A Unified Generative Framework for Aspect-Based Sentiment Analysis

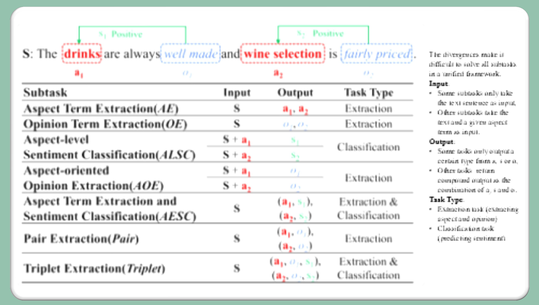
Hang Yan∗, Junqi Dai∗, Tuo Ji, Xipeng Qiu, Zheng Zhang
{hyan19, jqdai19, tji19, xpqiu}@fudan.edu.cn, zz@nyu.edu

FUDAN UNIVERSITY · FNLP

ABSA Subtasks

Input
Task Type
Output

"Divergence"



Subtask	Input	Output	Task Type
Aspect Term Extraction(*AE*)	S	a_1, a_2	Extraction
Opinion Term Extraction(*OE*)	S	o_1, o_2	Extraction
Aspect-level Sentiment Classification(*ALSC*)	S + a_1	s_1	Classification
	S + a_2	s_2	
Aspect-oriented Opinion Extraction(*AOE*)	S + a_1	o_1	Extraction
	S + a_2	o_2	
Aspect Term Extraction and Sentiment Classification(*AESC*)	S	(a_1, s_1), (a_2, s_2)	Extraction & Classification
Pair Extraction(*Pair*)	S	(a_1, o_1), (a_2, o_2)	Extraction
Triplet Extraction(*Triplet*)	S	(a_1, o_1, s_1), (a_2, o_2, s_2)	Extraction & Classification

The divergences make it difficult to solve all subtasks in a unified framework.

Input:
- Some subtasks only take the text sentence as input.
- Other subtasks take the text and a given aspect term as input.

Output:
- Some tasks only output a certain type from s, o or a.
- Other tasks return compound output as the combination of a, s and/or o.

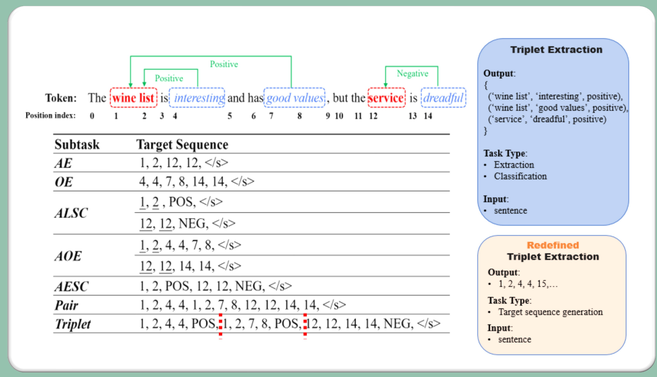
Task Type:
- Extraction task (extracting aspect and opinion)
- Classification task (predicting sentiment)

Unified

Generative

Framework

Unified Formulation



Subtask	Target Sequence
AE	1, 2, 12, 12, </s>
OE	4, 4, 7, 8, 14, 14, </s>
ALSC	1, 2 , POS, </s>
	12, 12, NEG, </s>
AOE	1, 2, 4, 4, 7, 8, </s>
	12, 12, 14, 14, </s>
AESC	1, 2, POS, 12, 12, NEG, </s>
Pair	1, 2, 4, 4, 1, 2, 7, 8, 12, 12, 14, 14, </s>
Triplet	1, 2, 4, 4, POS, 1, 2, 7, 8, POS, 12, 12, 14, 14, NEG, </s>

Triplet Extraction

Output:
{
('wine list', 'interesting', positive),
('wine list', 'good values', positive),
('service', 'dreadful', positive)
}

Task Type:
- Extraction
- Classification

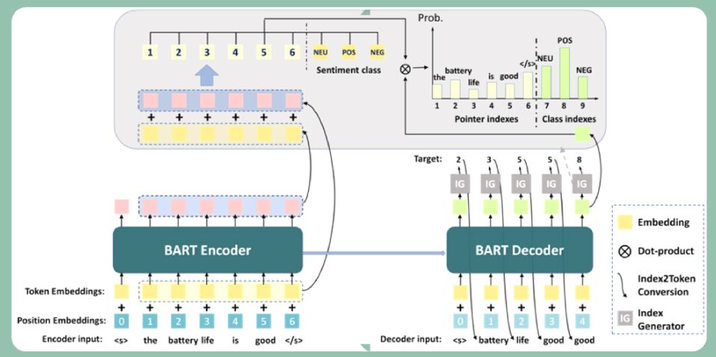
Input:
- sentence

Redefined Triplet Extraction

Output:
- 1, 2, 4, 4, 15,...

Task Type:
- Target sequence generation

Input:
- sentence



Generative Framework

A Unified ABSA Solution

Subtask	Input	Output	Task Type
Aspect Term Extraction(*AE*)	S	a_1, a_2	Extraction
Opinion Term Extraction(*OE*)	S	o_1, o_2	Extraction
Aspect-level Sentiment Classification(*ALSC*)	S + a_1	s_1	Classification
	S + a_2	s_2	
Aspect-oriented Opinion Extraction(*AOE*)	S + a_1	o_1	Extraction
	S + a_2	o_2	
Aspect Term Extraction and Sentiment Classification(*AESC*)	S	(a_1, s_1), (a_2, s_2)	Extraction & Classification
Pair Extraction(*Pair*)	S	(a_1, o_1), (a_2, o_2)	Extraction
Triplet Extraction(*Triplet*)	S	(a_1, o_1, s_1), (a_2, o_2, s_2)	Extraction & Classification

The divergences make it difficult to solve all subtasks in a unified framework.

Input:

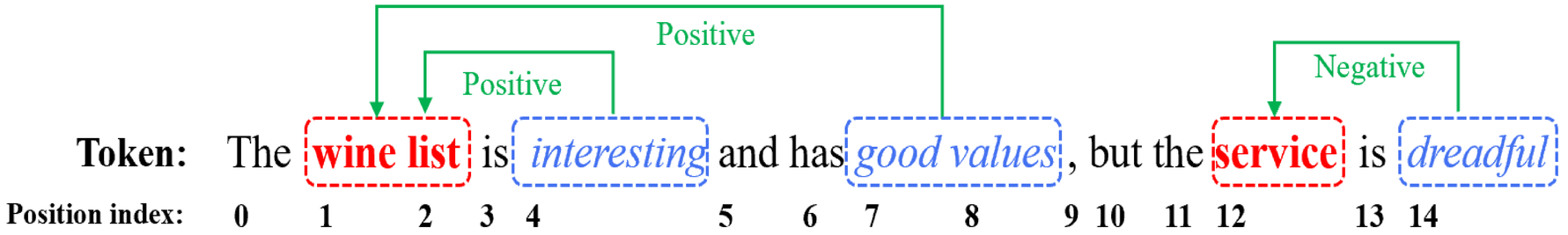
- Some subtasks only take the text sentence as input,
- Other subtasks take the text and a given aspect term as input.

Output:

- Some tasks only output a certain type from a, s or o,
- Other tasks return compound output as the combination of a, s and o.

Task Type:

- Extraction task (extracting aspect and opinion)
- Classification task (predicting sentiment)



Subtask	Target Sequence
AE	1, 2, 12, 12, </s>
OE	4, 4, 7, 8, 14, 14, </s>
ALSC	1, 2 , POS, </s>
	12, 12, NEG, </s>
AOE	1, 2, 4, 4, 7, 8, </s>
	12, 12, 14, 14, </s>
AESC	1, 2, POS, 12, 12, NEG, </s>
Pair	1, 2, 4, 4, 1, 2, 7, 8, 12, 12, 14, 14, </s>
Triplet	1, 2, 4, 4, POS, 1, 2, 7, 8, POS, 12, 12, 14, 14, NEG, </s>

Triplet Extraction

Output:
{
('wine list', 'interesting', positive),
('wine list', 'good values', positive),
('service', 'dreadful', positive)
}

Task Type:
- Extraction
- Classification

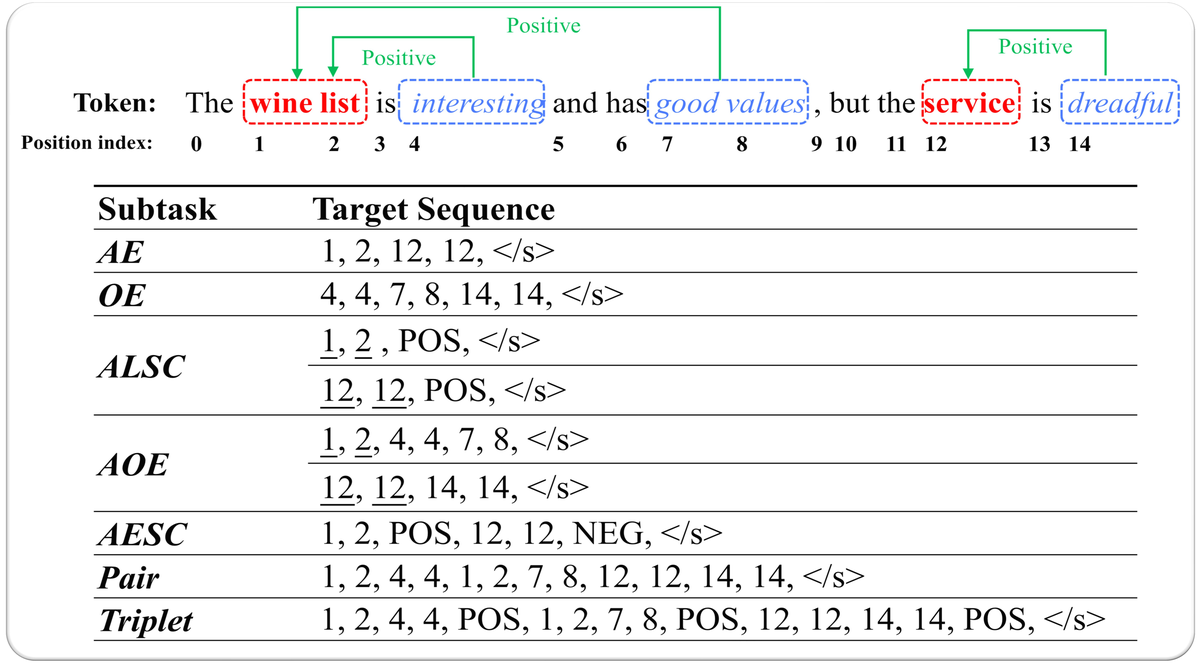
Input:
- sentence

Redefined
Triplet Extraction

Output:
- 1, 2, 4, 4, 15,…

Task Type:
- Target sequence generation

Input:
- sentence



Subtask	Target Sequence
AE	1, 2, 12, 12, </s>
OE	4, 4, 7, 8, 14, 14, </s>
ALSC	1, 2 , POS, </s>
	12, 12, POS, </s>
AOE	1, 2, 4, 4, 7, 8, </s>
	12, 12, 14, 14, </s>
AESC	1, 2, POS, 12, 12, NEG, </s>
Pair	1, 2, 4, 4, 1, 2, 7, 8, 12, 12, 14, 14, </s>
Triplet	1, 2, 4, 4, POS, 1, 2, 7, 8, POS, 12, 12, 14, 14, POS, </s>

~~Divergences~~

~~Output~~:
- ~~one output type~~
- ~~combination of a, s and o~~

~~Task Type~~:
- ~~Extraction~~
- ~~Classification~~

~~Input~~:
- ~~sentence~~
- ~~sentence and aspect term~~

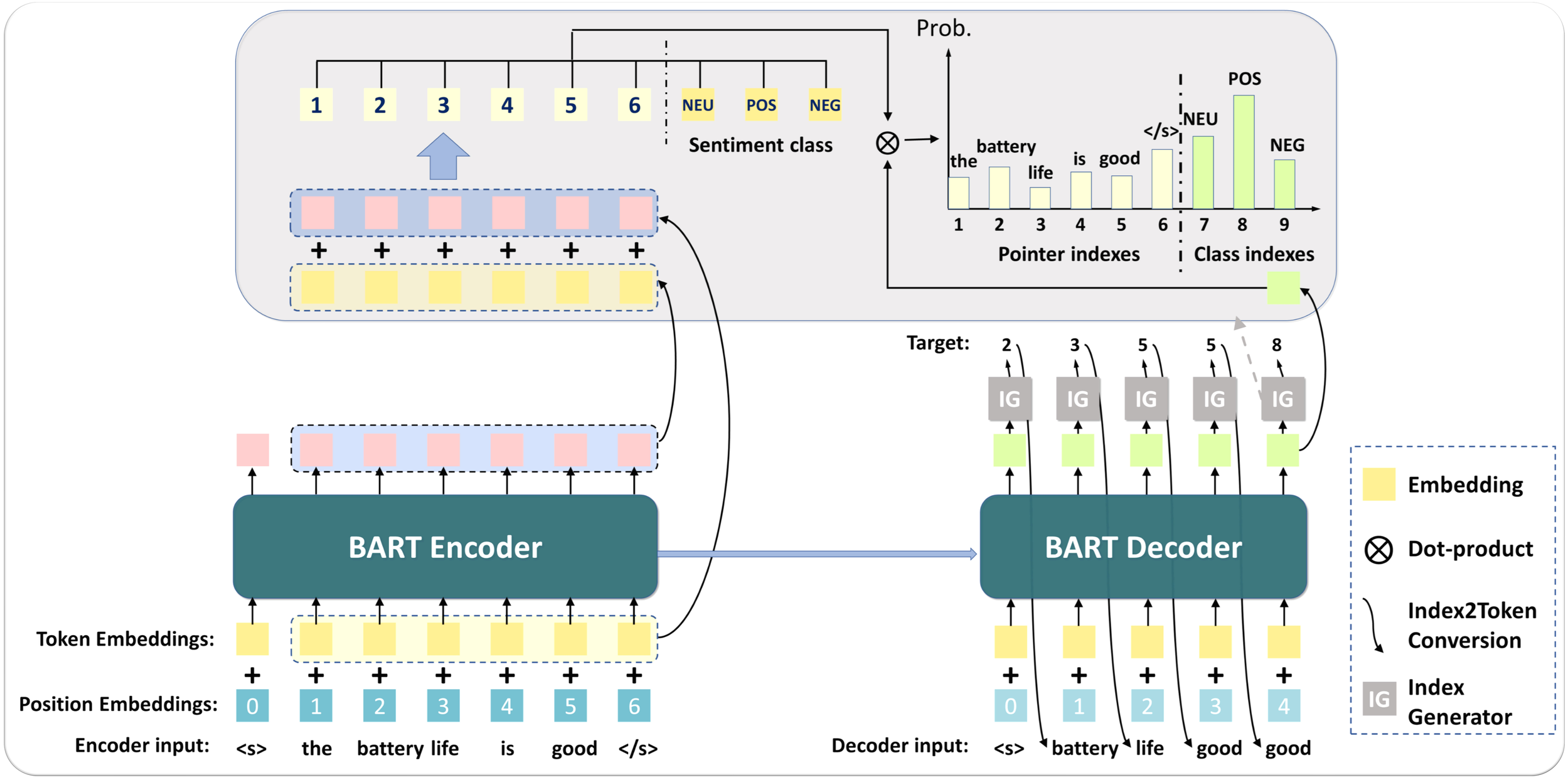
Unified Formulation

Unified Formulation

Output:
- Unified index sequence

Task Type:
- Unified generative task

Input:
- Unified sentence input


1
2
3
4
5
6
NEU
POS
NEG
Sentiment class
Prob.
the
battery
life
is
good
</s>
NEU
POS
NEG
1
2
3
4
5
6
7
8
9
Pointer indexes
Class indexes
Target:
2
3
5
5
8
IG
IG
IG
IG
IG
BART Encoder
BART Decoder
Token Embeddings:
Position Embeddings:
0
1
2
3
4
5
6
Encoder input:
<s>
the
battery
life
is
good
</s>
Decoder input:
<s>
battery
life
good
good
Embedding
Dot-product
Index2Token Conversion
IG
Index Generator

Datasets

Dataset		14res				14lap				15res				16res				Subtasks
		#s	*#a*	*#o*	*#p*	*#s*	*#a*	*#o*	*#p*	*#s*	*#a*	*#o*	*#p*	*#s*	*#a*	*#o*	*#p*	
$\mathcal{D}_{17}$	train	3044	3699	3484	-	3048	2373	2504	-	1315	1199	1210	-	-	-	-	-	*AE, OE, ALSC, AESC*
	test	800	1134	1008	-	800	654	674	-	685	542	510	-	-	-	-	-	
$\mathcal{D}_{19}$	train	1627	2643	-	-	1158	1634	-	-	754	1076	-	-	1079	1512	-	-	*AOE*
	test	500	865	-	-	343	482	-	-	325	436	-	-	329	457	-	-	
$\mathcal{D}_{20a}$	train	1300	-	-	2145	920	-	-	1265	593	-	-	923	842	-	-	1289	*AE, OE, ALSC, AOE, AESC, Pair, Triplet*
	dev	323	-	-	524	228	-	-	337	148	-	-	238	210	-	-	316	
	test	496	-	-	862	339	-	-	490	318	-	-	455	320	-	-	465	
$\mathcal{D}_{20b}$	train	1266	-	-	2338	906	-	-	1460	605	-	-	1013	857	-	-	1394	*AE, OE, ALSC, AOE, AESC, Pair, Triplet*
	dev	310	-	-	577	219	-	-	346	148	-	-	249	210	-	-	339	
	test	492	-	-	994	328	-	-	543	148	-	-	485	326	-	-	514	

Datasets

Baselines

Baselines	E2E	Task Formulation	Backbone	Datasets	*AE*	*OE*	*ALSC*	*AOE*	*AESC*	*Pair*	*Triplet*
SPAN-BERT	-	Span.Extraction	BERT	$\mathcal{D}_{17}$	✓	-	✓	-	✓	-	-
IMN-BERT	✓	Seq.Tagging	BERT	$\mathcal{D}_{17}$	✓	✓	✓	-	✓	-	-
RACL-BERT	-	Seq.Tagging	BERT	$\mathcal{D}_{17}$	✓	✓	✓	-	✓	-	-
IOG	✓	Seq.Tagging	LSTM	$\mathcal{D}_{19}$	-	-	-	✓	-	-	-
LOTN	✓	Seq.Tagging	LSTM	$\mathcal{D}_{19}$	-	-	-	✓	-	-	-
ONG	✓	Seq.Tagging	BERT	$\mathcal{D}_{19}$	-	-	-	✓	-	-	-
RINANTE+	-	Seq.Tagging	LSTM+CRF	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
CMLA+	-	Seq.Tagging	Attention	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
Li-unified+	-	Seq.Tagging	LSTM	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
Peng-two-stage	-	Seq.Tagging	LSTM+GCN	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
JET-BERT	✓	Seq.Tagging	BERT	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
Dual-MRC	-	Span.MRC	BERT	$\mathcal{D}_{17}$,$\mathcal{D}_{19}$,$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	-	✓	✓	✓	✓	✓
Ours	✓	Span.Generation	BART	$\mathcal{D}_{17}$,$\mathcal{D}_{19}$,$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	✓	✓	✓	✓

Baselines

Model	14res			14lap			15res			16res		
	AE	*OE*	*ALSC*	*AESC*	*AE*	*OE*	*ALSC*	*AESC*	*AE*	*OE*	*ALSC*	*AESC*
SPAN-BERT	86.71	-	71.75	73.68	82.34	-	-	62.5	61.25	74.63	-	50.28	62.29
IMN-BERT	84.06	85.10	75.67	70.72	77.55	**81.0**	75.56	61.73	69.90	73.29	70.10	60.22
RACL-BERT	86.38	87.18	81.61	75.42	81.79	79.72	73.91	63.40	73.99	76.0	**74.91**	66.05
Dual-MRC	86.60	-	**82.04**	**75.95**	82.51	-	75.97	65.94	75.08	-	73.59	65.08
Ours	**87.07**	**87.29**	75.56	73.56	**83.52**	77.86	**76.76**	**67.37**	**75.48**	**76.49**	73.91	**66.61**

Model	14res			14lap			15res			16res		
	AESC	*Pair*	*Triplet.*	*AESC*	*Pair*	*Triplet.*	*AESC*	*Pair*	*Triplet.*	*AESC*	*Pair*	*Triplet.*
CMLA+ †	70.62	48.95	43.12	56.90	44.10	32.90	53.60	44.60	35.90	61.20	50.00	41.60
RINANTE+ †	48.15	46.29	34.03	36.70	29.70	20.0	41.30	35.40	28.0	42.10	30.70	23.30
Li-unified+ †	73.79	55.34	51.68	63.38	52.56	42.47	64.95	56.85	46.69	70.20	53.75	44.51
Peng-two-stage †	74.19	56.10	51.89	62.34	53.85	43.50	65.79	56.23	46.79	71.73	60.04	53.62
JET-BERT ‡	-	-	63.92	-	-	50.0	-	-	54.67	-	-	62.98
Dual-MRC†	76.57	74.93	70.32	64.59	63.37	55.58	65.14	64.97	57.21	70.84	75.71	67.40
Ours	**78.47**	**77.68**	**72.46**	**68.17**	**66.11**	**57.59**	**69.95**	**67.98**	**60.11**	**75.69**	**77.38**	**69.98**

Model	14res			14lap			15res			16res		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
CMLA+	39.18	47.13	42.79	30.09	36.92	33.16	34.56	39.84	37.01	41.34	42.1	41.72
RINANTE+	31.42	39.38	34.95	21.71	18.66	20.07	29.88	30.06	29.97	25.68	22.3	23.87
Li-unified+	41.04	**67.35**	51.0	40.56	44.28	42.34	44.72	51.39	47.82	37.33	54.51	44.31
Peng-two-stage	43.24	63.66	51.46	37.38	50.38	42.87	48.07	57.51	52.32	46.96	64.24	54.21
JET-BERT	**70.56**	55.94	62.40	55.39	47.33	51.04	**64.45**	51.96	57.53	**70.42**	58.37	63.83
Ours	65.52	64.99	**65.25**	**61.41**	**56.19**	**58.69**	59.14	**59.38**	**59.26**	66.6	**68.68**	**67.62**

Model	14res			14lap			15res			16res		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
IOG	82.38	78.25	80.23	73.43	68.74	70.99	72.19	71.76	71.91	84.36	79.08	81.60
LOTN	84.0	80.52	82.21	77.08	67.62	72.02	76.61	70.29	73.29	86.57	80.89	83.62
ONG	83.23	81.46	82.33	73.87	77.78	75.77	76.63	**81.14**	78.81	87.72	84.38	86.01
Dual-MRC	**89.79**	78.43	83.73	78.21	**81.66**	79.90	77.19	71.98	74.50	86.07	80.77	83.33
Ours	86.01	**84.76**	**85.38**	**83.11**	78.13	**80.55**	**80.12**	80.93	**80.52**	**89.22**	**86.67**	**87.92**

D_{17}

D_{20a}

D_{20b}

D_{19}

Results

Datasets

Dataset		14res				14lap				15res				16res				Subtasks
		#s	#a	#o	#p	#s	#a	#o	#p	#s	#a	#o	#p	#s	#a	#o	#p	
$\mathcal{D}_{17}$	train	3044	3699	3484	-	3048	2373	2504	-	1315	1199	1210	-	-	-	-	-	*AE*, *OE*, *ALSC*,
	test	800	1134	1008	-	800	654	674	-	685	542	510	-	-	-	-	-	*AESC*
$\mathcal{D}_{19}$	train	1627	2643	-	-	1158	1634	-	-	754	1076	-	-	1079	1512	-	-	*AOE*
	test	500	865	-	-	343	482	-	-	325	436	-	-	329	457	-	-	
$\mathcal{D}_{20a}$	train	1300	-	-	2145	920	-	-	1265	593	-	-	923	842	-	-	1289	*AE*, *OE*, *ALSC*, *AOE*, *AESC*, *Pair*, *Triplet*
	dev	323	-	-	524	228	-	-	337	148	-	-	238	210	-	-	316	
	test	496	-	-	862	339	-	-	490	318	-	-	455	320	-	-	465	
$\mathcal{D}_{20b}$	train	1266	-	-	2338	906	-	-	1460	605	-	-	1013	857	-	-	1394	*AE*, *OE*, *ALSC*, *AOE*, *AESC*, *Pair*, *Triplet*
	dev	310	-	-	577	219	-	-	346	148	-	-	249	210	-	-	339	
	test	492	-	-	994	328	-	-	543	148	-	-	485	326	-	-	514	

Baselines

Baselines	E2E	Task Formulation	Backbone	Datasets	*AE*	*OE*	*ALSC*	*AOE*	*AESC*	*Pair*	*Triplet*
SPAN-BERT	-	Span.Extraction	BERT	$\mathcal{D}_{17}$	✓	-	✓	-	✓	-	-
IMN-BERT	✓	Seq.Tagging	BERT	$\mathcal{D}_{17}$	✓	✓	✓	-	✓	-	-
RACL-BERT	-	Seq.Tagging	BERT	$\mathcal{D}_{17}$	✓	✓	✓	-	✓	-	-
IOG	✓	Seq.Tagging	LSTM	$\mathcal{D}_{19}$	-	-	-	✓	-	-	-
LOTN	✓	Seq.Tagging	LSTM	$\mathcal{D}_{19}$	-	-	-	✓	-	-	-
ONG	✓	Seq.Tagging	BERT	$\mathcal{D}_{19}$	-	-	-	✓	-	-	-
RINANTE+	-	Seq.Tagging	LSTM+CRF	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
CMLA+	-	Seq.Tagging	Attention	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
Li-unified+	-	Seq.Tagging	LSTM	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
Peng-two-stage	-	Seq.Tagging	LSTM+GCN	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
JET-BERT	✓	Seq.Tagging	BERT	$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	-	✓	✓	✓
Dual-MRC	-	Span.MRC	BERT	$\mathcal{D}_{17}$,$\mathcal{D}_{19}$,$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	-	✓	✓	✓	✓	✓
Ours	✓	Span.Generation	BART	$\mathcal{D}_{17}$,$\mathcal{D}_{19}$,$\mathcal{D}_{20a}$,$\mathcal{D}_{20b}$	✓	✓	✓	✓	✓	✓	✓

Model	14res				14lap				15res			
	AE	*OE*	*ALSC*	*AESC*	*AE*	*OE*	*ALSC*	*AESC*	*AE*	*OE*	*ALSC*	*AESC*
SPAN-BERT	86.71	-	71.75	73.68	82.34	-	62.5	61.25	74.63	-	50.28	62.29
IMN-BERT	84.06	85.10	75.67	70.72	77.55	**81.0**	75.56	61.73	69.90	73.29	70.10	60.22
RACL-BERT	86.38	87.18	81.61	75.42	81.79	79.72	73.91	63.40	73.99	76.0	**74.91**	66.05
Dual-MRC	86.60	-	**82.04**	**75.95**	82.51	-	75.97	65.94	75.08	-	73.59	65.08
Ours	**87.07**	**87.29**	75.56	73.56	**83.52**	77.86	**76.76**	**67.37**	**75.48**	**76.49**	73.91	**66.61**

D17

Model	14res			14lap			15res			16res		
	AESC	*Pair*	*Triple.*	*AESC*	*Pair*	*Triple.*	*AESC*	*Pair*	*Triple.*	*AESC*	*Pair*	*Triple.*
CMLA+ †	70.62	48.95	43.12	56.90	44.10	32.90	53.60	44.60	35.90	61.20	50.00	41.60
RINANTE+ †	48.15	46.29	34.03	36.70	29.70	20.0	41.30	35.40	28.0	42.10	30.70	23.30
Li-unified+ †	73.79	55.34	51.68	63.38	52.56	42.47	64.95	56.85	46.69	70.20	53.75	44.51
Peng-two-stage †	74.19	56.10	51.89	62.34	53.85	43.50	65.79	56.23	46.79	71.73	60.04	53.62
JET-BERT ♯	-	-	63.92	-	-	50.0	-	-	54.67	-	-	62.98
Dual-MRC†	76.57	74.93	70.32	64.59	63.37	55.58	65.14	64.97	57.21	70.84	75.71	67.40
Ours	**78.47**	**77.68**	**72.46**	**68.17**	**66.11**	**57.59**	**69.95**	**67.98**	**60.11**	**75.69**	**77.38**	**69.98**

D20a

Model	14res			14lap			15res			16res		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
CMLA+	39.18	47.13	42.79	30.09	36.92	33.16	34.56	39.84	37.01	41.34	42.1	41.72
RINANTE+	31.42	39.38	34.95	21.71	18.66	20.07	29.88	30.06	29.97	25.68	22.3	23.87
Li-unified+	41.04	**67.35**	51.0	40.56	44.28	42.34	44.72	51.39	47.82	37.33	54.51	44.31
Peng-two-stage	43.24	63.66	51.46	37.38	50.38	42.87	48.07	57.51	52.32	46.96	64.24	54.21
JET-BERT	**70.56**	55.94	62.40	55.39	47.33	51.04	**64.45**	51.96	57.53	**70.42**	58.37	63.83
Ours	65.52	64.99	**65.25**	**61.41**	**56.19**	**58.69**	59.14	**59.38**	**59.26**	66.6	**68.68**	**67.62**

D20b

Model	14res			14lap			15res			16res		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
IOG	82.38	78.25	80.23	73.43	68.74	70.99	72.19	71.76	71.91	84.36	79.08	81.60
LOTN	84.0	80.52	82.21	77.08	67.62	72.02	76.61	70.29	73.29	86.57	80.89	83.62
ONG	83.23	81.46	82.33	73.87	77.78	75.77	76.63	**81.14**	78.81	87.72	84.38	86.01
Dual-MRC	**89.79**	78.43	83.73	78.21	**81.66**	79.90	77.19	71.98	74.50	86.07	80.77	83.33
Ours	86.01	**84.76**	**85.38**	**83.11**	78.13	**80.55**	**80.12**	80.93	**80.52**	**89.22**	**86.67**	**87.92**

D19

Thanks !

Q & A

Paper link : https://arxiv.org/abs/2106.04300
Github link : https://github.com/yhcc/BARTABSA

Reporter: jqdai19@fudan.edu.cn

ACL 2021