaker than shorter ones
- S nodes enable the construction of shorter edge representations