Document-Level Event Role Filler Extraction using Multi-Granularity Contextualized Encoding

Xinya Du and Claire Cardie
Department of Computer Science
Cornell University

ACL2020

Document-level Event Extraction

- Given an article, the task requires:
  - + Role filler extraction.
  - + Noun phrase coreference resolution.
  - + Event Tracking.
- Dataset: Message Understanding Conference 3+4

Including 1,700 articles describing terrorist incidents.

Document-level Event Extraction

TST1-MUC3-0080

BOGOTA, 3 APR 90 (INRAVISION TELEVISION CADENA 1) -- [REPORT] [JORGE ALONSO SIERRA VALENCIA] [TEXT] LIBERAL SENATOR FEDERICO ESTRADA VELEZ WAS KIDNAPPED ON 3 APRIL AT THE CORNER OF 60TH AND 48TH STREETS IN WESTERN MEDELLIN, ONLY 100 METERS FROM A METROPOLITAN POLICE CAI [IMMEDIATE ATTENTION CENTER]. THE ANTIOQUIA DEPARTMENT LIBERAL PARTY LEADER HAD LEFT HIS HOUSE WITHOUT ANY BODYGUARDS ONLY MINUTES EARLIER. AS HE WAITED FOR THE TRAFFIC LIGHT TO CHANGE, THREE HEAVILY ARMED MEN FORCED HIM TO GET OUT OF HIS CAR AND GET INTO A BLUE RENAULT.

HOURS LATER, THROUGH ANONYMOUS TELEPHONE CALLS TO THE METROPOLITAN POLICE AND TO THE MEDIA, THE EXTRADITABLES CLAIMED RESPONSIBILITY FOR THE KIDNAPPING. IN THE CALLS, THEY ANNOUNCED THAT THEY WILL RELEASE THE SENATOR WITH A NEW MESSAGE FOR THE NATIONAL GOVERNMENT.

LAST WEEK, FEDERICO ESTRADA VELEZ HAD REJECTED TALKS BETWEEN THE GOVERNMENT AND THE DRUG TRAFFICKERS.

0. MESSAGE ID	TST1-MUC3-0080
1. TEMPLATE ID	1
2. DATE OF INCIDENT	03 APR 90
3. TYPE OF INCIDENT	KIDNAPPING
4. CATEGORY OF INCIDENT	TERRORIST ACT
5. PERPETRATOR: ID OF INDIV(S)	"THREE HEAVILY ARMED MEN"
6. PERPETRATOR: ID OF ORG(S)	"THE EXTRADITABLES" / "EXTRADITABLES"
7. PERPETRATOR: CONFIDENCE	CLAIMED OR ADMITTED: "THE EXTRADITABLES" / "EXTRADITABLES"
8. PHYSICAL TARGET: ID(S)	*
9. PHYSICAL TARGET: TOTAL NUM	*
10. PHYSICAL TARGET: TYPE(S)	*
11. HUMAN TARGET: ID(S)	"FEDERICO ESTRADA VELEZ" ("LIBERAL SENATOR" / "ANTIOQUIA DEPARTMENT LIBERAL PARTY LEADER" / "SENATOR" / "LIBERAL PARTY LEADER" / "PARTY LEADER")
12. HUMAN TARGET: TOTAL NUM	1
13. HUMAN TARGET: TYPE(S)	GOVERNMENT OFFICIAL / POLITICAL FIGURE: "FEDERICO ESTRADA VELEZ"
14. TARGET: FOREIGN NATION(S)	-
15. INSTRUMENT: TYPE(S)	*
16. LOCATION OF INCIDENT	COLOMBIA: MEDELLIN (CITY)
17. EFFECT ON PHYSICAL TARGET(S)	*
18. EFFECT ON HUMAN TARGET(S)	-

Document-level Roll Filler Extraction

[S1] ... by special urban troops, four terrorists have been arrested in soacha.

[S2] They are responsible for the car bomb attack on the Newspaper El Espectador, to a series of bogota dynamite attacks, to the freeing of a group of paid assassins.

[S3] The terrorists are also connected to the murder of Teofilo Forero Castro, ...

[S4] General Ramon is the commander of the 13th infantry brigade.

[S5] He said that at least two of those arrested have fully confessed to having taken part in the accident of Luis Carlos Galan Sarmiento in soacha, Cundinamarca.

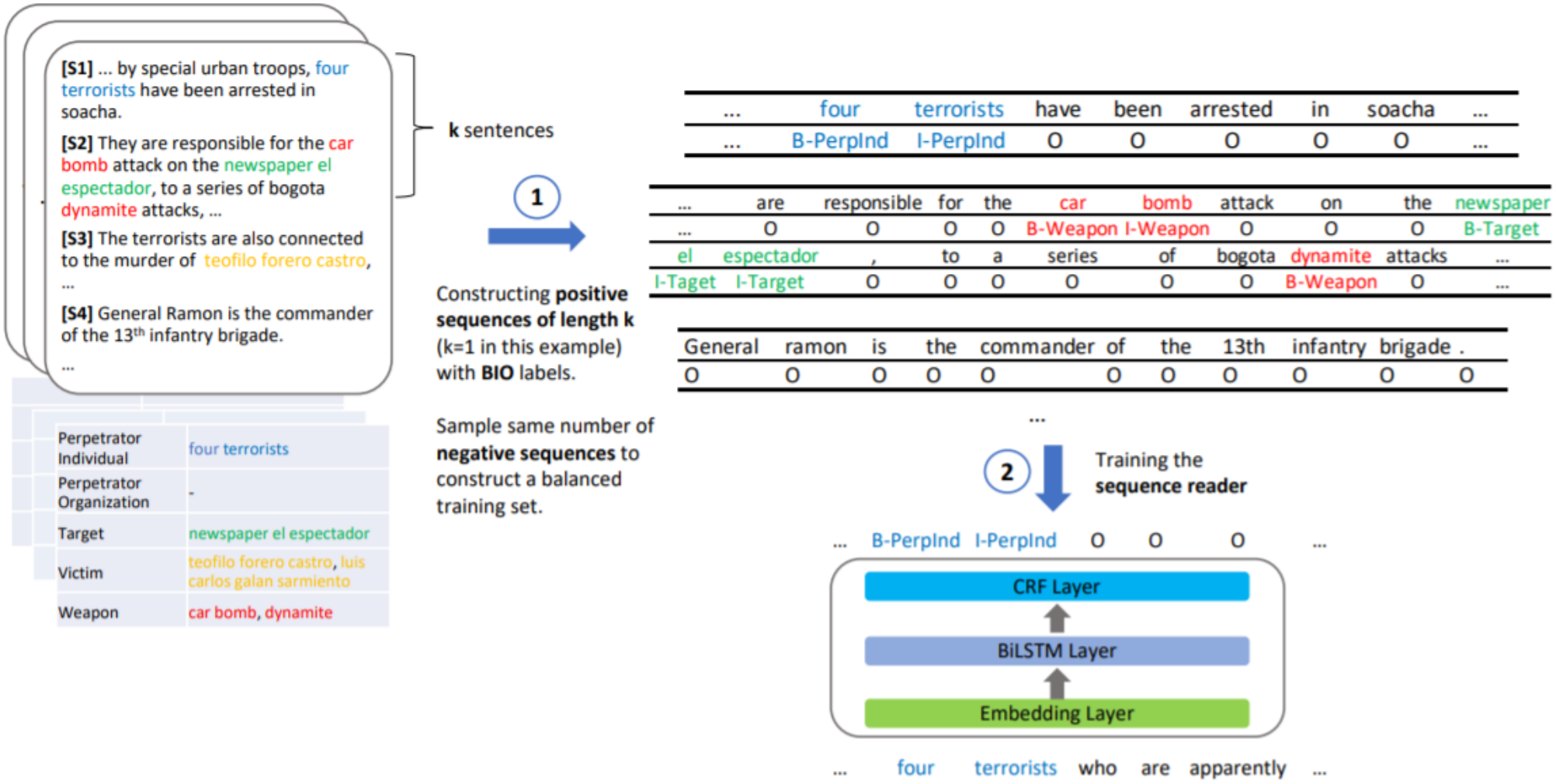
[S6] .. triumph over organized crime, its accomplices and its protectors.

...

Machine reader reads through the document

Perpetrator Individual	four terrorists
Perpetrator Organization	-
Target	Newspaper El Espectador
Victim	Teofilo Forero Castro, Luis Carlos Galan Sarmiento
Weapon	car bomb, dynamite

Create Training Data



Model: k-sentence Reader

- Embedding layer:

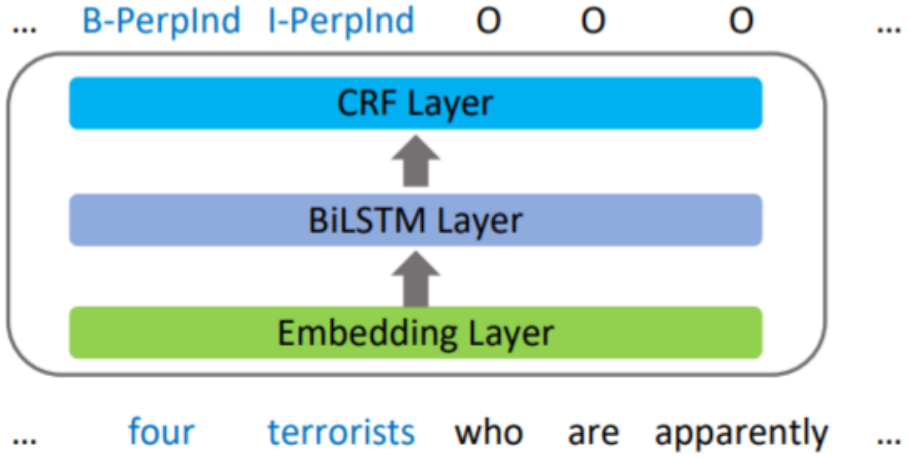
$$\mathbf{xe}_i = \mathbf{E}(x_i)$$
$$\mathbf{xb}_1, \mathbf{xb}_2, ..., \mathbf{xb}_m = \texttt{BERT}(x_1, x_2, ..., x_m)$$
$$\mathbf{x}_i = \texttt{concat}(\mathbf{xe}_i, \mathbf{xb}_i)$$

- BiLSTM layer:

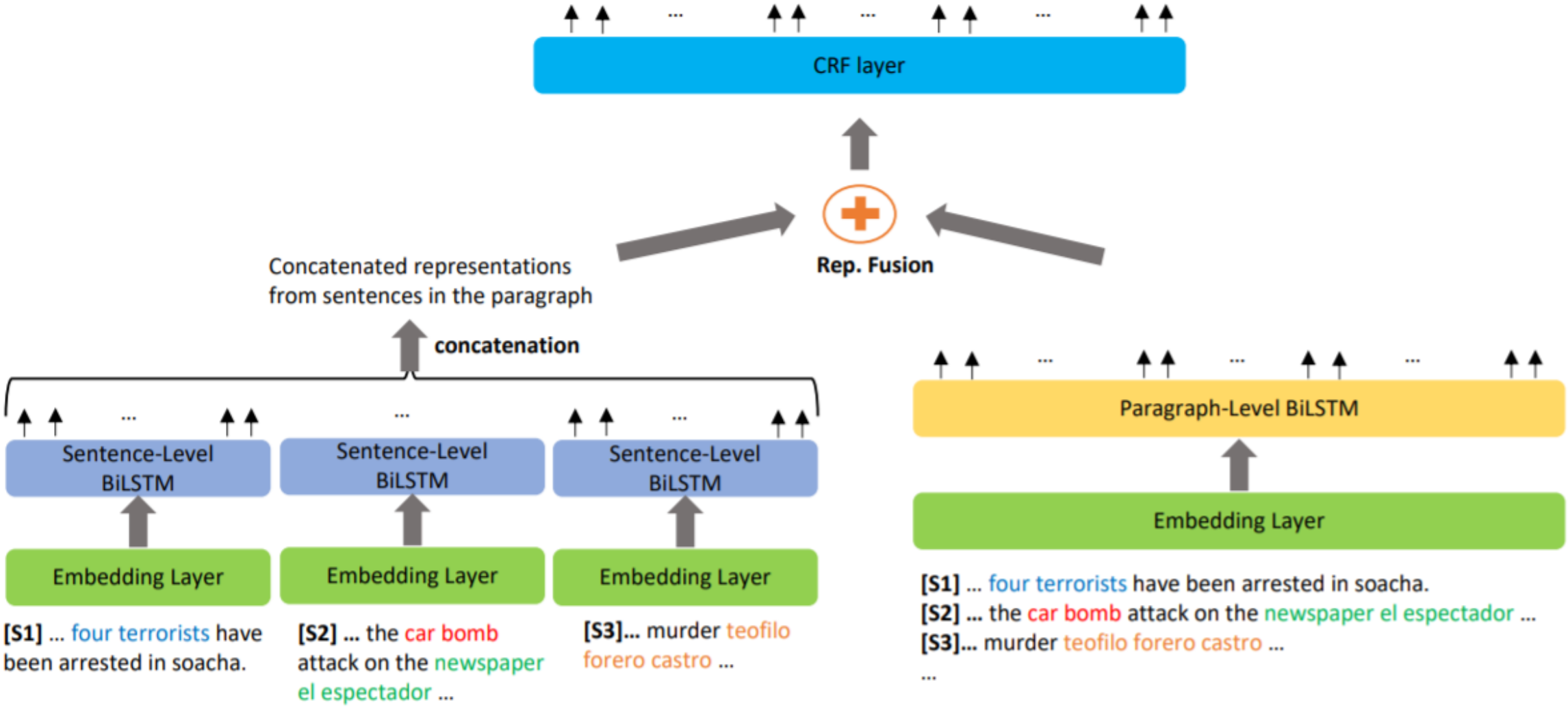
$$\{\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_m\} = \texttt{BiLSTM}(\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_m\})$$

- CRF layer:

$$score(\mathbf{X}, \mathbf{y}) = \sum_{i=0}^{m} \mathbf{A}_{y_i, y_{i+1}} + \sum_{i=1}^{m} \mathbf{P}_{i, y_i}$$



Model: Multi-Granularity Reader



Model: Multi-Granularity Reader

- Sentence-level BiLSTM:

$$\{\tilde{\mathbf{p}}_1^{(1)}, \tilde{\mathbf{p}}_2^{(1)}, ..., \tilde{\mathbf{p}}_{l_1}^{(1)}\}$$
$$= \texttt{BiLSTM}_{sent.}(\{\tilde{\mathbf{x}}_1^{(1)}, \tilde{\mathbf{x}}_2^{(1)}, ..., \tilde{\mathbf{x}}_{l_1}^{(1)}\})$$
$$...$$
$$\{\tilde{\mathbf{p}}_1^{(k)}, \tilde{\mathbf{p}}_2^{(k)}, ..., \tilde{\mathbf{p}}_{l_k}^{(k)}\}$$
$$= \texttt{BiLSTM}_{sent.}(\{\tilde{\mathbf{x}}_1^{(k)}, \tilde{\mathbf{x}}_2^{(k)}, ..., \tilde{\mathbf{x}}_{l_k}^{(k)}\})$$

- Paragraph-level BiLSTM:

$$\{\hat{\mathbf{p}}_1^{(1)}, ..., \hat{\mathbf{p}}_{l_1}^{(1)}, ..., \hat{\mathbf{p}}_1^{(k)}, ..., \hat{\mathbf{p}}_{l_k}^{(k)}\}$$
$$= \texttt{BiLSTM}_{para.}(\{\hat{\mathbf{x}}_1^{(1)}, ..., \hat{\mathbf{x}}_{l_1}^{(1)}, ..., \hat{\mathbf{x}}_{l_k}^{(k)}, ..., \hat{\mathbf{x}}_{l_k}^{(k)}\})$$

- Fusion layer:

$$\mathbf{g}_i^{(j)} = sigmoid(\mathbf{W}_1\tilde{\mathbf{p}}_i^{(j)} + \mathbf{W}_2\hat{\mathbf{p}}_i^{(j)} + b)$$
$$\mathbf{p}_i^{(j)} = \mathbf{g}_i^{(j)} \odot \tilde{\mathbf{p}}_i^{(j)} + (1 - \mathbf{g}_i^{(j)}) \odot \hat{\mathbf{p}}_i^{(j)}$$

$\odot$: element-wise product

Results

	Head Noun Match			Exact Match		
	Prec.	Recall	F-1	Prec.	Recall	F-1
GLACIER (Patwardhan and Riloff, 2009)	47.80	57.20	52.08	-	-	-
TIER (Huang and Riloff, 2011)	50.80	61.40	55.60	-	-	-
Cohesion Extract (Huang and Riloff, 2012)	57.80	59.40	58.59	-	-	-
w/o contextualized embedding						
Single-Sentence Reader	48.69	56.11	52.14	46.16	53.16	49.41
Double-sentence Reader	56.37	47.53	51.57	53.70	43.95	48.34
Paragraph Reader	53.19	53.16	53.17	49.45	49.26	49.35
Chunk Reader	**61.76**	37.04	46.31	**56.91**	34.92	43.28
w/ contextualized embedding						
Contextualized Single-Sentence Reader	47.32	61.26	53.39	44.40	**57.67**	50.17
Contextualized Double-sentence Reader	57.17	53.36	55.20	53.38	49.22	51.22
Contextualized Paragraph Reader	56.78	52.64	54.64	53.36	49.65	51.44
Contextualized Chunk Reader	60.90	41.10	49.07	55.18	37.51	44.66
Multi-Granularity Reader	56.44	**62.77**	**59.44**	52.03	56.81	**54.32**