Event, Emotion and Causality in Commonsense View

虞扬

20221019

Commonsense Knowledge

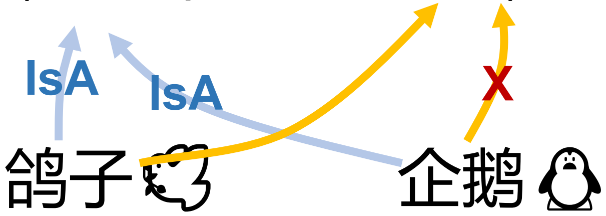
Definition

“Millions of basic facts and understanding possessed by most people”

Encyclopedic Knowledge: (《猛虎行》, 作者, 李白)

Factual Knowledge: (中华人民共和国, 首都, 北京)

Commonsense Knowledge: (鸟, CapableOf, 飞翔)



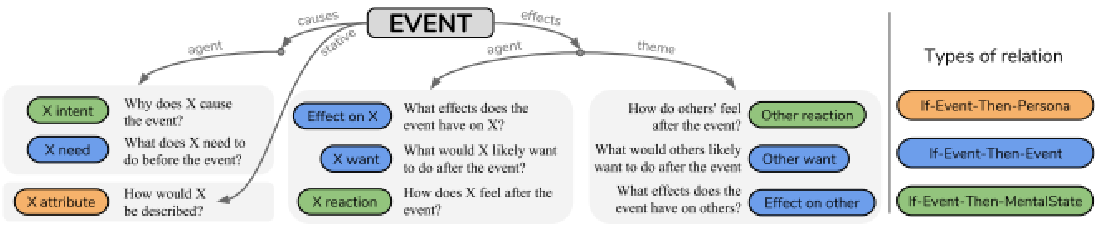
A. The scale of commonsense knowledge could be enormous

B. A kind of preference rather than fixed fact

C. Implicit in Text, and more conceptual

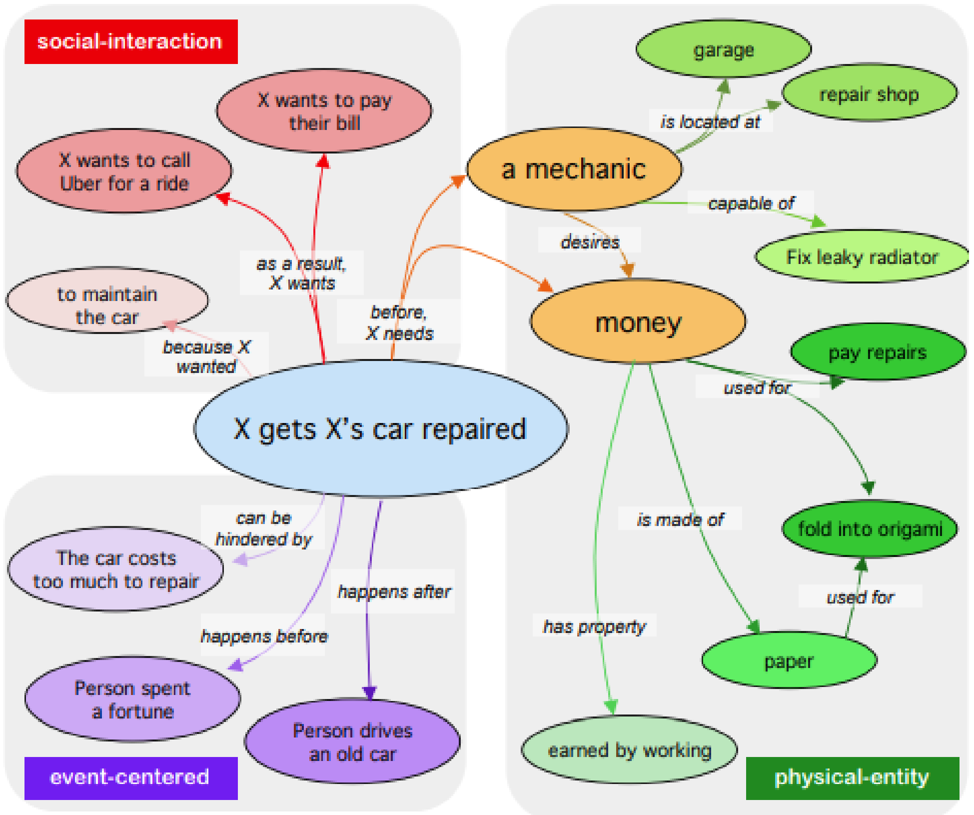
Eventuality Dataset

[2020] Event-Centered



	Count	#words
# triples: If-Event-Then-*	877,108	-
- Mental-State	212,598	-
- Event	521,334	-
- Persona	143,176	-
# nodes: If-Event-Then-*	309,515	2.7
- Mental-State	51,928	2.1
- Event	245,905	3.3
- Persona	11,495	1.0
Base events	24,313	4.6
# nodes appearing > 1	47,356	–

[2020] Mental-Event-Entity

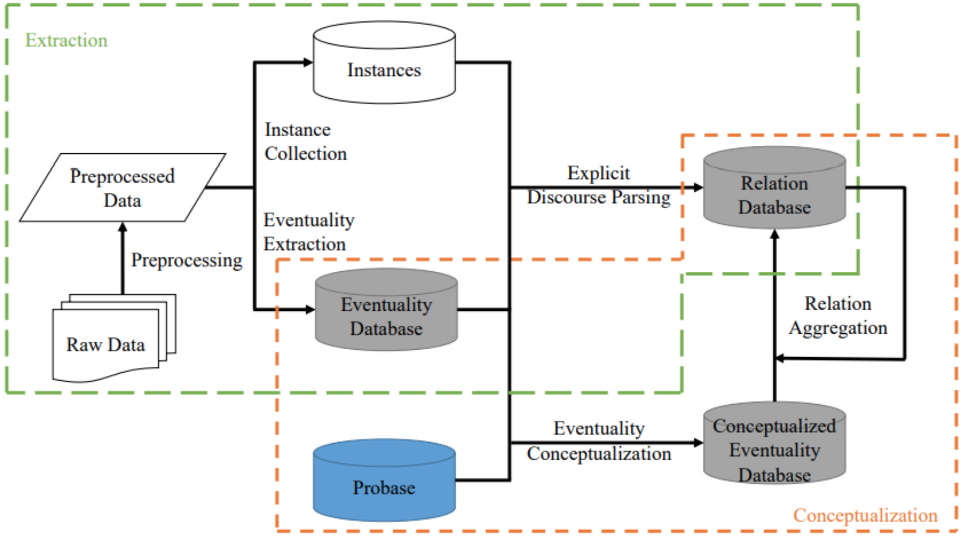


	Head	Relation
PHYSICAL-ENTITY	bread	ObjectUse
		AtLocation*
		MadeUpOf
		HasProperty*
	baker	CapableOf*
		Desires*
		Not Desires*
EVENT-CENTERED	X runs out of steam	IsAfter
		HasSubEvent
		IsBefore
		HinderedBy
		Causes
		xReason
	X watches ___ anyway	isFilledBy
SOCIAL-INTERACTION	X runs out of steam	xNeed
		xAttr
		xEffect
		xReact
		xWant
	X votes for Y	xIntent
		oEffect
		oReact
		oWant

ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning. AAAI. 2019
(Comet-) Atomic 2020: On Symbolic and Neural Commonsense Knowledge Graphs. AAAI. 2021

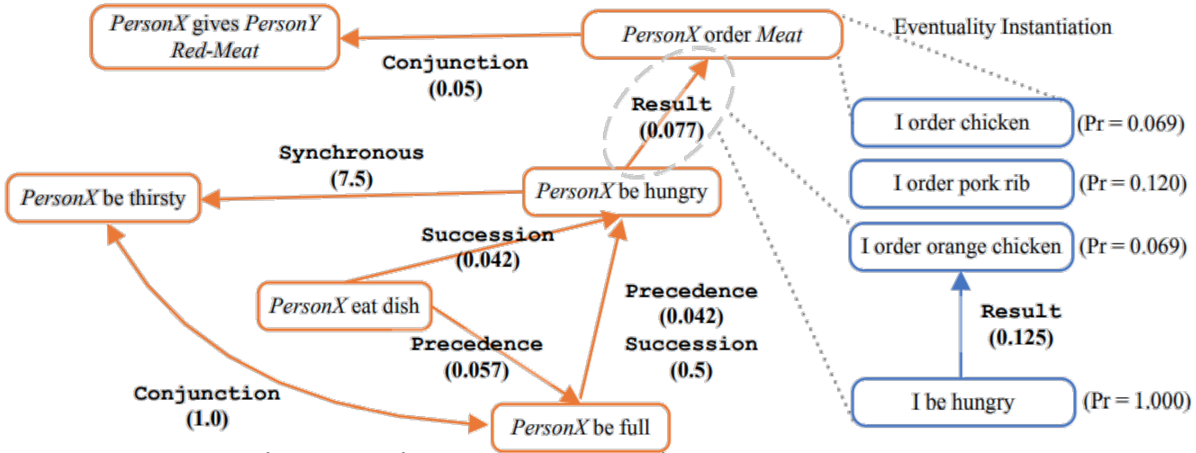
Event Mining

ASER

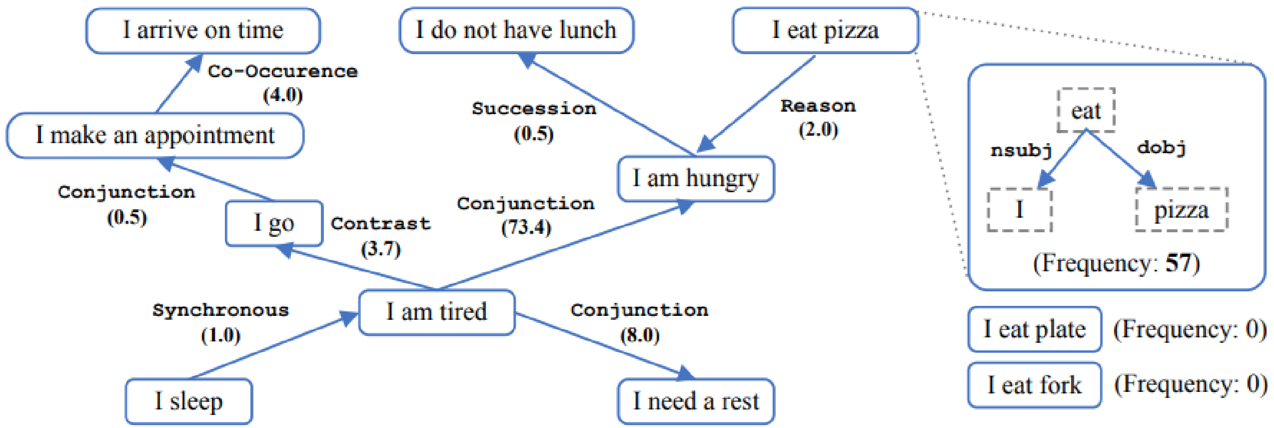


	# Eventuality	# Relation	# Relation Types
FrameNet	27,691	1,709	7
ACE	3,290	0	0
PropBank	112,917	0	0
NomBank	114,576	0	0
TimeBank	7,571	8,242	1
OMCS (Only include edges about eventualities)	74,989	116,097	4
Event2Mind	24,716	57,097	3
ProPora	2,406	16,269	1
ATOMIC	309,515	877,108	9
ATOMIC-2020	638,128	1,331,113	23
GLUECOSE	286,753	304,099	10
Knowlywood	964,758	2,644,415	4
ASER (core)	52,940,258	52,296,498	14
ASER (full)	438,648,952	648,514,465	14
ASER (concept)	15,640,017	224,213,142	14

Concept



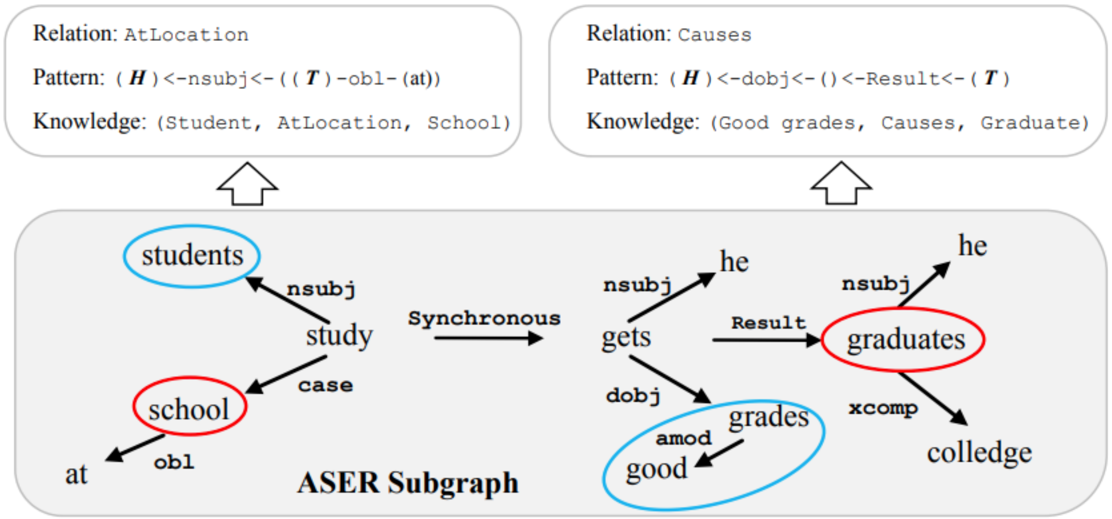
Instance



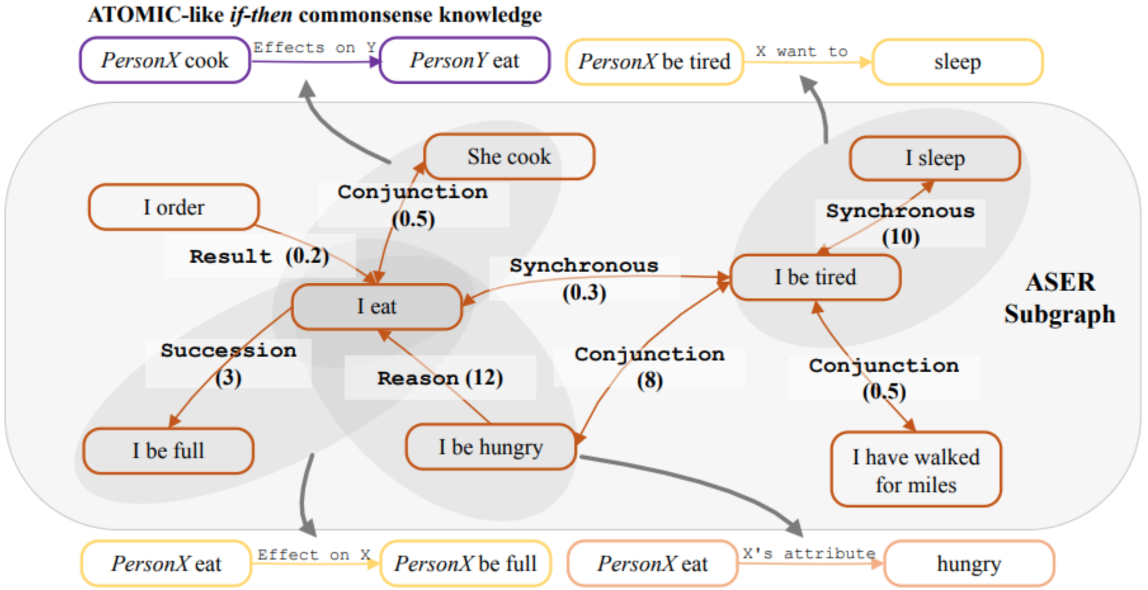
ASER: A Large-scale Eventuality Knowledge Graph. WWW. 2020
ASER: Towards Large-scale Commonsense Knowledge Acquisition via Higher-order Selectional Preference over Eventualities. Artificial Intelligence. 2022

Event Mining

TransOMCS
To ConceptNet-like Commonsense



DISCOS
To ATOMIC-like Commonsense



TransOMCS: From Linguistic Graphs to Commonsense Knowledge. IJCAI. 2020
DISCOS: Bridging the Gap between Discourse Knowledge and Commonsense Knowledge. WWW. 2021

Event-centric Dataset

In Downstream Task (Conversation)

Problems:
when a friend in **distress** tells us that he recently **adopted a cat** .

1. finds out he has allergies
2. becomes less lonely
3. finds a cat at the animal shelter

ATOMIC

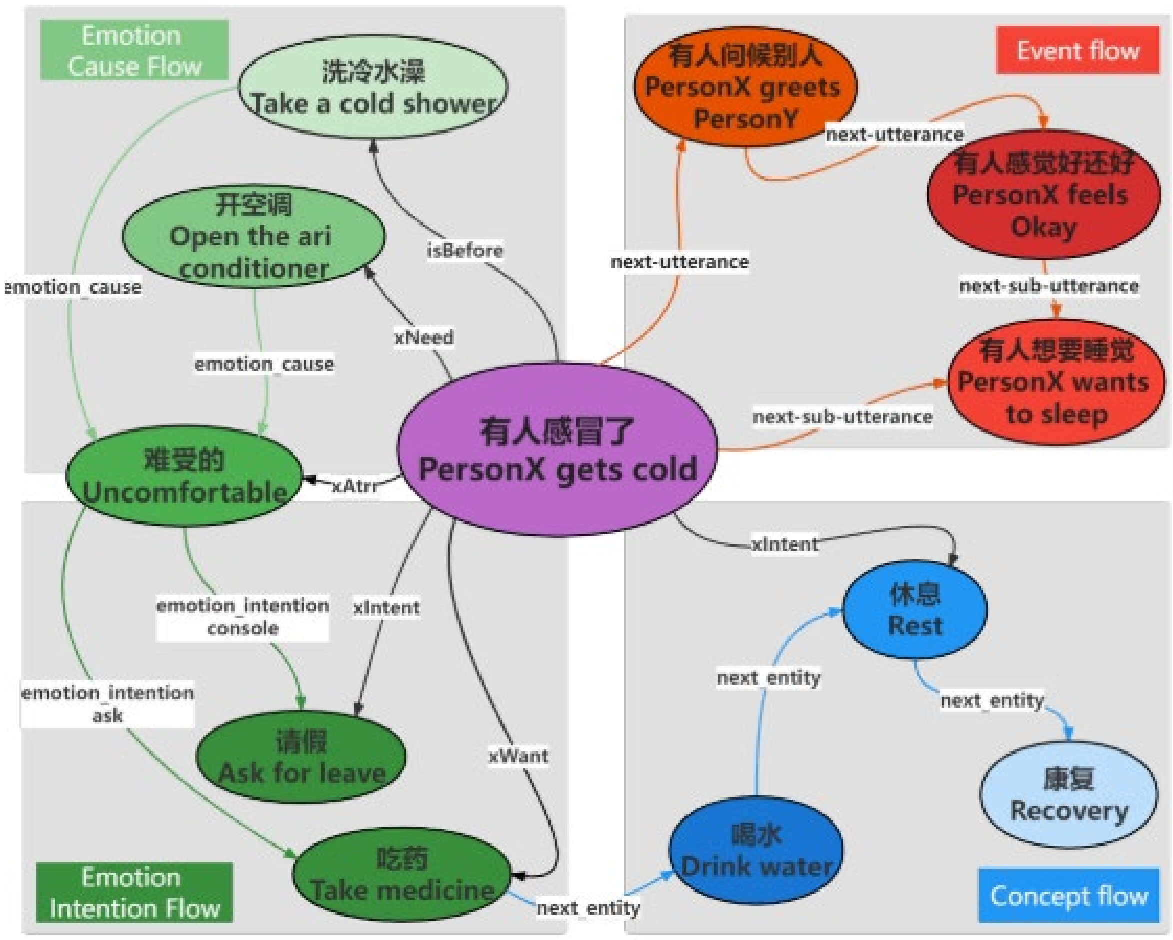
Selectional Preference

A. Event Centric
   Emotion can help (Persona)

B. Events in ATOMIC are isolated
   follow Dialog Flow (Context)

How ?

Flow-Centered



Difficulties

Relation	ASER (full)	ASER (core)	ASER (concept)
Precedence	14,058,213	1,790,016	4,798,015
Succession	4,939,291	663,183	1,963,820
Synchronous	19,464,898	3,123,042	8,013,943
Reason	9,775,829	2,205,076	6,439,128
Result	16,153,925	2,012,311	6,718,666
Condition	18,052,484	3,160,271	8,063,967
Contrast	59,333,901	8,655,661	24,978,311
Concession	5,684,395	477,155	1,499,276
Conjunction	82,121,343	13,978,907	45,597,200
Instantiation	1,278,381	18,496	93,266
Restatement	1,304,095	65,753	242,301
Alternative	3,539,892	583,174	123,883
ChosenAlternative	647,228	35,406	1,843,140
Exception	106,000	20,155	93,412
Co-Occurrence	412,054,590	49,232,161	113,744,814
Overall	648,514,465	86,020,767	224,213,142

Mapping statistics of ATOMIC nodes and edges in ASER

Relation	Nodes	Edges	Avg. Shortest Path Length
oEffect	31.1%	25.36%	2.41
oReact	87.3%	51.53%	2.22
oWant	61.6%	36.95%	2.47
xAttr	95.8%	53.67%	2.38
xEffect	33.1%	21.81%	2.51
xIntent	33.8%	21.06%	2.56
xNeed	52.9%	24.91%	2.67
xReact	88.7%	52.66%	2.25
xWant	58.8%	30.60%	2.59
Average	62.9%	35.91%	2.44

1. **Bad Coverage on Causality-related Relations**
2. **Emotion-related clues are easier to mine**

Causality

1. Word

   Suicide is one of the leading causes of death.

   Cause

   Effect

2. Phrase

   Financial stress is one of the main causes of divorce.

3. Clause

   We play with a steady beat so that dancers can follow it.

4. Event

   A car travelling from Guizhou to Guangdong collided head-on with a bush results the ten people, six men and four women, including the driver, died at the scene.

Causality Extraction

1.Simple Causality

[Financial stress] is one of the main causes of **[divorce]**.

{[Financial stress], Cause-Effect, [divorce]}

2.Embedded Causality

{[Lesions], Cause-Effect, [a contralateral, severe leg-predominant hemiplegia]}

[Lesions] involving the medial part of the premotor cortex, the supplementary motor area and the rear portion of the medial part of the precentral gyrus caused **[a contralateral, severe leg-predominant hemiplegia]**, **[distally predominant and a less severe proximal weakness]** of the arm.

3.Cross Causality

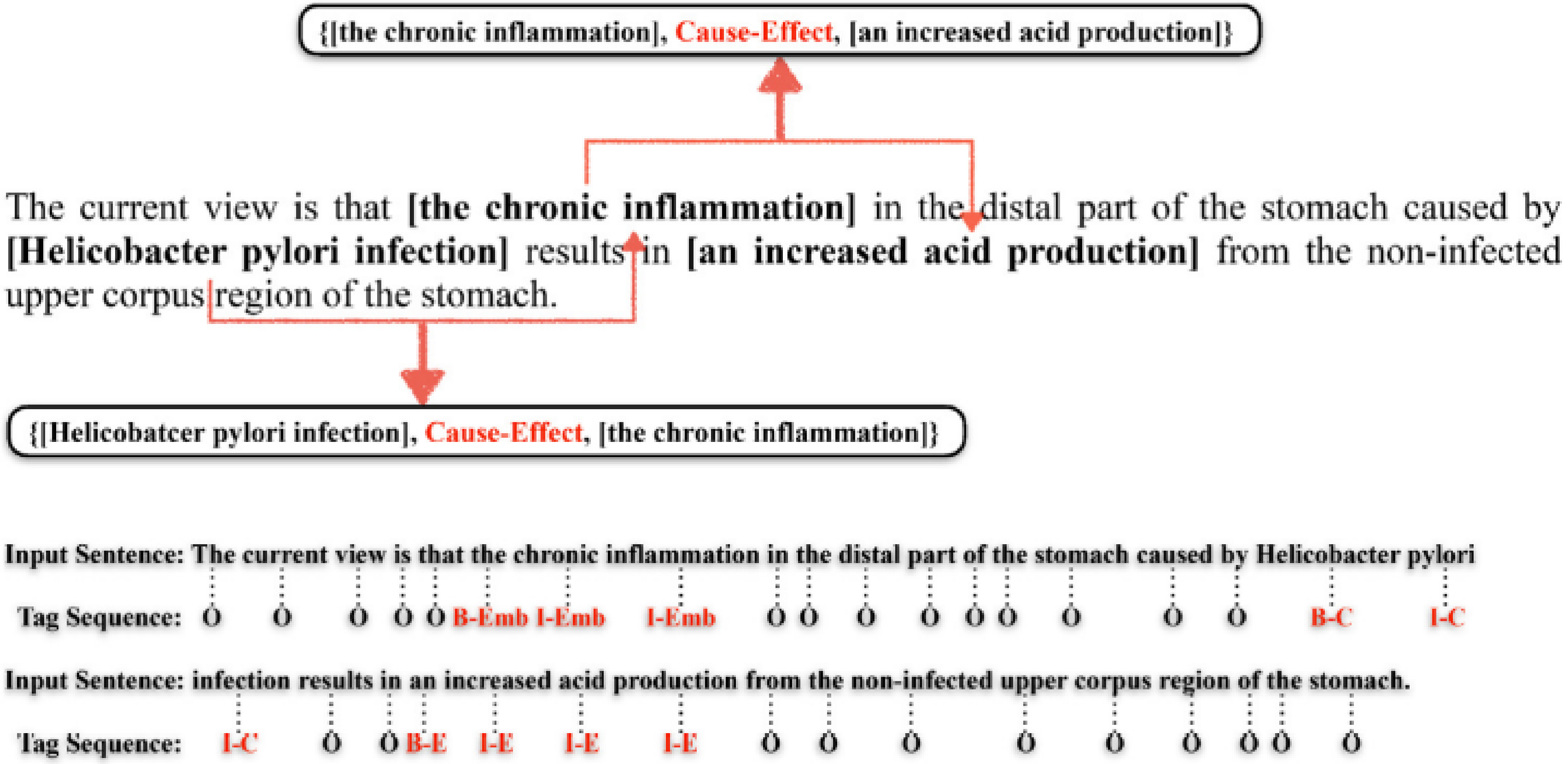
{[Lesions], Cause-Effect, [distally predominant and a less severe proximal weakness]}

Tagging

Traditional BIO Tagging
"B" for Begin
" I " for Inside
"O" for Other

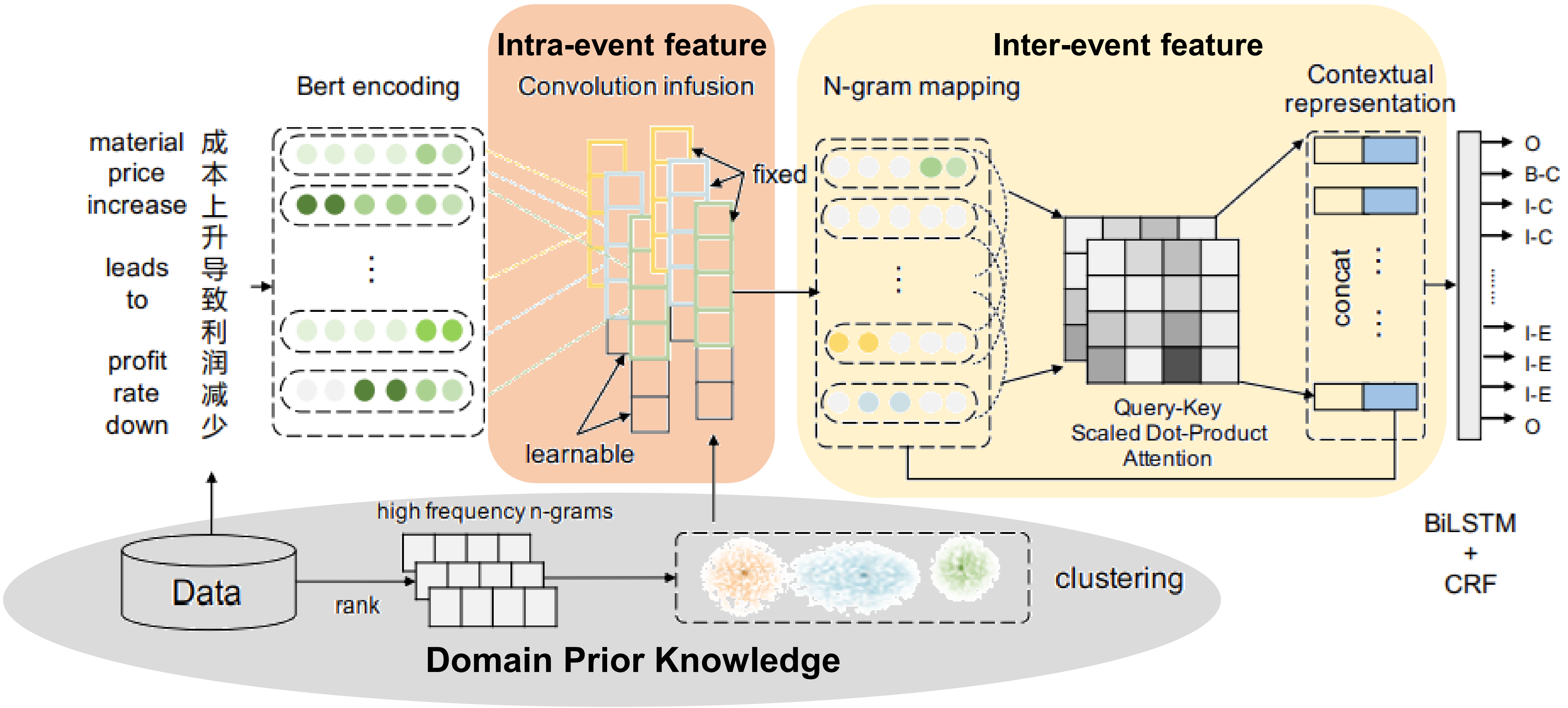
Causality Extraction
C for Cause
E for Effect

Embedded Causality
Emb for this



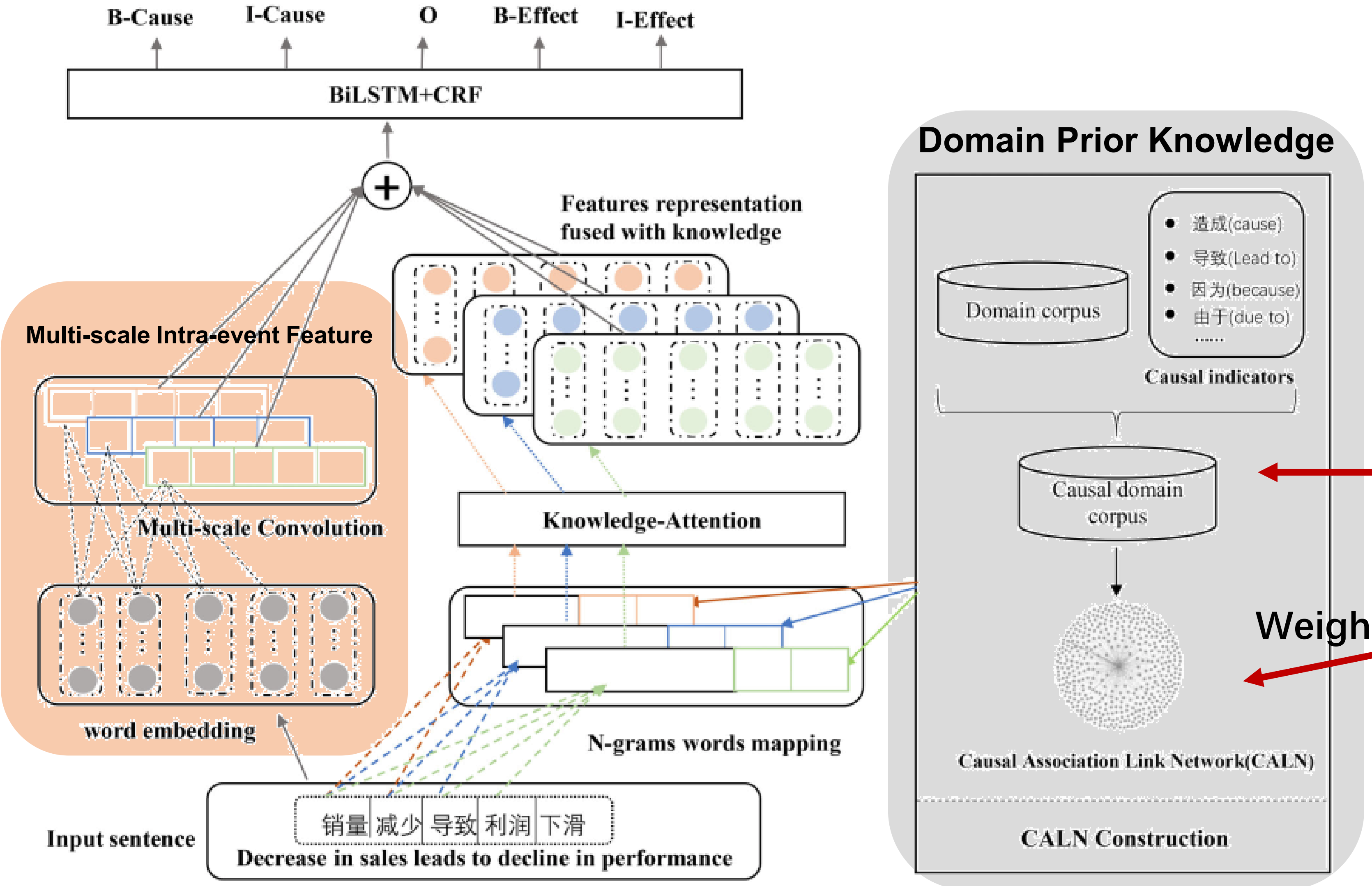
Domain-augmented

CISAN



Back to Prior Knowledge: Joint Event Causality Extraction via Convolutional Semantic Infusion. PAKDD. 2021

Domain-augmented



Noun-Verb Graph
exclude Named Entity

Weight $A^{w_i,w_j} = \frac{Co(w_i,w_j)}{\sqrt{DF(w_i)\cdot DF(w_j)}}$

Multi-scale event causality extraction via simultaneous knowledge-attention and convolutional. Expert Systems. 2022

Model	SemEval2010 (1)	CEC (2)	Financial (4)
BiLSTM+CRF	73.20	68.74	74.75
CNN+BiLSTM+CRF	74.20	71.68	74.31
CNN+Att+BiLSTM+CRF (CSNN)	73.71	70.61	74.59
CISAN(w/o Bert)	74.20	72.49	75.99
Bert-CSNN	75.69	74.61	76.23
CISAN (unigram)	**77.65**	75.26	76.07
CISAN (bigram)	77.26	**75.93**	76.70
CISAN (trigram)	77.14	74.45	76.29
CISAN (quagram)	77.35	75.27	**77.09**
KA-CNN (w/o CNN)	-	73.68	77.90
KA-CNN (w/o CALN)	-	72.03	75.49
KA-CNN	-	**74.39**	**80.83**

Causality Score:

Hits Score:

H(X, Y) / H(Y, X)

Similarity Score:

$\sum_{\mathrm{i=k}}^{i=0} Similarity\_Score(\text{"X may cause Y"}, S_i)$ / K

NLI Score(CaKNowLI):

(X, Y) => *“**X** causes **Y**”* => $\sum_{\mathrm{i=k}}^{i=0} NLI\_Score(prompt(X,Y))$ / K

*“**X** is the reason for **Y**”* **BERT**

… (NLI)

Causal Knowledge Extraction through Large-Scale Text Mining. AAAI. 2020
Unsupervised Causal Knowledge Extraction from Text using Natural Language Inference. AAAI(Student Abstract). 2021

Emotion-Cause Pair Extraction

c_1 : Yesterday morning,

c_2^C: a police man visited the old man with the lost money,

c_3^C: and told him that the thief was caught.

c_4^E: The old man was very happy,

c_5 : and deposited the money in the bank.

Document

$d = [c_1, c_2, \dots, c_{|d|}]$

{

(c_4^E: The old man was very happy,
c_2^C: a police man visited the old man with the lost money)

(c_4^E: The old man was very happy,
c_3^C: and told him that the thief was caught)

}

Clause Pairs

$P = \{\dots, (c^e, c^c), \dots\}$

Emotion-Cause Pair Extraction

c_1 : At first I thought this restaurant was good,

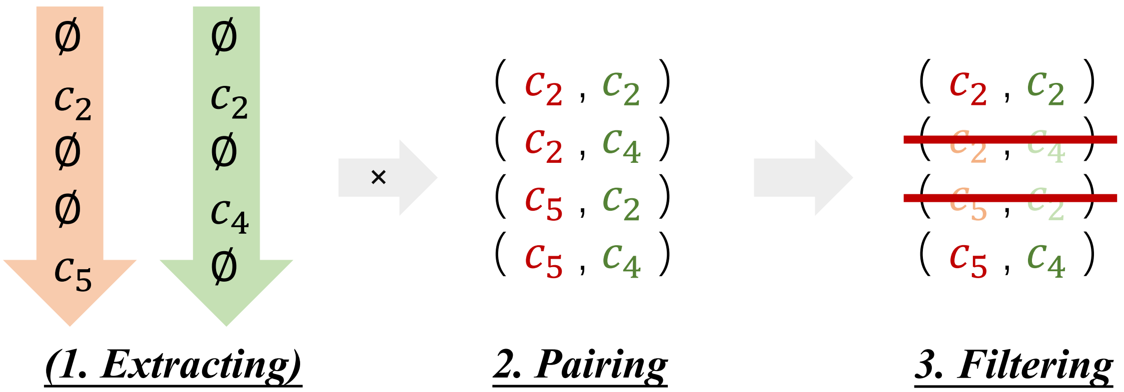
c_2^C: but I was tired of going to the same restaurant always,

c_3^C: I want to change to another restaurant.

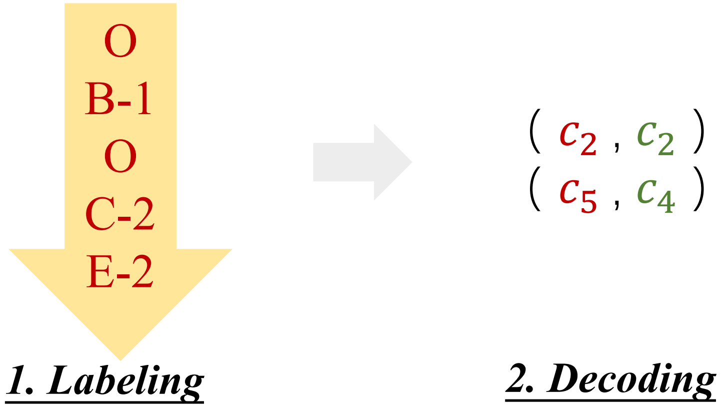
c_4^E: but my friend says that it is affordable,

c_5 : which makes me disgusted.

A. Pipeline



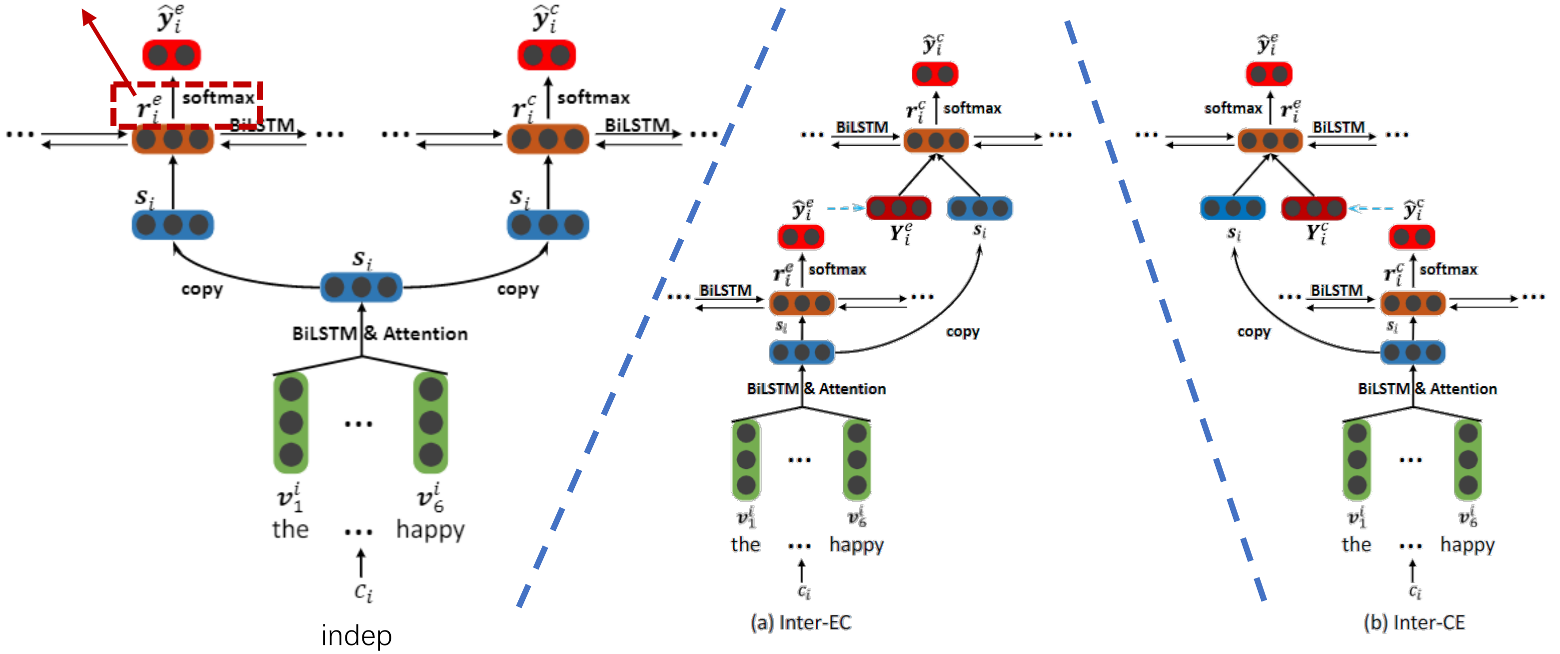
B. Sequence Labeling



ECPE-2Step

A. Emotion & cause Extraction

$$\hat{y}_i^e = softmax(W^e r_i^e + b^e)$$



$$c_i = [v_{i,1}, v_{i,2}, v_{i,3}, v_{i,4}, v_{i,5}, v_{i,6}]$$

…was caught. The old man was very happy, and deposited …

B. Pairing & Filtering

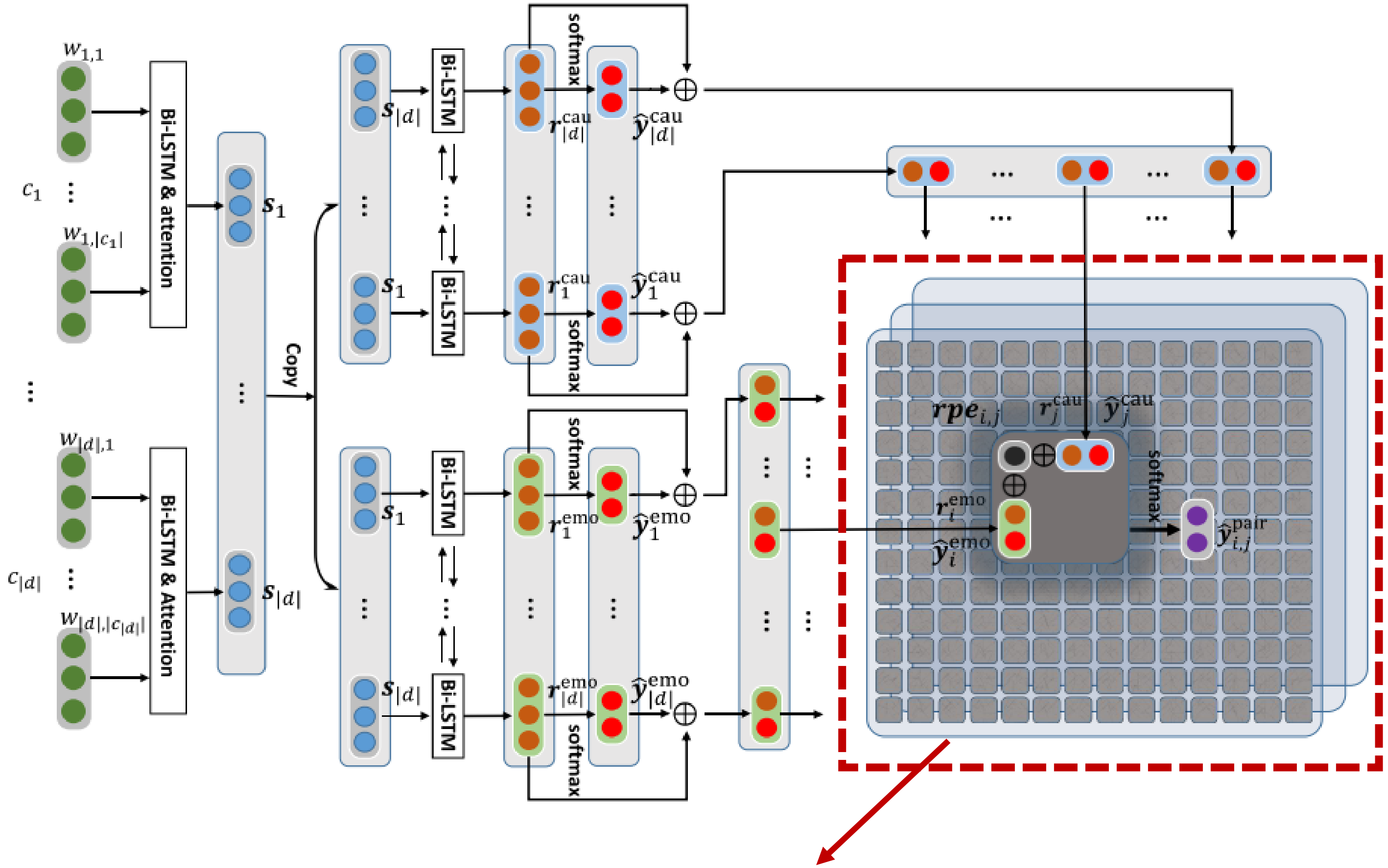
$$P_{all} = \{\dots, (c_i^e, c_j^c), \dots\}$$

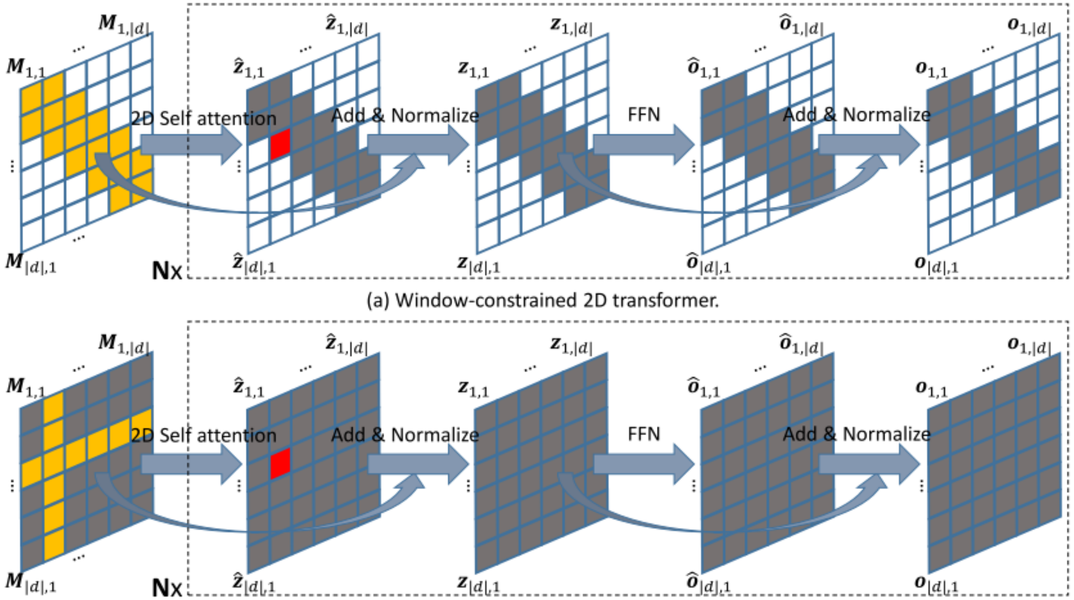
$$x_{(c_i^e, c_j^c)} = [s_i^e; s_j^c; v^d]$$

$$\hat{y}_{(c_i^e, c_j^c)} = \delta(\theta^T x_{(c_i^e, c_j^c)})$$

Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts. ACL. 2019

ECPE-2D





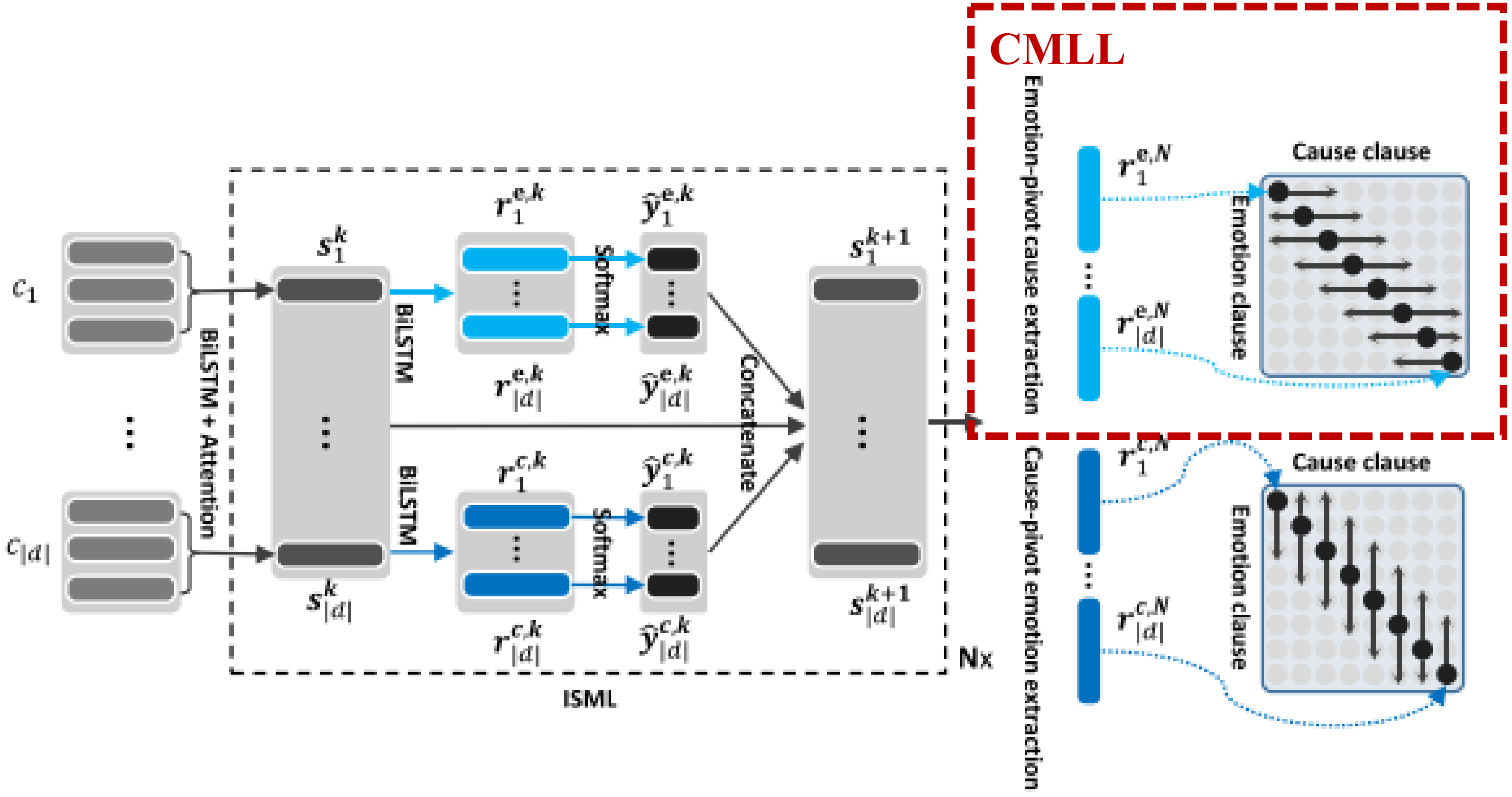
(a) Window-constrained 2D transformer.

(b) Cross-road 2D transformer.

$$M_{i,j} = r_i^{emo} \oplus \hat{y}_i^{emo} \oplus r_j^{cau} \oplus \hat{y}_j^{cau} \oplus rpe_{i,j}$$

ECPE-2D: Emotion-Cause Pair Extraction based on Joint Two-Dimensional Representation, Interaction and Prediction. ACL. 2020

ECPE-MLL



$$c_{i-|w|}, \cdots, c_{i-1}, c_i, c_{i+1}, \cdots, c_{i+|w|}$$

$$\mathbf{y}_i^{\mathrm{cml}} = [y_i^{\mathrm{cml}_{-|w|}}, \cdots, y_i^{\mathrm{cml}_0}, \cdots, y_i^{\mathrm{cml}_{|w|}}]$$

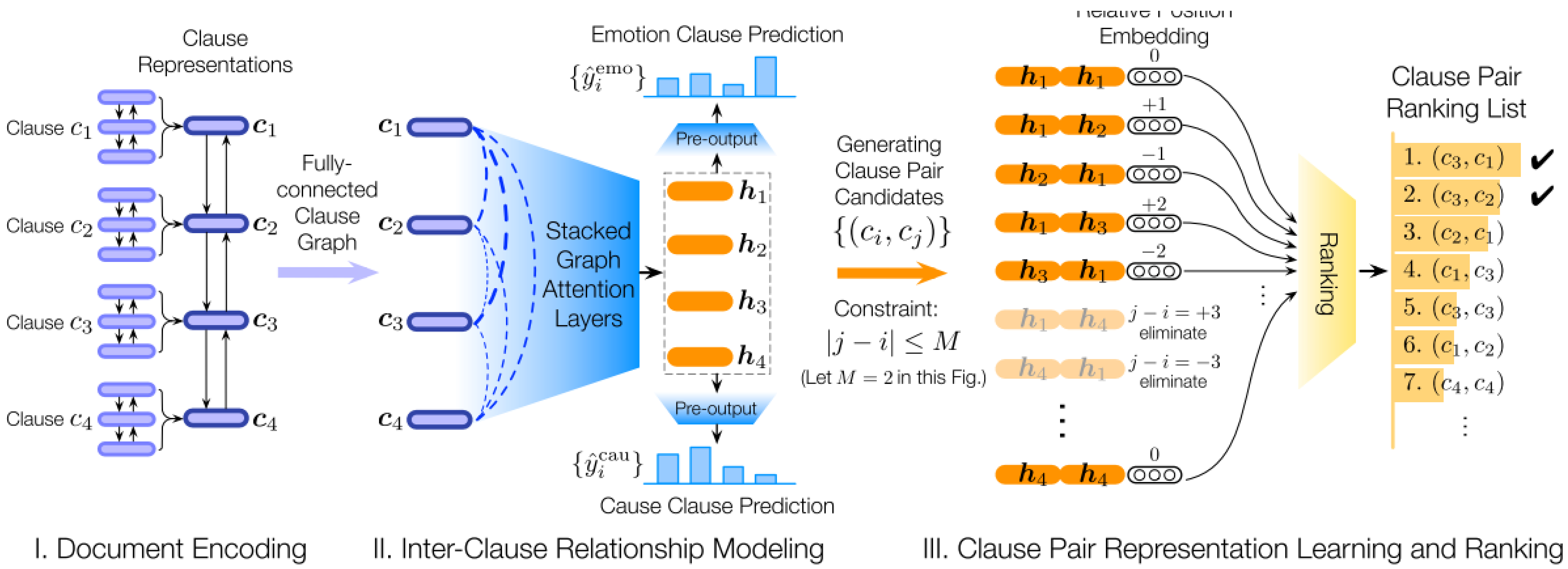
$$p(y_i^{\mathrm{cml}_j} = 1|c_i) = \hat{y}_i^{\mathrm{cml}_j} = \frac{1}{1 + e^{\mathbf{W}^{\mathrm{cml}_j}\mathbf{r}_i^{e,N} + \mathbf{b}^{\mathrm{cml}_j}}}, \tag{7}$$

$$p(y_i^{\mathrm{cml}_j} = 0|c_i) = 1 - \hat{y}_i^{\mathrm{cml}_j}, \tag{8}$$

$$L^{\mathrm{CMLL}} = -\sum_{i=1}^{|d|}\sum_{j=-|w|}^{|w|}(y_i^{\mathrm{cml}_j} \cdot \log\hat{y}_i^{\mathrm{cml}_j} + (1 - y_i^{\mathrm{cml}_j}) \cdot \log(1 - \hat{y}_i^{\mathrm{cml}_j})). \tag{9}$$

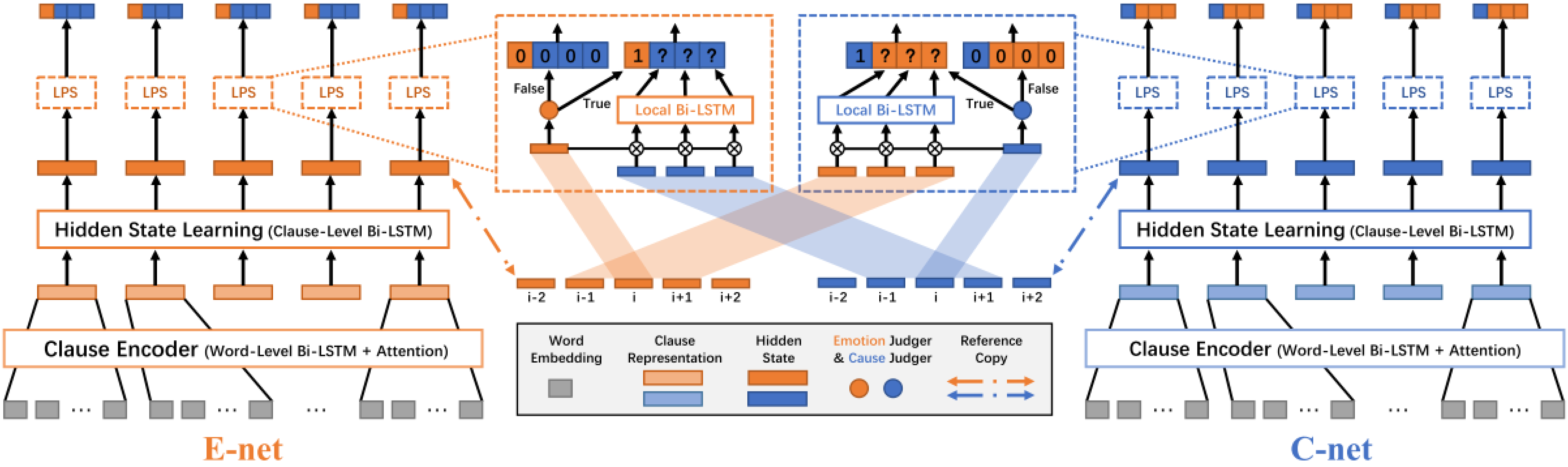
End-to-End Emotion-Cause Pair Extraction based on Sliding Window Multi-Label Learning. EMNLP. 2020

RankCP



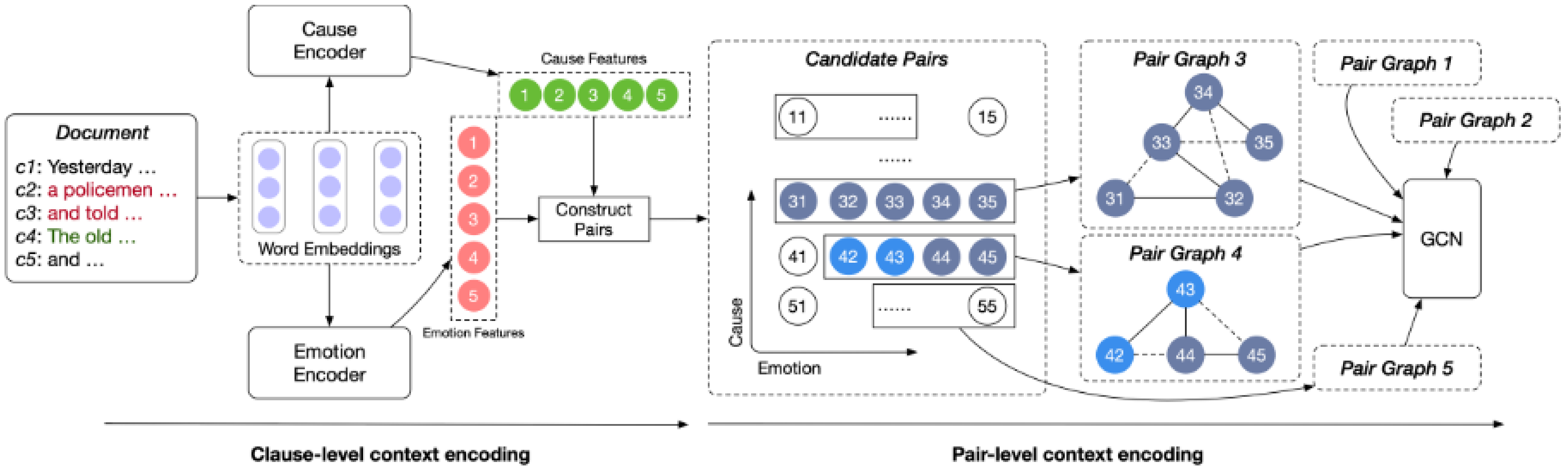
Effective Inter-Clause Modeling for End-to-End Emotion-Cause Pair Extraction. ACL. 2020

SLSN



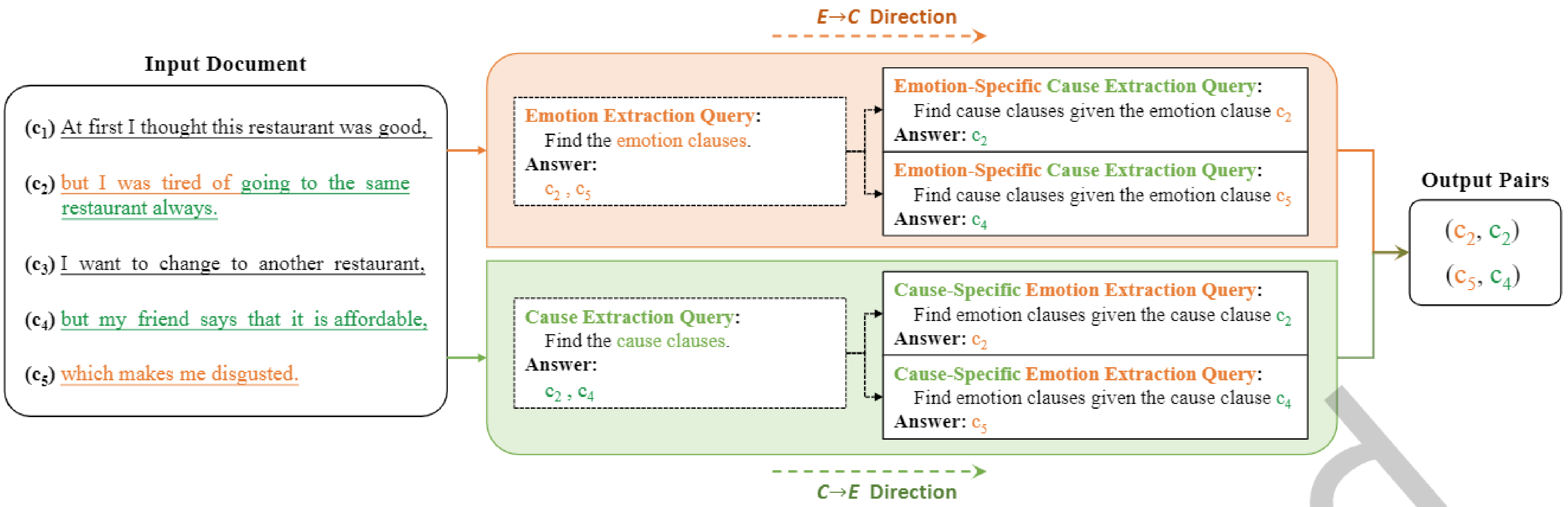
A Symmetric Local Search Network for Emotion-Cause Pair Extraction. COLING. 2020

PairGCN



End-to-End Emotion-Cause Pair Extraction with Graph Convolutional Network. COLING. 2020

DC-MRC



A. Exposure Bias
Use Pseudo Cause-Specific Emotion Extraction Query
Emotion-Specific Cause

B. Set Combination
1. Intersection 2.Union
3. Harmonic 4.Complementary

Leaderboard

Model		Emotion Extraction			Cause Extraction			Emotion-Cause Pair Extraction		
		P	R	F1	P	R	F1	P	R	F1
ECPE-2Step	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**
ECPE-2D	Inter-EC	0.8511	0.8237	0.8365	0.7133	0.6285	0.6672	0.7118	0.5984	0.6494
	+Bert	**0.8627**	**0.9221**	**0.8910**	**0.7336**	**0.6934**	**0.7123**	**0.7292**	**0.6544**	**0.6889**
ECPE-MLL	Inter-EC	0.8494	0.8297	0.8390	0.7256	0.6417	0.6799	0.6814	0.7257	0.6521
	ISML-6	0.8582	0.8429	0.8500	0.7248	0.6702	0.6950	0.7091	0.6441	0.6740
	+Bert	**0.8608**	**0.9191**	**0.8886**	**0.7382**	**0.7912**	**0.7630**	**0.7700**	**0.7235**	**0.7452**
SLSN (w/o Bert)		0.8406	0.7980	0.8118	0.6992	0.6588	0.6778	0.6836	0.6291	0.6545
PairGCN		0.8857	0.7958	0.8375	0.7907	0.6928	0.7375	0.7692	0.6791	0.7202
RankCP		0.9123	0.8999	0.9057	0.7461	0.7788	0.7615	0.7119	0.7630	0.7360
CD-MRC		0.9692	0.9398	0.9537	0.8101	0.8068	0.8077	0.8249	0.7800	0.8013

Sequence Tagging

	Yuan et al.	IE-CNN	UTOS
c_1 : Yesterday morning,	O	O	N
c_2^C: a police man visited the old man with the lost money,	C,2	C-H	1-C
c_3^C: and told him that the thief was caught.	C,1	C-H	1-C
c_4^E: The old man was very happy,	O	E-H	1-E
c_5 : and deposited the money in the bank.	O	O	O

Emotion-Cause Pair Extraction as Sequence Labeling Based on A Novel Tagging Scheme. EMNLP. 2020
A Unified Sequence Labeling Model for Emotion Cause Pair Extraction. COLING. 2020
A Unified Target-Oriented Sequence-to-Sequence Model for Emotion-Cause Pair Extraction. TASLP. .2021

TransECPE

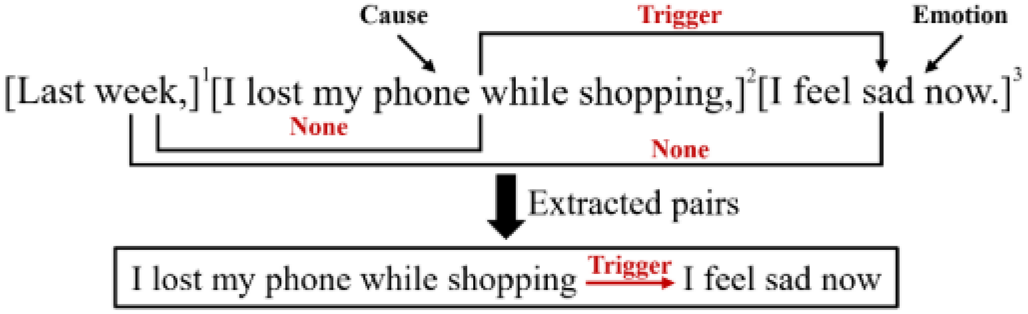
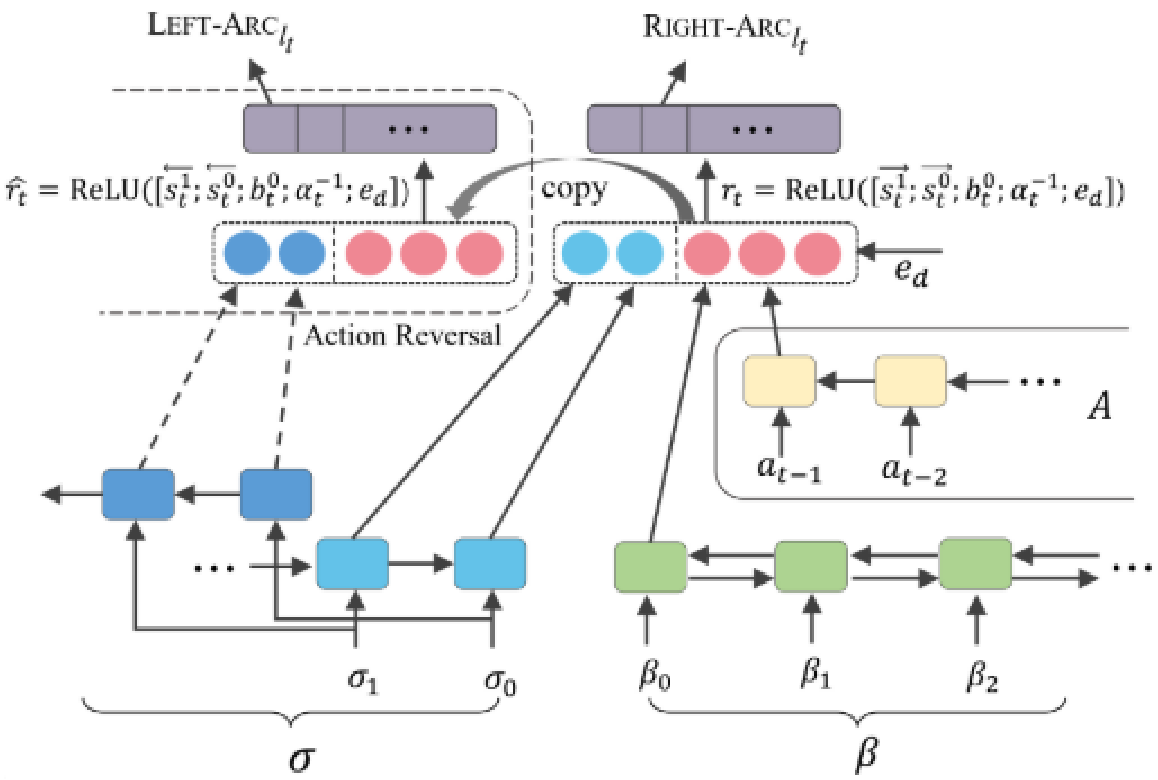


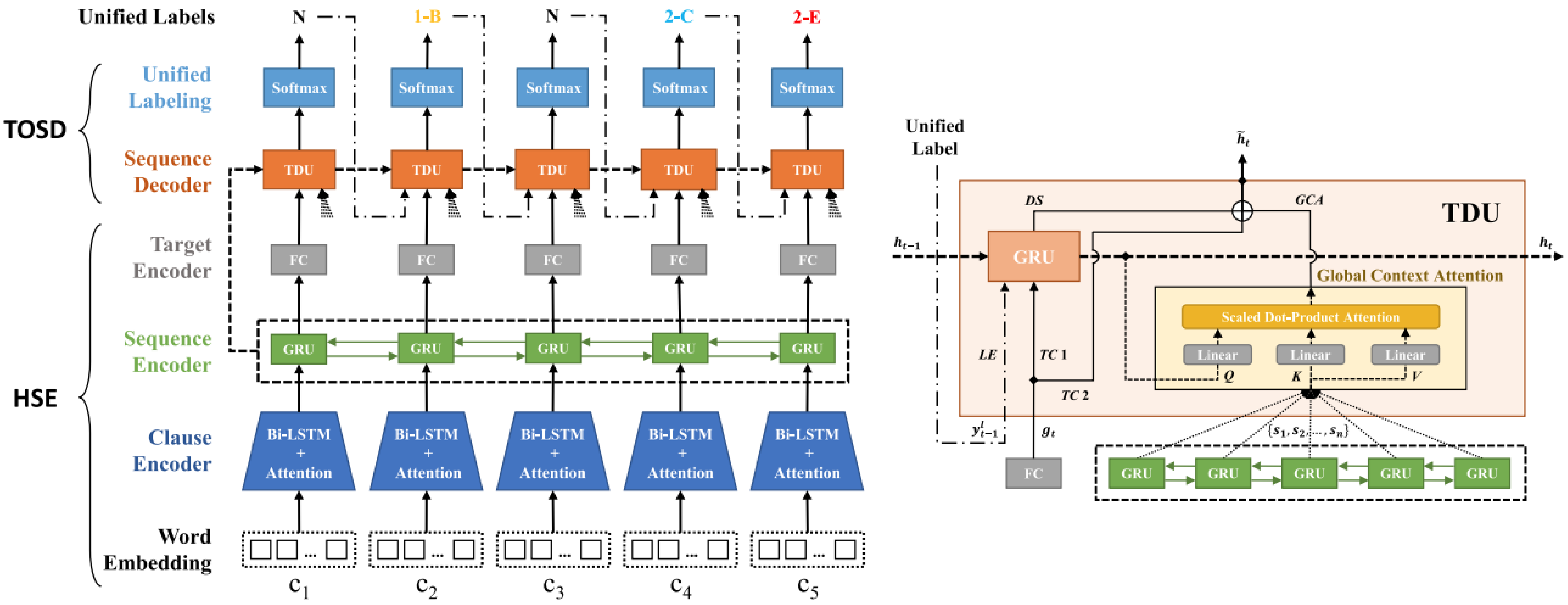
Figure 1: An example of emotion-cause pair extraction.

Stack σ	Buffer β	Action A	Emotion	Cause	Edge
[]	[1,2,3,$]	SH	$\varnothing$	$\varnothing$	$\varnothing$
[1]	[2,3,$]	SH	$\varnothing$	$\varnothing$	$\varnothing$
[1,2]	[3,$]	SH	$\varnothing$	$\varnothing$	$\varnothing$
[1,2,3]	[$]	RA_{l_t}	$\varnothing \cup \{3\}$	$\varnothing \cup \{2\}$	$2 \xrightarrow{l_t} 3$
[1,3]	[$]	RA_{l_n}	$\{3\} \cup \{3\}$	–	$1 \xrightarrow{l_n} 3$
[3]	[$]	SH	–	–	–
[3,$]	[]	–	–	–	–



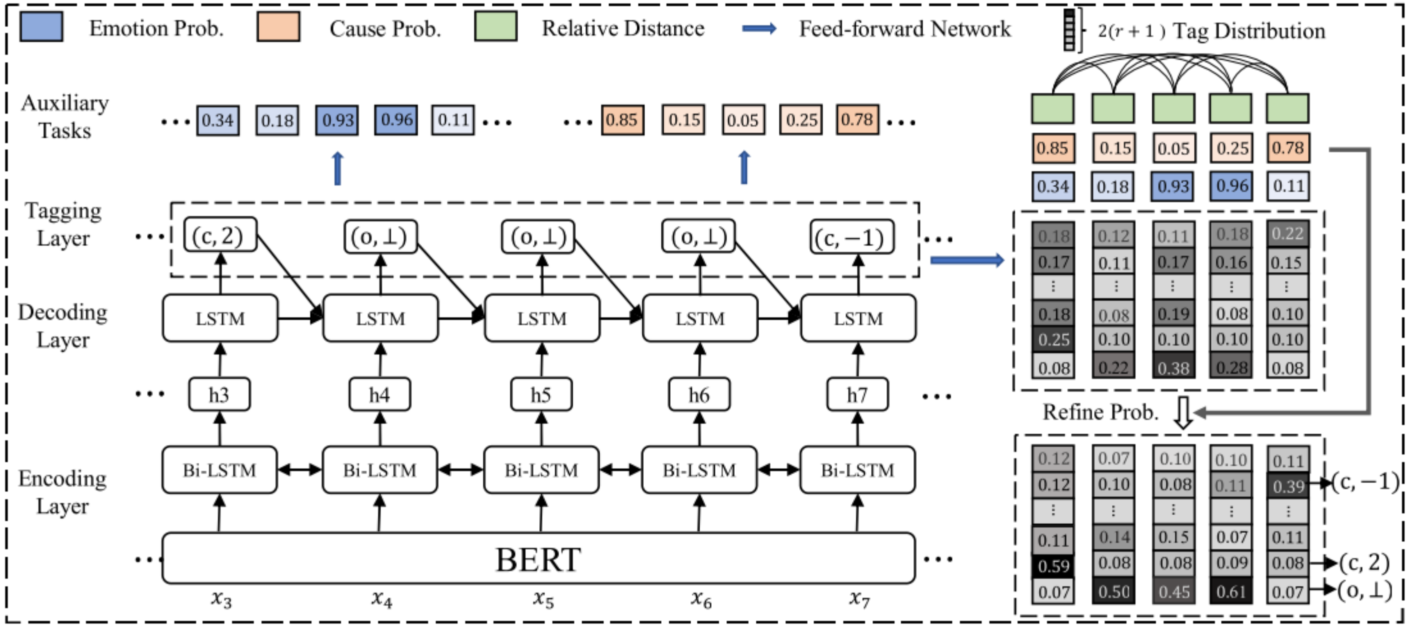
Transition-based Directed Graph Construction for Emotion-Cause Pair Extraction. ACL. 2020

UTOS



A Unified Target-Oriented Sequence-to-Sequence Model for Emotion-Cause Pair Extraction. TASLP(CCF-B). 2021

MTST



$$\tilde{\mathbf{p}}_k^t = \begin{cases} \mathbf{p}_\mathbf{k}^\mathbf{t} + \mathbf{v}_\mathbf{k} & p(y_i^e = 1|x_i) > 0.5 \\ \mathbf{p}_k^t - \mathbf{v}_k & p(y_i^e = 1|x_i) \leq 0.5 \end{cases}$$

If $p(y_i^e = 1|x_i) > 0.5$,

$$v_k^u = \begin{cases} \alpha_k \cdot (1 - p(y_k^t = j|x_k)) & u = j \\ -\alpha_k \cdot (1 - p(y_k^t = j|x_k))/(N_t - 1) & u \neq j \end{cases}$$

$$\alpha_k = w_k^i \cdot p(y_i^e = 1|x_i) \cdot p(y_k^c = 1|x_k)$$

when $p(y_i^e = 1|x_i) \leq 0.5$,

$$v_k^u = \begin{cases} \alpha_k \cdot p(y_k^t = j|x_k) & u = j \\ -\alpha_k \cdot p(y_k^t = j|x_k)/(N_t - 1) & u \neq j \end{cases}$$

$$\alpha_k = (1 - w_k^i) \cdot (1 - p(y_i^e = 1|x_i)) \cdot (1 - p(y_k^c = 1|x_k))$$

Model		Emotion Extraction			Cause Extraction			Emotion-Cause Pair Extraction		
		P	R	F1	P	R	F1	P	R	F1
ECPE-2Step	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**
Yuan et al.		0.8196	0.7329	0.7739	0.7490	0.6602	0.7018	0.7243	0.6366	0.6776
IE-CNN (w/o BERT)		0.8614	0.7811	0.8188	0.7348	0.5841	0.6496	0.7149	0.6279	0.6686
TransECPE		0.8716	0.8244	0.8474	0.7562	0.6471	0.6974	0.7374	0.6307	0.6799
	+Bert	**0.8879**	**0.8315**	**0.8588**	**0.7874**	**0.6689**	**0.7233**	**0.7708**	**0.6532**	**0.7072**
UTOS		0.8610	0.7925	0.8250	0.7189	0.6496	0.6802	0.6911	0.6193	0.6524
	+Bert	**0.8815**	**0.8321**	**0.8556**	**0.7671**	**0.7320**	**0.7471**	**0.7389**	**0.7062**	**0.7203**
MTST		0.8711	0.8178	0.8436	0.7947	0.7404	0.7666	0.7746	0.7109	0.7463

Question & Answer

Form	Explicit Connectives	Ambiguous Connectives
Verb	cause, result, arise, trigger	increase, affect, effect, make, induce, derive, reveal
Conjunction	because, so	hence, therefore, thus, thereby, since
Preposition	for, because of,	from, as, with, through, after
Adverb	accordingly, consequently	
Verb Phrase	result in(from), lead to	stem from, bring about, give rise to, thanks to
Prepositional Phrase	for the reason(alone), as a result of,	owing to, due to, in consequence of, in view of, as a consequence of, on account of, in as much as
Clause	that’s why, the result is, by reason that, so that	on this(that) account, in this way, in that, now that, on the grounds that, for fear that, if... then