Multi-Granularity Semantic Aware Graph Model for Reducing Position Bias in Emotion-Cause Pair Extraction

Multi-Granularity Semantic Aware Graph Model for Reducing Position Bias in Emotion-Cause Pair Extraction

Yinan Bao[1,2], Qianwen Ma[1,2], Lingwei Wei[1,2], Wei Zhou[1,*] and Songlin Hu[1,2]
[1]Institute of Information Engineering, Chinese Academy of Sciences
[2]School of Cyber Security, University of Chinese Academy of Sciences
{baoyinan, maqianwen, weilingwei, zhouwei, husonglin}@iie.ac.cn

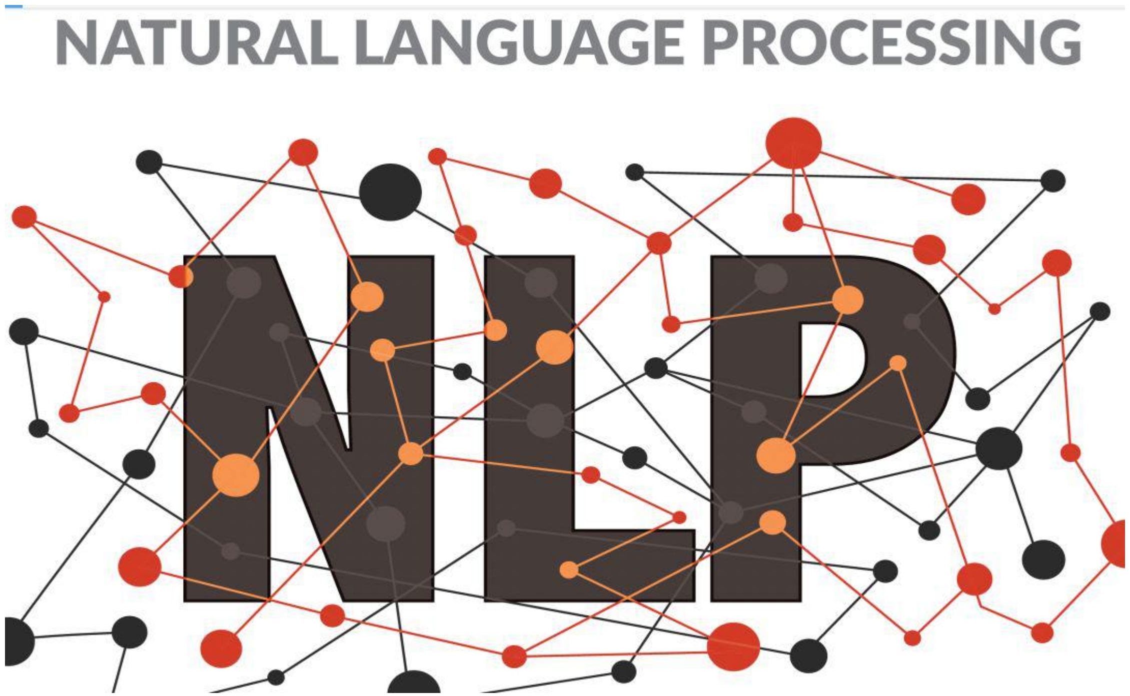
2022.04.17 • ChongQing

2022_ACL

Reported by Yidan Liu

Code:None

Artificial Intelligence



1.Introduction

2.Method

3.Experiments

Introduction

Example. *When the driver was about to start the bus to leave the station* (c_1), *an old lady ran to the front of the bus with a fast speed and sat down on the ground* (c_2). *Passengers standing in the front of the bus can see this scene clearly* (c_3). *Seeing this scene* (c_4), *the passengers in the car immediately became restless* (c_5), *and had a heated debate* (c_6). *Some of the passengers were* ***angry*** (c_7), *and told the driver he shouldn't be meddlesome* (c_8).

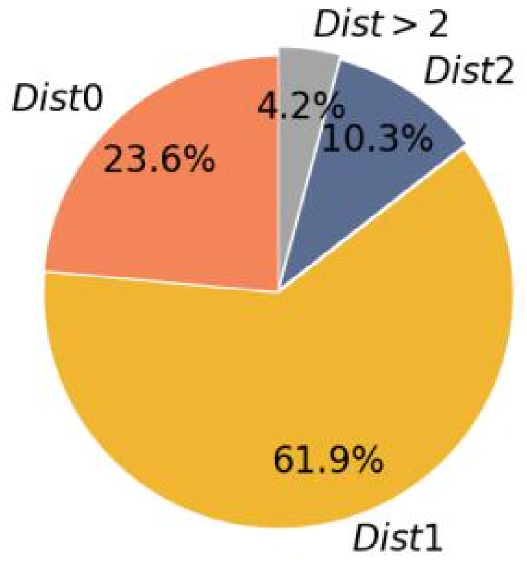


Figure 1: The distribution of the relative distance of an emotion clause and a cause clause that comprise a pair in the ECPE dataset (Xia and Ding, 2019). $Dist0$, $Dist1$, and $Dist2$ mean the relative distances between the two clauses are 0, 1, and 2 respectively. $Dist > 2$ means the relative distances are larger than 2.

Method

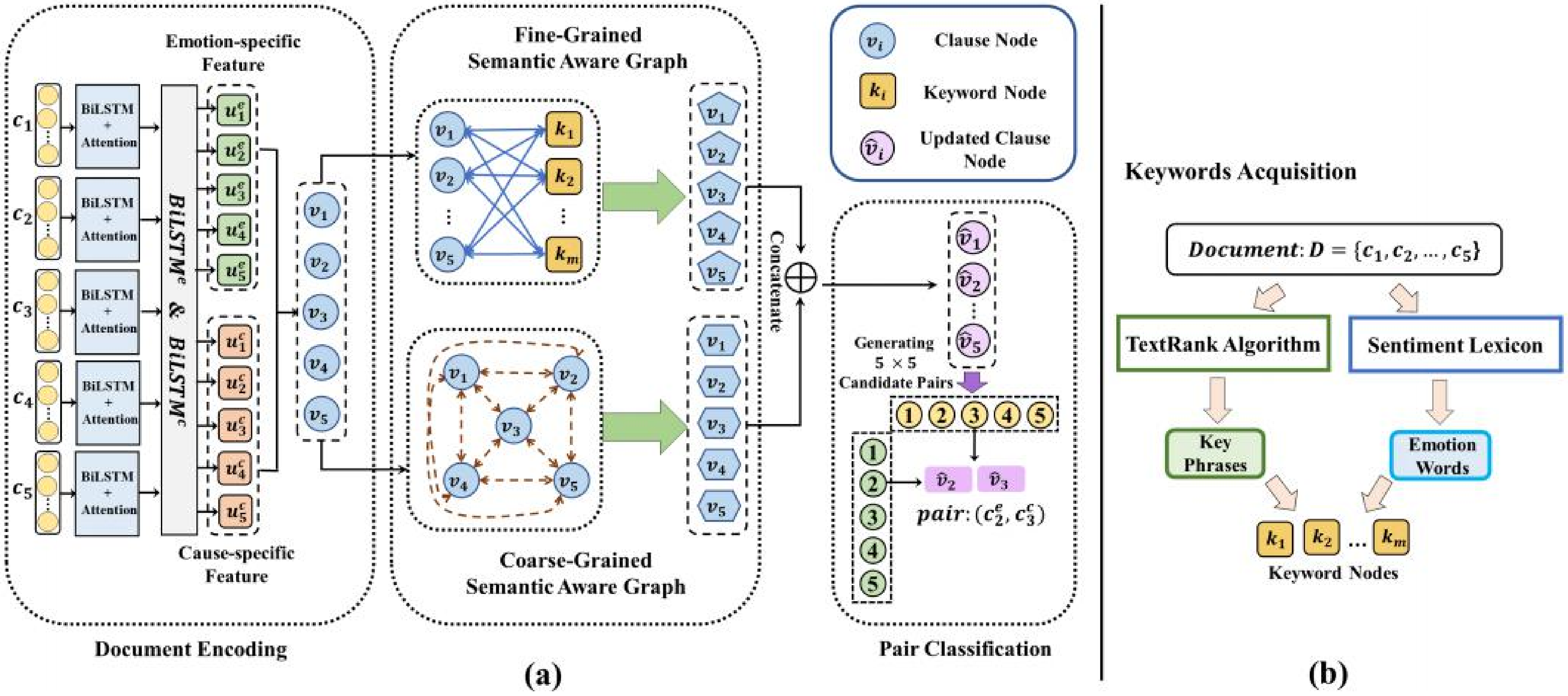
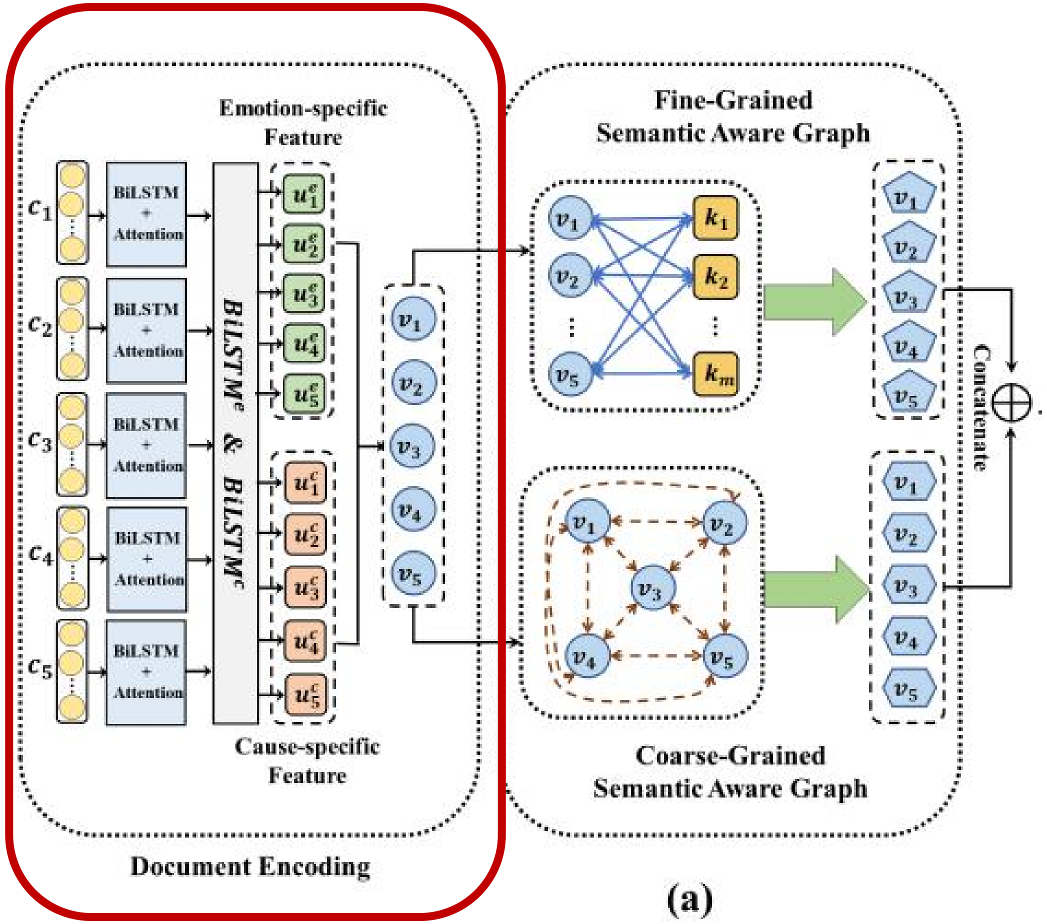


Figure 2: (a) shows an overview of MGSAG. (b) shows the process of keywords acquisition.

Method



(a)

Word-Level Encoder

For each clause $c_i = \{w_1^i, w_2^i, ..., w_{|c_i|}^i\}$

$(h_1^i, h_2^i, ..., h_{|c_i|}^i)$

$\mathbf{h}_i = \sum_{j=1}^{|c_i|} \alpha_j h_j^i \qquad \alpha_j = \mathrm{softmax}(\mathbf{W}_a h_j^i)$

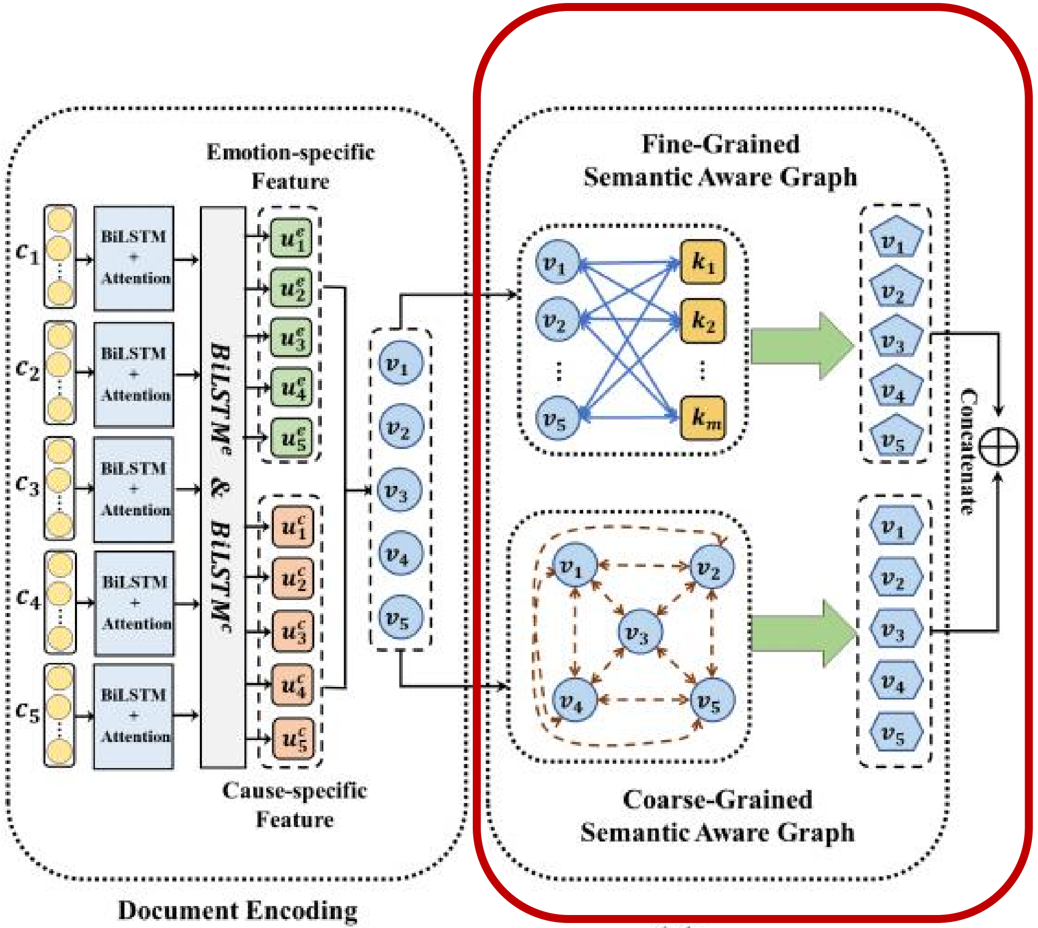
Clause-Level Encoder

$$\mathbf{u}_i^e = \mathrm{BiLSTM}^e(\mathbf{h}_i), \quad \mathbf{u}_i^c = \mathrm{BiLSTM}^c(\mathbf{h}_i), \tag{1}$$

$$\mathbf{g}_i = \sigma(\mathbf{W}_g \mathbf{u}_i^e + \mathbf{b}_g), \quad \mathbf{v}_i = \mathbf{g}_i \mathbf{u}_i^c + (1 - \mathbf{g}_i)\mathbf{u}_i^e, \tag{2}$$

$$\hat{\mathbf{y}}_i^e = \mathrm{softmax}(\mathbf{W}_e \mathbf{u}_i^e + \mathbf{b}_e), \quad \hat{\mathbf{y}}_i^c = \mathrm{softmax}(\mathbf{W}_c \mathbf{u}_i^c + \mathbf{b}_c), \tag{3}$$

Method



(a)

Fine-Grained Semantic Aware Graph

$$\alpha_{ij} = \frac{\exp(w^\top[\mathbf{W}_1\mathbf{v}_i; \mathbf{W}_2\mathbf{k}_j])}{\sum_{t=1}^{|D|}\exp(w^\top[\mathbf{W}_1\mathbf{v}_t; \mathbf{W}_2\mathbf{k}_j])}, \quad (4)$$

$$\mathbf{v}_i^b = \mathbf{tanh}((\mathbf{v}_i + \sum_{j=1}^{m}(\alpha_{ij}(\sum_{t=1}^{|D|}\alpha_{tj}\mathbf{W}_3\mathbf{v}_t))) + \mathbf{b}), \quad (5)$$

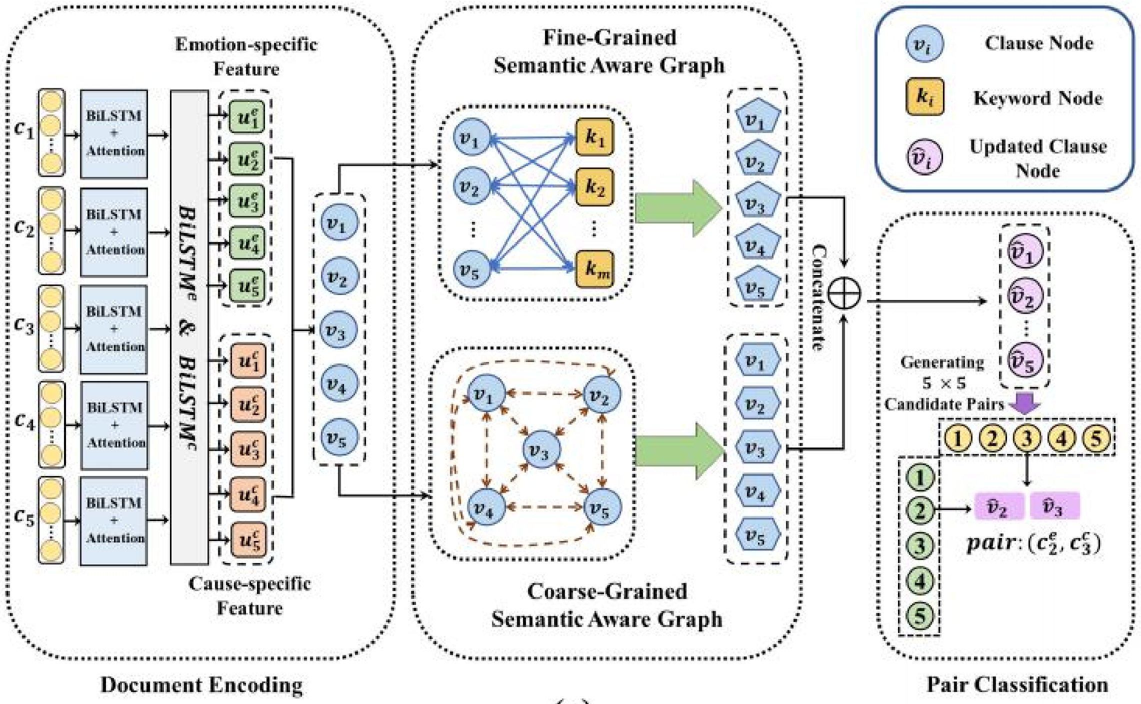
Coarse-Grained Semantic Aware Graph

$$\mathbf{v}_i^{(t)} = \mathbf{ReLU}(\sum_{j\in\mathcal{N}(i)}\alpha_{ij}^{(t)}\mathbf{W}_1^{(t)}\mathbf{v}_j^{(t-1)} + \mathbf{b}^{(t)}), \quad (6)$$

$$e_{ij}^{(t)} = w^{(t)\top}\mathbf{tanh}([\mathbf{W}_2^{(t)}\mathbf{v}_i^{(t-1)}; \mathbf{W}_3^{(t)}\mathbf{v}_j^{(t-1)}]),$$

$$\alpha_{ij}^{(t)} = \frac{\exp(\mathbf{LeakyReLU}(e_{ij}^{(t)}))}{\sum_{k\in\mathcal{N}(i)}\exp(\mathbf{LeakyReLU}(e_{ik}^{(t)}))}, \quad (7)$$

Method



(a)

$$\hat{\mathbf{v}}_i = [\mathbf{v}_i^b; \mathbf{v}_i^c]$$

Emotion Cause Pair Extraction

$$\hat{\mathbf{p}}_{ij} = \mathrm{softmax}(\mathbf{W}_p^\top \mathbf{v}_{ij}^p + \mathbf{b}_p)\,, \qquad (8)$$

$$\mathcal{L} = \mathcal{L}_{pair} + \mathcal{L}_{emo} + \mathcal{L}_{cau}\,. \qquad (9)$$

Emotion Extraction and Cause Extraction

$$\hat{\mathbf{E}}_i = \begin{cases} 1, & if \sum_{j=1}^{|D|}(\hat{\mathbf{EC}}_{ij}) > 0 \\ 0, & otherwise \end{cases} \quad . \qquad (10)$$

Experiments

Category	Model	Emotion Ext.			Cause Ext.			EC Pair Ext.		
		P	R	F_1	P	R	F_1	P	R	F_1
Position-insensitive Baselines	Indep	0.8375	0.8071	0.8210	0.6902	0.5673	0.6205	0.6832	0.5082	0.5818
	Inter-CE	0.8494	0.8122	0.8300	0.6809	0.5634	0.6151	0.6902	0.5135	0.5901
	Inter-EC	0.8364	0.8107	0.8230	0.7041	0.6083	0.6507	0.6721	0.5705	0.6128
	IE-CNN	0.8614	0.7811	0.8188	0.7348	0.5841	0.6496	0.7149	0.6279	0.6686
Position-sensitive Baselines	PairGCN	0.8587	0.7208	0.7829	0.7283	0.5953	0.6541	0.6999	0.5779	0.6321
	ECPE-2D	0.8512	0.8220	0.8358	0.7272	0.6298	0.6738	0.6960	0.6118	0.6496
	SLSN-U	0.8406	0.7980	0.8181	0.6992	0.6588	0.6778	0.6836	0.6291	0.6545
	RankCP	0.8703	0.8406	**0.8548**	0.6927	**0.6743**	0.6824	0.6698	**0.6546**	0.6610
	ECPE-MLL	0.8582	**0.8429**	0.8500	0.7248	0.6702	0.6950	0.7090	0.6441	0.6740
Our Model	**MGSAG**	**0.8721**	0.7911	0.8287	**0.7510**	0.6713	**0.7080**	**0.7243**	0.6507	**0.6846**

Table 1: Comparison of varying approaches on the original test set ($Test_{all}$).

Experiments

Model	$Test_{Bias}$	$Test_{NoBias}$
Inter-EC	0.6783	0.3318
IE-CNN	0.7666	0.3484
PairGCN	0.7246	0.3355
ECPE-2D	0.7590	0.3830
SLSN-U	0.7456	0.3978
RankCP	0.7467	0.3857
ECPE-MLL	0.7673	0.3988
MGSAG	**0.7730**	**0.4301**

Table 2: F_1 results of varying approaches on $Test_{Bias}$ and $Test_{NoBias}$, focusing on EC Pair Ext.

Model	$Test_{Bias}$	$Test_{NoBias}$	$Test_{all}$
w/o FGSAG	0.7594	0.3894	0.6519
w/o CGSAG	0.7654	0.4027	0.6529
w/o FGSAG+CGSAG	0.7264	0.3269	0.6242
MGSAG	**0.7730**	**0.4301**	**0.6846**

Table 3: F_1 results of ablation study on $Test_{Bias}$, $Test_{NoBias}$, and $Test_{all}$, focusing on EC Pair Ext.

Experiments

Loss Function	P	R	F_1
$\mathcal{L}_{pair}$	0.6940	**0.6533**	0.6720
$\mathcal{L}_{pair} + \mathcal{L}_{emo} + \mathcal{L}_{cau}$	**0.7243**	0.6507	**0.6846**

Table 4: Comparison of different supervised signals for our method.

Model	$Test_{Bias}$	$Test_{NoBias}$	$Test_{all}$
w/ RW	0.7596	0.4078	0.6674
w/o EW	0.7669	0.3920	0.6686
w/o TW	0.7658	0.4271	0.6771
MGSAG	**0.7730**	**0.4301**	**0.6846**

Table 5: Comparative F_1 results on $Test_{Bias}$, $Test_{NoBias}$, and $Test_{all}$ of our variant models, focusing on EC Pair Ext. "w/ RW" means using random embeddings for keyword feature initialization. "w/o EW" and "w/o TW" means removing emotion words and key phrases obtained by TextRank, respectively.

Experiments

(c_1) When Mr. Chen drove from south to north on Yangtze River Avenue,
(c_2) it was dark and there was no street lamp.
(c_3) His slight was not clear.
(c_4) He ran into a stone and overturned.
(c_5) After that,
(c_6) he was trapped in the car,
(c_7) broke the window glass,
(c_8) and climbed out of the car.
(c_9) He said that,
(c_{10}) because of the sudden incident,
(c_{11}) he was shocked and injured.

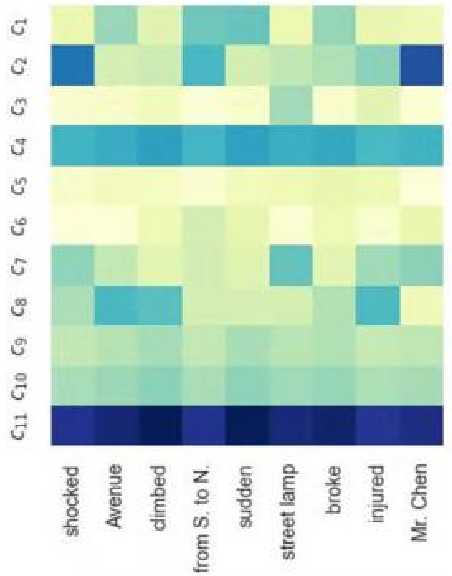


Figure 4: An example that MGSAG extracts the emotion cause pair (c_{11}, c_4) correctly, while ECPE-MLL fails. Words shaded in yellow are keywords. The heatmap presents attention scores in the clause-keyword bipartite graph. Rows of c_{11} and c_4 are the top-two darkest rows, means that keywords pay more attention to them and facilitate MGSAG to extract pair (c_{11}, c_4) correctly.

Thank you!