Making Better Use of Training Corpus: Retrieval-based Aspect Sentiment Triplet Extraction via Label Interpolation

Guoxin Yu[1,2], Lemao Liu[3*], Haiyun Jiang[3], Shuming Shi[3], Xiang Ao[1,2*]

1 Key Lab of Intelligent Information Processing of Chinese Academy of Sciences
(CAS), Institute of Computing Technology, CAS, Beijing 100190, China.
2 University of Chinese Academy of Sciences, Beijing 100049, China.
3 Tencent AI Lab, China.

Aspect Sentiment Triplets Extraction(ASTE) aims to extract all triplets containing aspect terms, the corresponding opinion terms, and their sentiment polarities.

An example sentence:

The **food** is **spicy** and **delicious** ... including ...
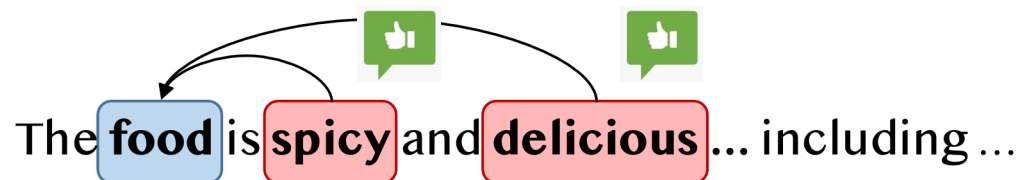
Extracted Results: (**food**, **spicy**,positive) (**food**, **delicious**,positive)...

The intra-sentence information is insufficient.

Long distance

The **food** is **spicy** and **delicious**, including the **shrimp appetizer**, the **cod**, and the **scallop roll**.

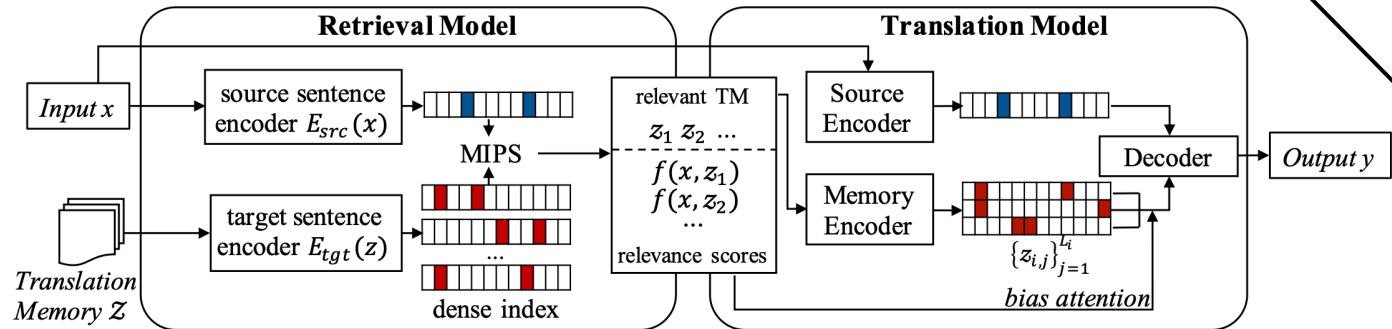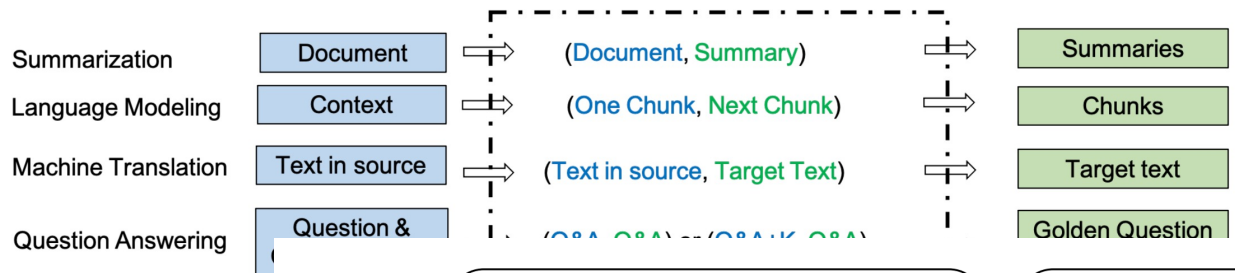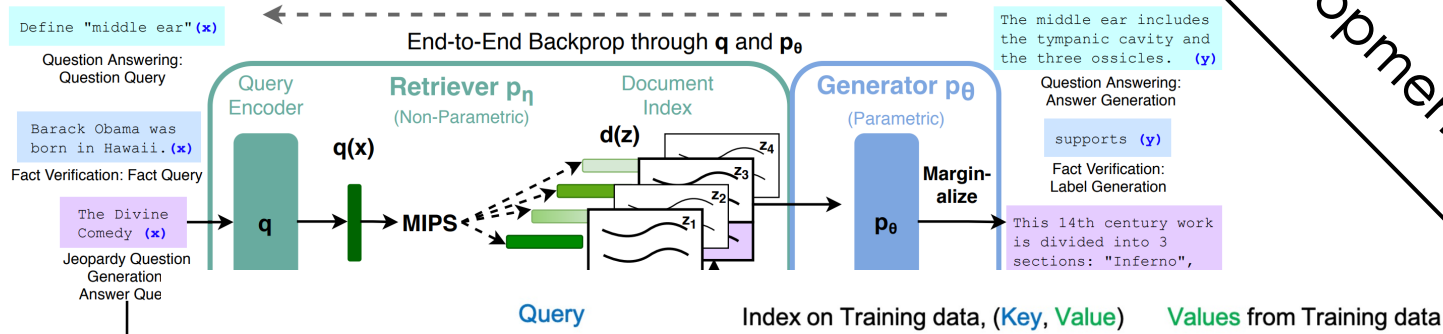An uncommon aspect term

Knowledge Source

💡 It is promising to introduce inter-sentence information.

Define "middle ear" **(x)**

Question Answering: Question Query

Barack Obama was born in Hawaii. **(x)**

Fact Verification: Fact Query

The Divine Comedy **(x)**

Jeopardy Question Generation Answer Que

End-to-End Backprop through **q** and **p_θ**

Query Encoder
Retriever p_η (Non-Parametric)
Document Index
d(z)
Generator p_θ (Parametric)

q(x)

q → MIPS → z_4 z_3 z_2 z_1 → **p_θ** **Marginalize**

The middle ear includes the tympanic cavity and the three ossicles. **(y)**

Question Answering: Answer Generation

supports **(y)**

Fact Verification: Label Generation

This 14th century work is divided into 3 sections: "Inferno",

	Query	Index on Training data, (Key, Value)		Values from Training data
Summarization	Document	⇒	(Document, Summary) ⇒	Summaries
Language Modeling	Context	⇒	(One Chunk, Next Chunk) ⇒	Chunks
Machine Translation	Text in source	⇒	(Text in source, Target Text) ⇒	Target text
Question Answering	Question &		(Q&A, Q&A) or (Q&A+K, Q&A)	Golden Question

The development of retrieval augmented methods

Retrieval Model

Input x → source sentence encoder $E_{src}(x)$ → MIPS

Translation Memory $\mathcal{Z}$ → target sentence encoder $E_{tgt}(z)$ → ... → dense index

relevant TM
z_1 z_2 ...
$f(x, z_1)$
$f(x, z_2)$
...
relevance scores

Translation Model

Source Encoder → Decoder → Output y

Memory Encoder → $\{z_{i,j}\}_{j=1}^{L_i}$ → bias attention
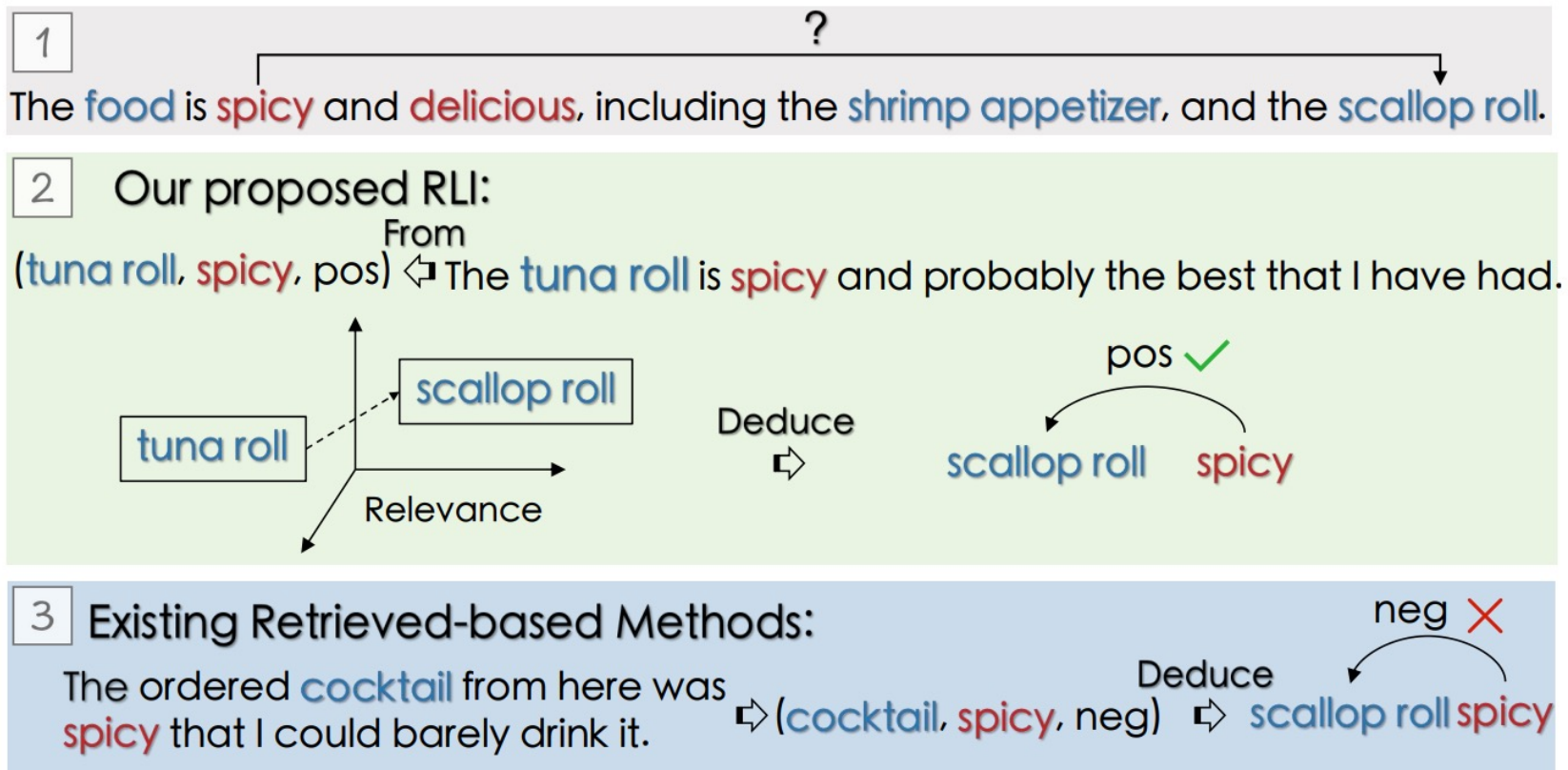
ASTE has its own challenging when adapting retrieval-augmented methods:
its purpose includes predicting the sentiment polarities and it is usually aspect-dependent.
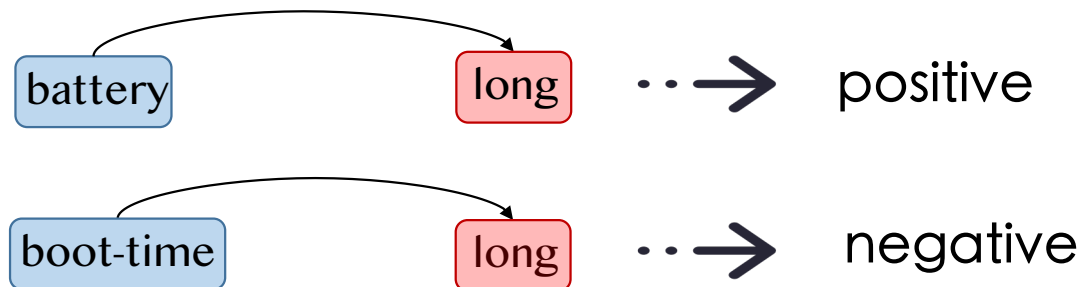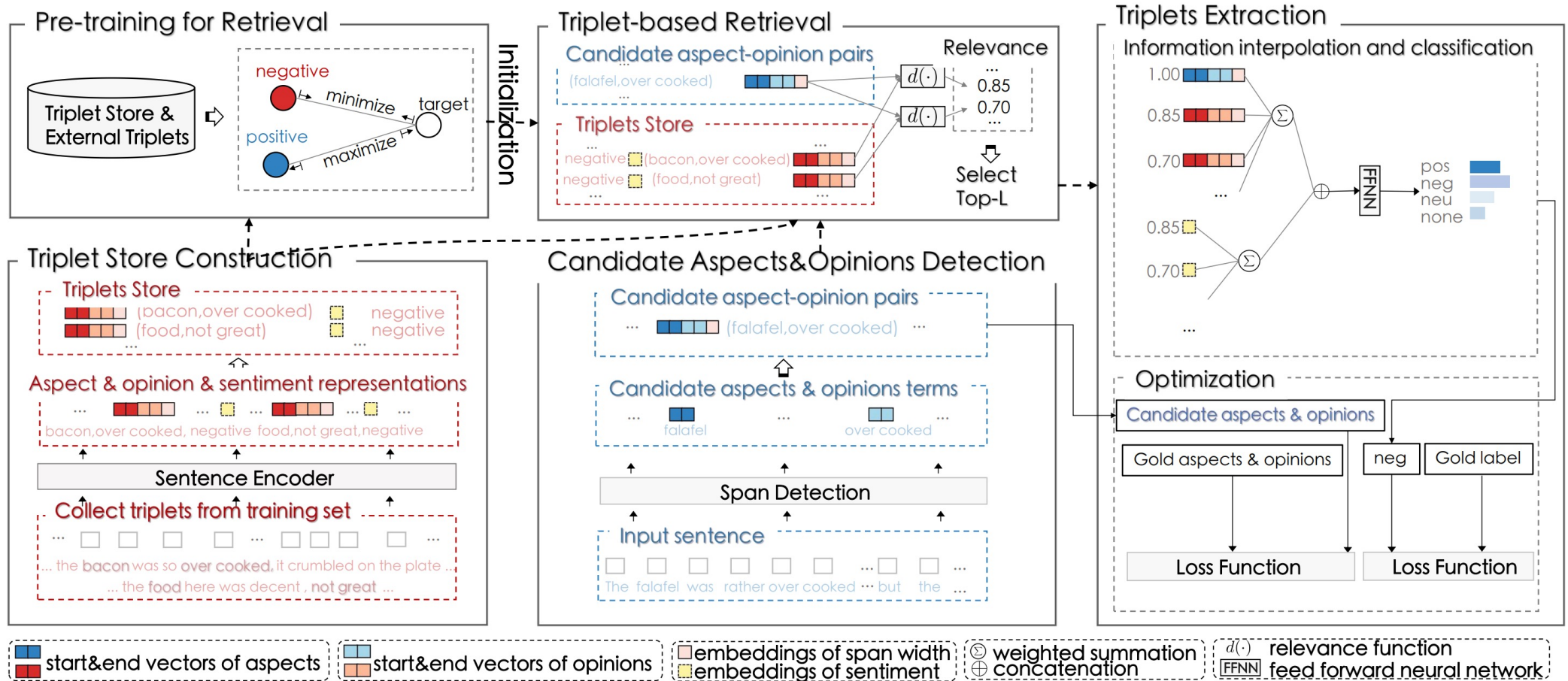

Triplets with similar semantic similarities may have conflict sentiment polarities.

We proposed a ASTE approach named RLI (Retrieval-based Aspect Sentiment Triplets Extraction via Label Interpolation).

1 ?
The food is spicy and delicious, including the shrimp appetizer, and the scallop roll.

2 Our proposed RLI:
From
(tuna roll, spicy, pos) ⇦ The tuna roll is spicy and probably the best that I have had.

scallop roll

tuna roll

Relevance

Deduce ⇨

pos ✓

scallop roll   spicy

3 Existing Retrieved-based Methods:
The ordered cocktail from here was spicy that I could barely drink it. ⇨ (cocktail, spicy, neg) ⇨ Deduce neg ✗ scallop roll spicy

ASTE-DATA-V1

Model	Res14		Lap14		Res15		Res16	
	Pair	*Triplet*	*Pair*	*Triplet*	*Pair*	*Triplet*	*Pair*	*Triplet*
WhatHowWhy$^{\diamond}$	56.10	51.89	53.85	43.50	56.23	46.79	60.04	53.62
CMLA+$^{\diamond}$	48.95	43.12	44.10	32.90	44.60	35.90	50.00	41.60
RINANTE+$^{\diamond}$	46.29	34.03	29.70	20.00	35.40	28.00	30.70	23.30
Unified+$^{\diamond}$	55.34	51.68	52.56	42.47	56.85	46.69	53.75	44.51
Dual-MRC$^{\diamond}$	74.93	70.32	63.37	55.58	64.97	57.21	75.71	67.40
Generative$^{\triangledown}$	77.68	72.46	66.11	57.59	67.98	60.11	77.38	69.98
GAS$^{\sharp}$	–	70.20	–	54.50	–	59.10	–	65.00
LEGO$^{\sharp}$	–	72.60	–	59.50	–	63.20	–	71.50
JET$^{t}_{M=6}$ $^{\triangledown}$	–	60.41	–	46.65	–	53.68	–	63.41
JET$^{o}_{M=6}$ $^{\triangledown}$	–	63.92	–	50.00	–	54.67	–	62.98
SPAN*	78.62	73.96	69.48	60.59	71.56	64.50	78.85	70.48
RLI(Ours)	**79.92**	**74.98**	**70.27**	**61.97**	**72.66**	**65.71**	**81.29**	**73.33**

ASTE-DATA-V2

Model	Res14			Lap14			Res15			Res16		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
WhatHowWhy$^{\diamond}$	43.24	63.66	51.46	37.38	50.38	42.87	48.07	57.51	52.32	46.96	64.24	54.21
TOP$^{\sharp}$	63.59	73.44	68.16	57.84	**59.33**	58.58	54.53	63.30	58.59	63.57	71.98	67.52
BMRC$^{\sharp}$	72.17	65.43	68.64	**65.91**	52.15	58.18	62.48	55.55	58.79	69.87	65.68	67.35
RL$^{\sharp}$	70.60	68.65	69.61	64.80	54.99	59.50	65.45	60.29	62.72	67.21	69.69	68.41
GAS$^{\flat}$	–	–	72.16	–	–	60.78	–	–	62.10	–	–	70.10
OTE-MTL$^{\sharp}$	63.07	58.25	60.56	54.26	41.07	46.75	60.88	42.68	50.18	65.65	54.28	59.42
GTS$^{\diamond}$	67.76	67.29	67.50	57.82	51.32	54.36	62.59	57.94	60.15	66.08	69.91	67.93
JET$^{t}_{M=6}$ $^{\diamond}$	63.44	54.12	59.41	53.53	43.28	47.86	**68.20**	42.89	52.66	65.28	51.95	63.83
JET$^{o}_{M=6}$ $^{\diamond}$	70.56	55.94	62.40	55.39	47.33	51.04	64.45	51.96	57.53	70.42	58.37	63.83
SPAN$^{\diamond}$	72.89	70.89	71.85	63.44	55.84	59.38	62.18	64.45	63.27	69.45	71.17	70.26
EMC-GCN$^{\triangledown}$	71.21	**72.39**	71.78	61.70	56.26	58.81	61.54	62.47	61.93	65.62	71.30	68.33
RLI (Ours)	**77.46**	71.97	**74.34**	63.32	57.43	**60.96**	60.08	**70.66**	**65.41**	70.50	**74.28**	**72.34**

Dataset	Model		Dev F1	Test F1
Res14	w/o joint		66.85	72.07
	joint	w/o sentiment	67.55	73.70
		w/o pre-training	67.12	72.58
		full model	**68.00**	**74.34**
Lap14	w/o joint		60.03	60.33
	joint	w/o sentiment	61.90	60.54
		w/o pre-training	61.06	60.02
		full model	**62.55**	**60.96**
Res15	w/o joint		70.83	63.99
	joint	w/o sentiment	71.54	65.04
		w/o pre-training	71.24	64.48
		full model	**72.21**	**65.41**
Res16	w/o joint		70.47	70.30
	joint	w/o sentiment	71.44	71.39
		w/o pre-training	70.75	71.69
		full model	**73.04**	**72.34**

Table 3: Ablation test. The displayed scores are *F1 score for triplets extraction* on $\mathcal{D}_b$.

Models	Base	Base+Aug	Ours
Avg. F1	66.19	66.99(+0.80)	**68.26**(+2.07)

Table 4: Average *F1 score for triplets extraction* on $\mathcal{D}_b$.
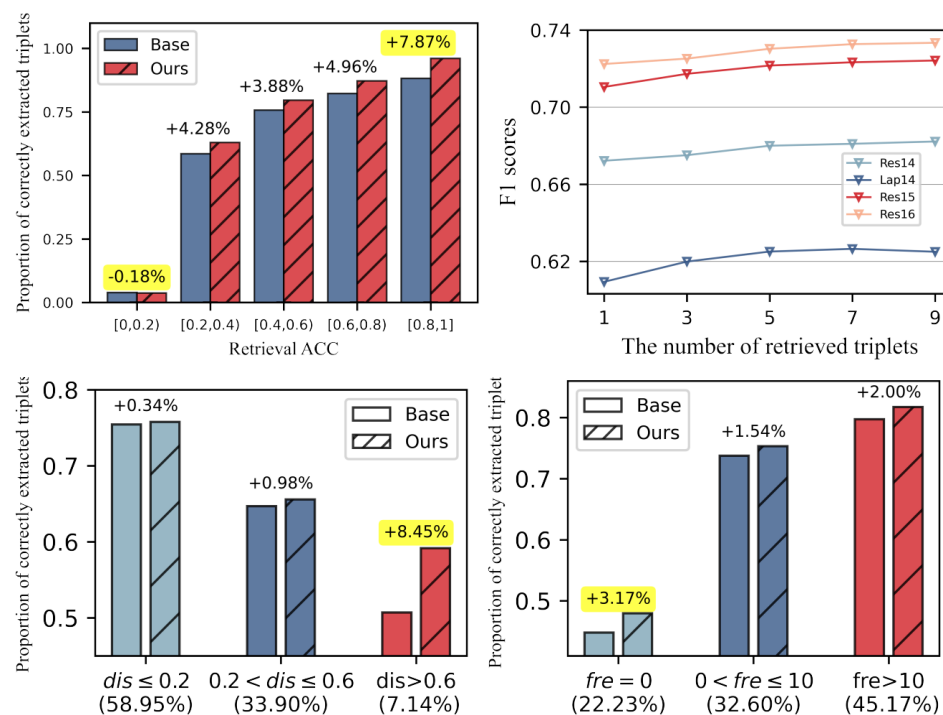
Sentence	BEFORE	AFTER	RETRIEVED
Long distance Service was good and so was the atmosphere.	(Service, good, pos) ✓ (null) ✗	(Service, good, pos) ✓ (atmosphere, good, pos) ✓ ←	From beginning appetizers, the scallops were incredible… (beginning appetizers, incredible, pos)
And these are not small, wimpy, fast food type burgers - these are real, full sized patties. **Uncommon aspect and opinion**	(null) ✗	(patties, real, pos) ✓ ←	… you go for the large amounts of food , the amiable atmosphere, … (atmosphere, aimable, pos)
	(null) ✗	(patties, full sized, pos) ✓ ←	The Prix Fixe menu is worth every penny and you get more than enough (both in quantity AND quality) . (quality, more than enough, pos)
Overlapped triplets & **Long distance** The icing made this cake, it was flurry, not ultra sweet, creamy and light.	(icing, flurry, pos) ✗	(cake, flurry, pos) ✓ ←	and yes Dal Bukhara is so dam good and so are all the kababs. (dal bukhara, good, pos)
	(cake, not ultra sweet, neg) ✗	(cake, not ultra sweet, pos) ✓ ←	I plan to come here again and look forward to trying their assortment of bruschetta. (bruschetta, look forward, pos)
	(null) ✗	(cake, creamy, pos) ✓ ←	Highly recommended is the Spicy Fried Clam Rolls and Spider Rolls. (spider rolls, recommended, pos)
	(null) ✗	(cake, light, pos) ✓ ←	Highly recommended is the Spicy Fried Clam Rolls and Spider Rolls. (spider rolls, recommended, pos)

Figure 5: Case Study. "Long distance" denotes the distance between aspect and opinion terms is considerably long. "Uncommon aspect/opinion" represents that the term appears in the training set once or less. "Overlapped triplets" denotes there are multiple triplets containing the same aspect or opinion in the sentence.

THANK YOU

Making Better Use of Training Corpus: Retrieval-based Aspect Sentiment Triplet Extraction via Label Interpolation

Guoxin Yu[1,2], Lemao Liu[3*], Haiyun Jiang[3], Shuming Shi[3], Xiang Ao[1,2*]

1 Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China.
2 University of Chinese Academy of Sciences, Beijing 100049, China.
3 Tencent AI Lab, China.