Extracting Relational Facts by an End-to-End Neural Model with Copy Mechanism

Xiangrong Zeng, Daojian Zeng, Shizhu He, Kang Liu, Jun Zhao

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science

Changsha University of Science & Technology

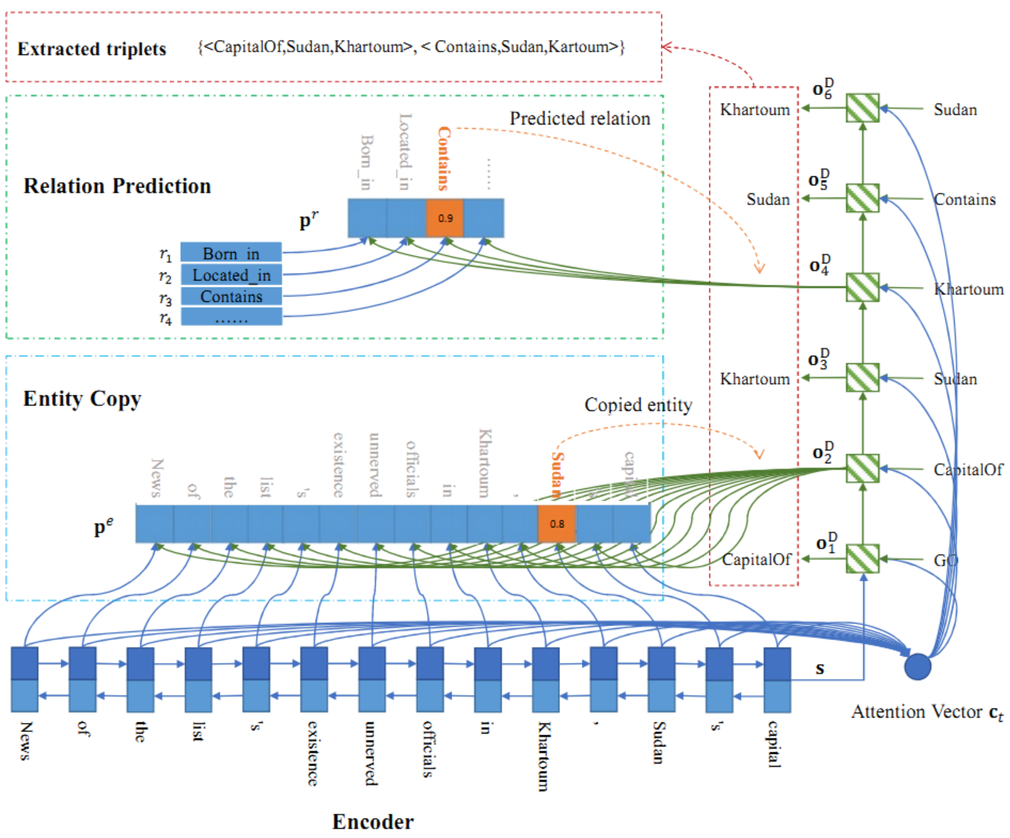
Motivation

- Relational triplets may have overlaps in a sentence.
- We divide the sentences into three types according to triplet overlap degree: Normal, EntityPairOverlap (EPO) and SingleEntityOverlap (SEO).
- Current work mainly concentrate on relation extraction of Normal class.

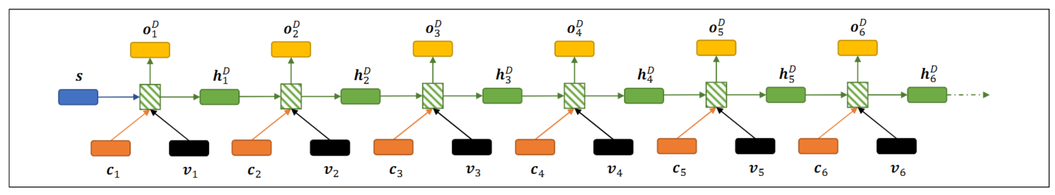
Normal	S1: Chicago is located in the United States. {<Chicago, country, United States>}	Chicago — country — United States
EPO	S2: Celery is an ingredient of Bakso, a dish from the country of Indonesia . {<Bakso, region, Indonesia>, <Bakso, country, Indonesia>}	Bakso — region / country — Indonesia
SEO	S3: Aarhus airport serves the city of Aarhus who's leader is Jacob Bundsgaard. {<Aarhus, leaderName, Jacob Bundsgaard>, <Aarhus Airport, cityServed, Aarhus>}	Jacob Bundsgaard — leaderName — Aarhus — cityServed — Aarhus Airport

Methods

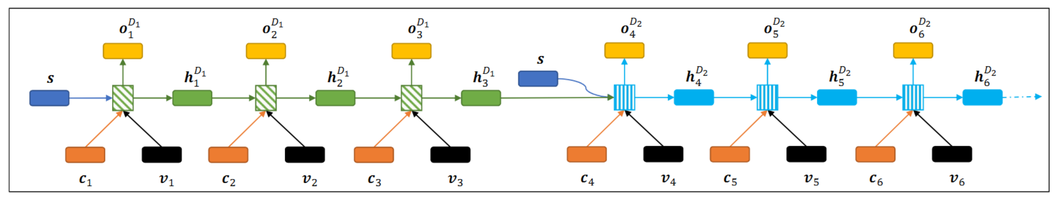
- We aim to design a model that could extract triplets from sentences of Normal, EPO and SEO classes.
- We propose an end2end model based on Seq2Seq learning with copy mechanism.
  - The encoder converts a natural language sentence (the source sentence) into a fixed length semantic vector
  - Then, the decoder reads in this vector and generates triplets directly.
    - To generate a triplet, firstly, the decoder generates the relation
    - Secondly, by adopting the copy mechanism, the decoder copies the first entity (head entity) from the source sentence
    - Lastly, the decoder copies the second entity (tail entity) from the source sentence.
- We adopt two different strategies in decoding process:
  - Employing only one unifies decoder (OneDecoder) to generate all triplets
  - Or, applying multiple separated decoders (MultiDecoder) and each of them generating one triplet.



- The inputs and outputs of the decoder(s) of OneDecoder model and MultiDecoder model.
  - (a) is the decoder of OneDecoder model. Only one decoder (the green rectangle with shadows) is used and this encoder is initialized with the sentence representation s.
  - (b) is the decoders of MultiDecoder model. There are two decoders (the green rectangle and blue rectangle with shadows). The first decoder is initialized with s; Other decoder(s) are initialized with s and previous decoder's state.



(a)



(b)

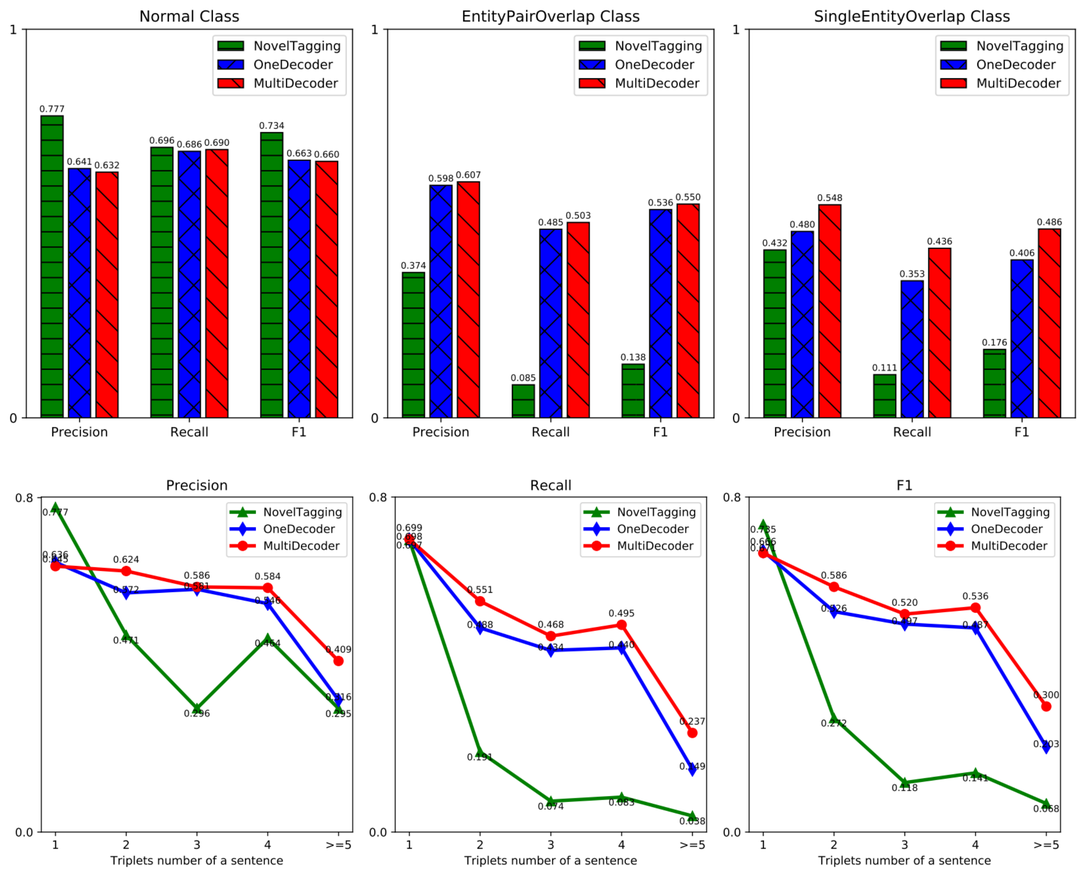
Experiments

- Datasets: NYT and WebNLG
- Baseline: NovelTagging (ACL2017)

Class	NYT		WebNLG	
	Train	Test	Train	Test
Normal	37013	3266	1596	246
EPO	9782	978	227	26
SEO	14735	1297	3406	457
ALL	56195	5000	5019	703

- Results of different models

Model	NYT			WebNLG		
	Precision	Recall	F1	Precision	Recall	F1
NovelTagging	**0.624**	0.317	0.420	**0.525**	0.193	0.283
OneDecoder	0.594	0.531	0.560	0.322	0.289	0.305
MultiDecoder	0.610	**0.566**	**0.587**	0.377	**0.364**	**0.371**



Model	NYT	WebNLG
OneDecoder	0.858	0.745
MultiDecoder	0.862	0.821

Table 3: F1 values of entity generation.

Model	NYT	WebNLG
OneDecoder	0.874	0.759
MultiDecoder	0.870	0.751

Table 4: F1 values of relation generation.

Contribution

- **We propose an end2end neural model based on sequence-to-sequence learning with copy mechanism to extract relational facts from sentences, where the entities and relations could be jointly extracted.**
- **Our model could consider the relational triplet overlap problem through copy mechanism. In our knowledge, the relational triplet overlap problem has never been addressed before.**
- **We conduct experiments on two public datasets. Experimental results show that our model outperforms the state-of-the-arts with 39.8% and 31.1% improvements respectively.**