RACV 2021

復旦大學
FUDAN UNIVERSITY

Transformer促进范式迁移

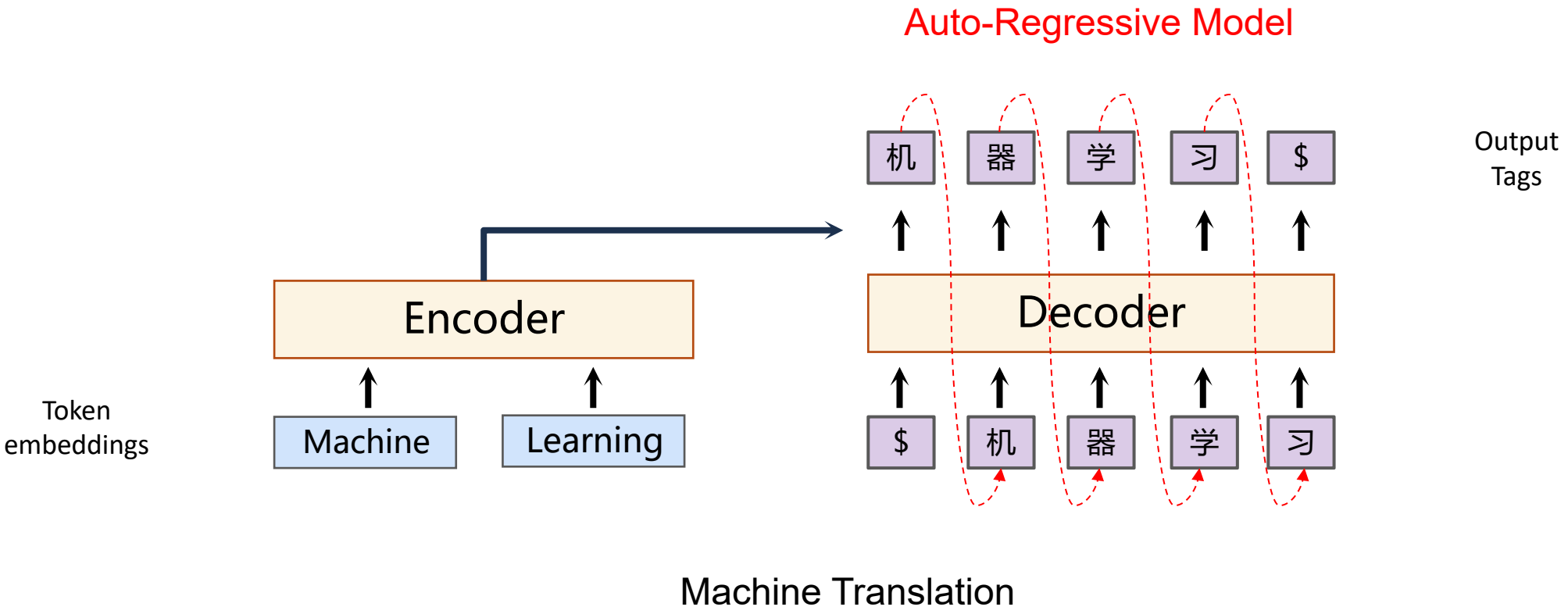
邱锡鹏

复旦大学

https://xpqiu.github.io/

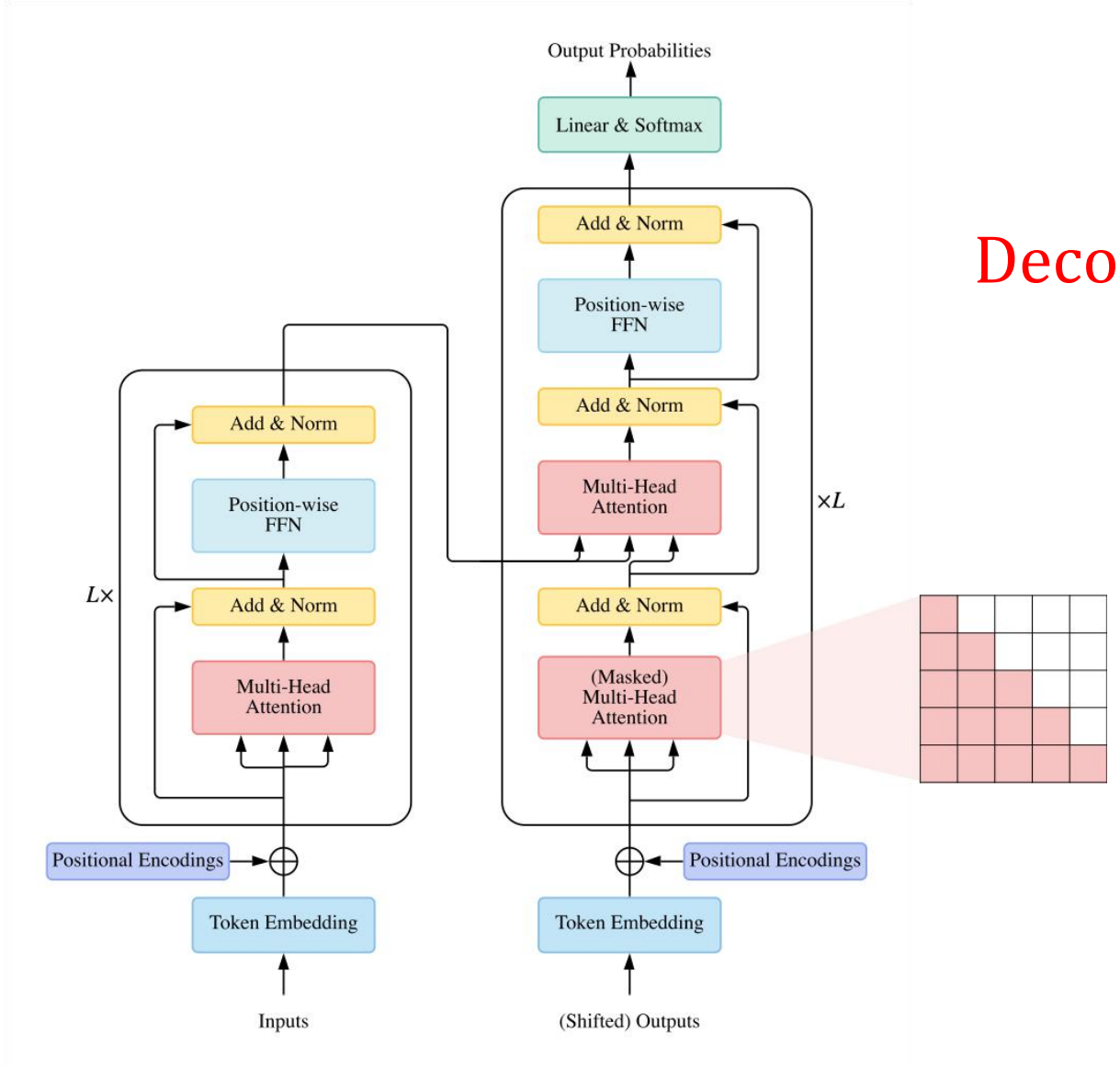
16/10/2021

Sequence-to-Sequence (Seq2Seq) In NLP



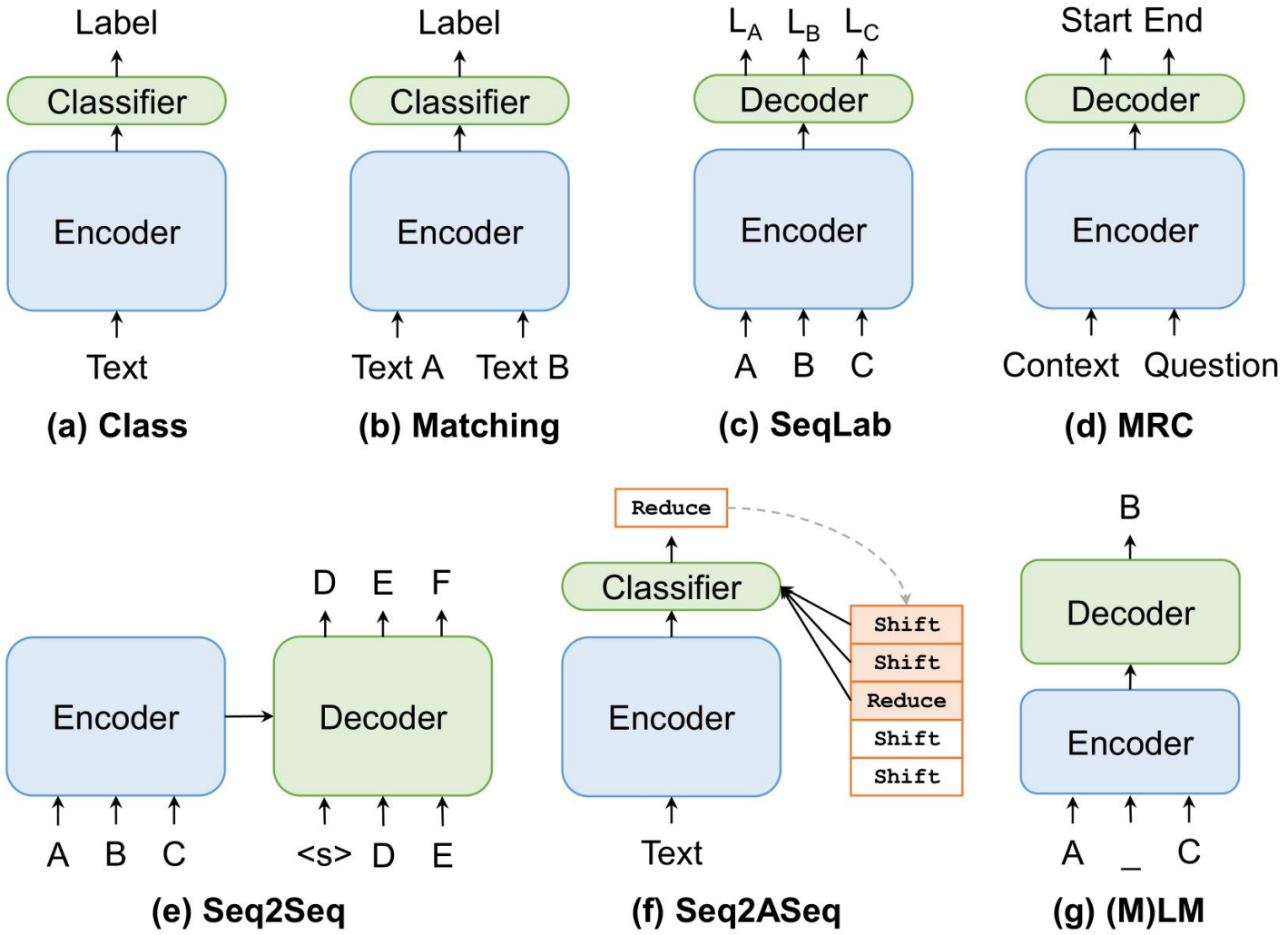
Machine Translation

Transformer

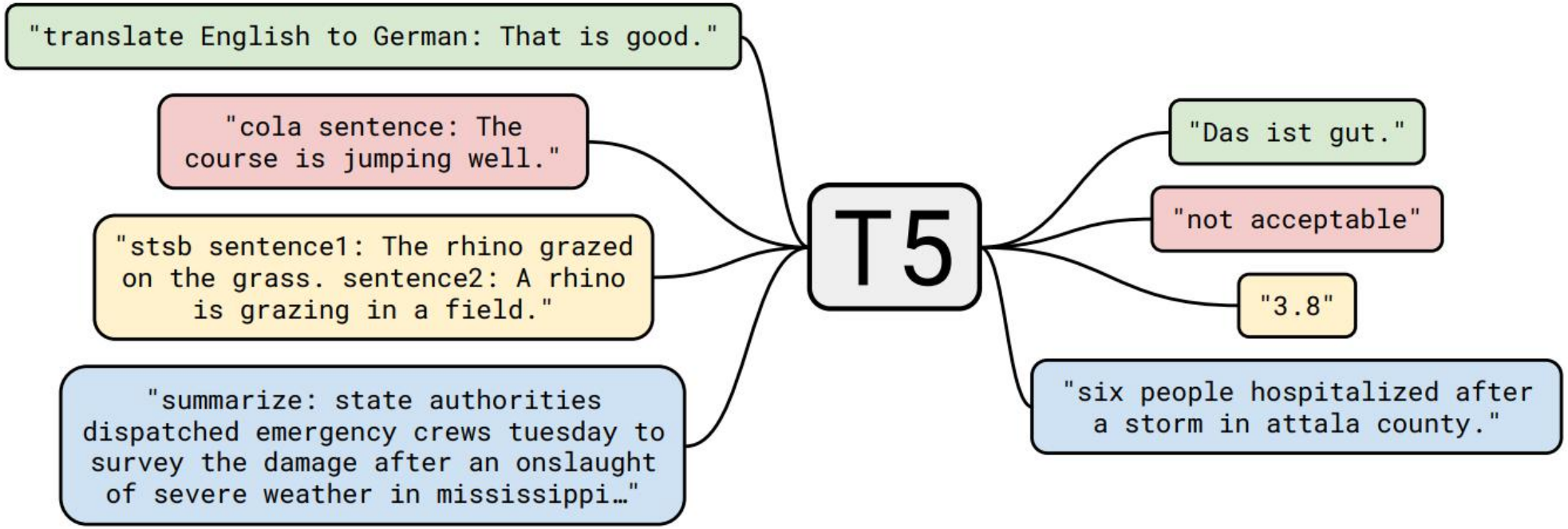


Decoder part is neglected in CV.

Seven Main Paradigms In NLP



Text-to-text framework

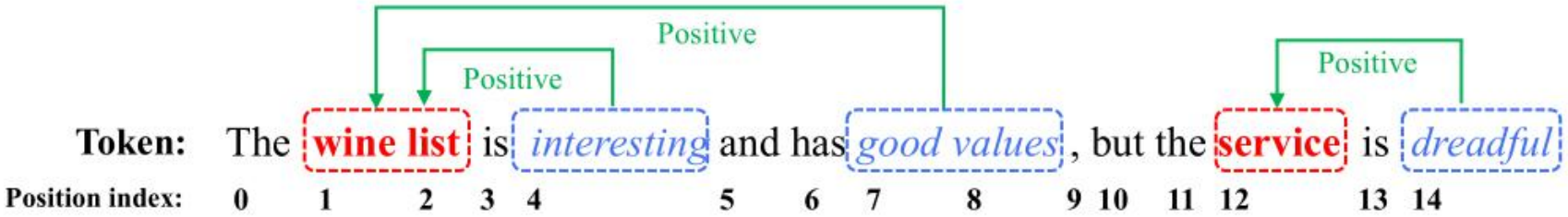


T5：Text-to-Text Transfer Transformer

ABSA

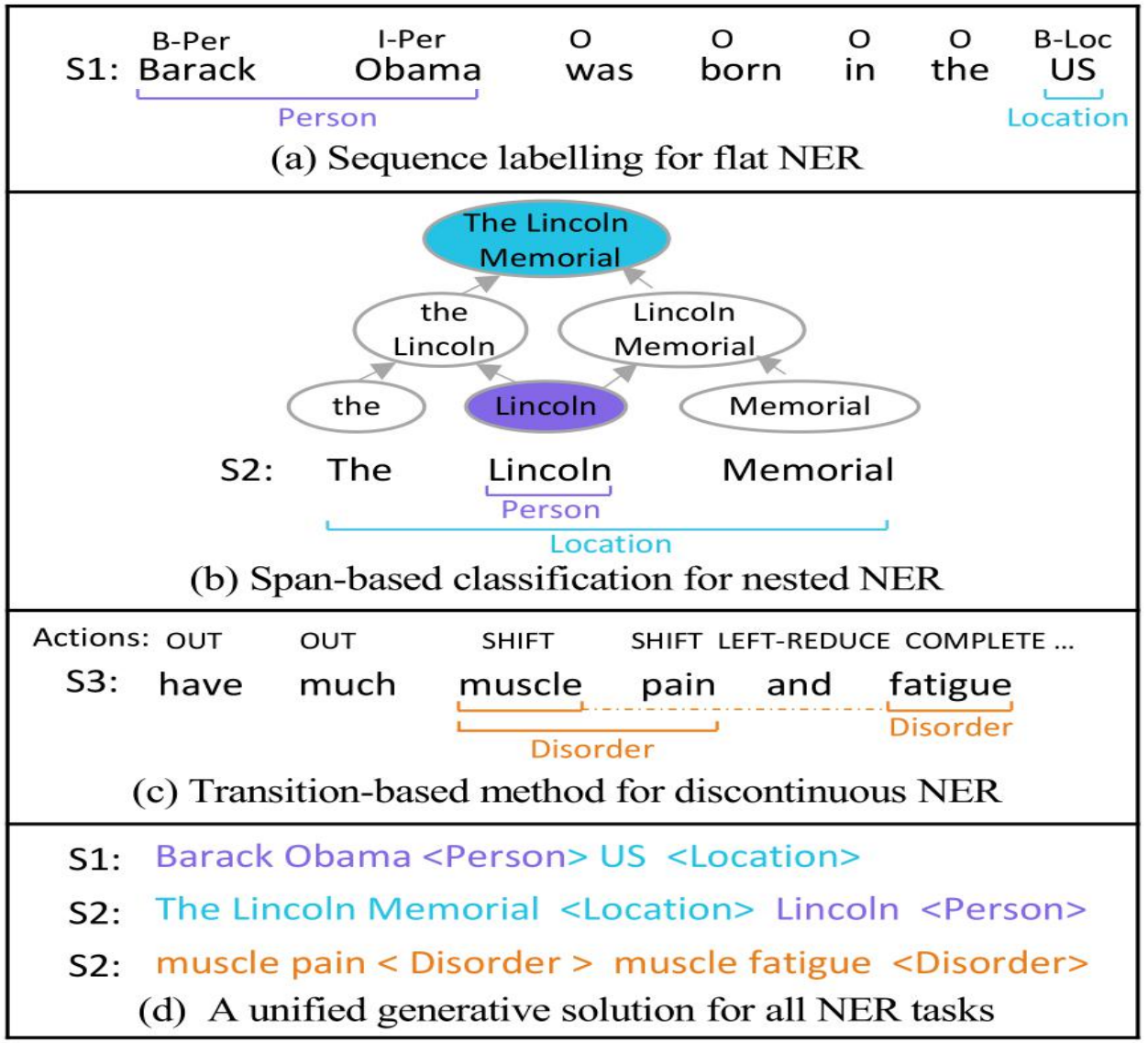
Subtask	Input	Output	Task Type
Aspect Term Extraction(*AE*)	**S**	$\mathbf{a_1}$, $\mathbf{a_2}$	Extraction
Opinion Term Extraction(*OE*)	**S**	o_1, o_2	Extraction
Aspect-level Sentiment Classification(*ALSC*)	**S** + $\mathbf{a_1}$	s_1	Classification
	S + $\mathbf{a_2}$	s_2	
Aspect-oriented Opinion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

ABSA

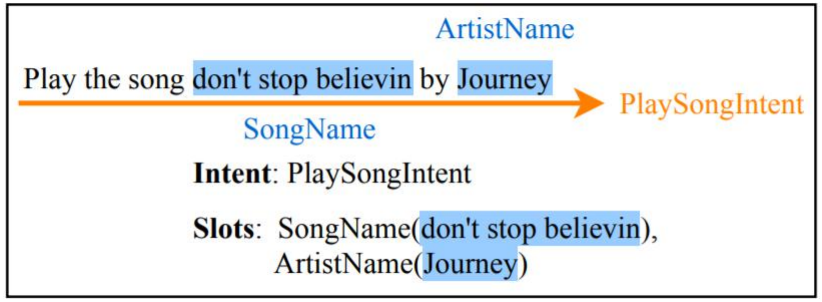


Subtask	Target Sequence
AE	1, 2, 12, 12, </s>
OE	4, 4, 7, 8, 14, 14, </s>
ALSC	1, 2 , POS, </s>
	12, 12, POS, </s>
AOE	1, 2, 4, 4, 7, 8, </s>
	12, 12, 14, 14, </s>
AESC	1, 2, POS, 12, 12, NEG, </s>
Pair.	1, 2, 4, 4, 1, 2, 7, 8, 12, 12, 14, 14, </s>
Triplet.	1, 2, 4, 4, POS, 1, 2, 7, 8, POS, 12, 12, 14,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

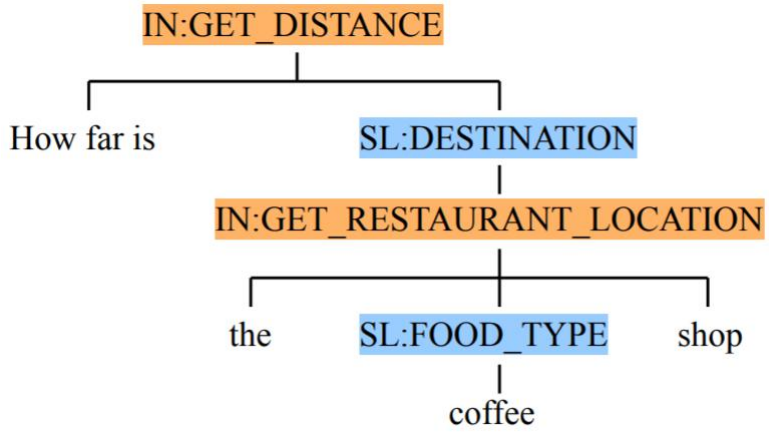
NER


B-Per I-Per O O O O B-Loc
S1: Barack Obama was born in the US
Person
Location
(a) Sequence labelling for flat NER
The Lincoln Memorial
the Lincoln
Lincoln Memorial
the
Lincoln
Memorial
S2: The Lincoln Memorial
Person
Location
(b) Span-based classification for nested NER
Actions: OUT OUT SHIFT SHIFT LEFT-REDUCE COMPLETE ...
S3: have much muscle pain and fatigue
Disorder
Disorder
(c) Transition-based method for discontinuous NER
S1: Barack Obama <Person> US <Location>
S2: The Lincoln Memorial <Location> Lincoln <Person>
S2: muscle pain < Disorder > muscle fatigue <Disorder>
(d) A unified generative solution for all NER tasks

Semantic parsing as Seq2Seq



“simple” query

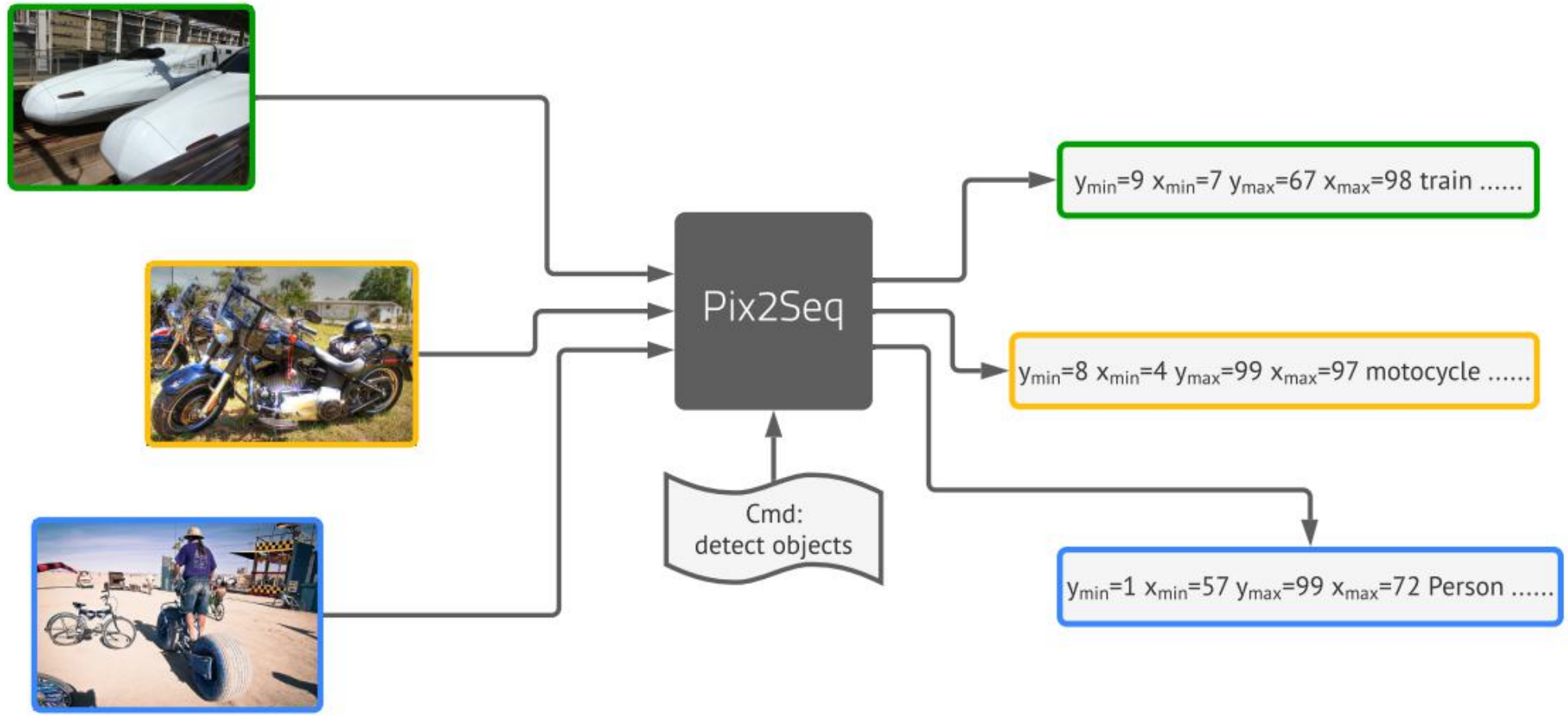


“complex” query

```
Source: play the song don't stop believin by journey
Target: PlaySongIntent SongName( @ptr3 @ptr4 @ptr5
        )SongName ArtistName( @ptr7 )ArtistName
```

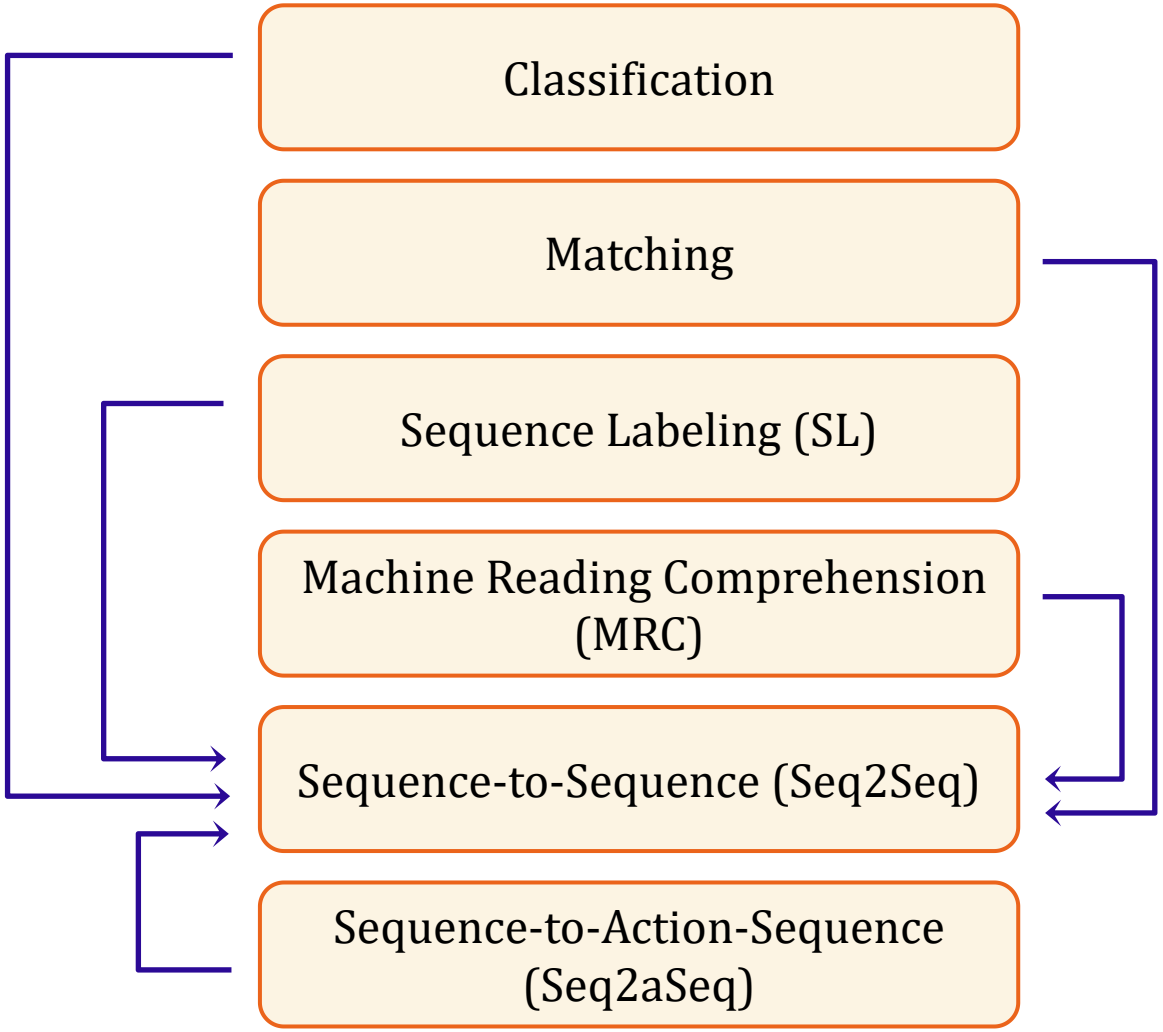
```
Source: How far is the coffee shop
Target: [IN:GET_DISTANCE @ptr0 @ptr1 @ptr2 [SL:DESTINATION
    [IN:GET_RESTAURANT_LOCATION @ptr3 [SL:TYPE_FOOD
    @ptr4 SL:TYPE_FOOD] @ptr5 IN:GET_RESTAURANT_LOCATION]
    SL:DESTINATION] IN:GET_DISTANCE]
```

A recent work in CV



Pix2Seq framework for object detection

Summary



Paradigm Shift brings various NLP tasks a unified framework.
The pretrained backbone models accelerate this shift.

Seq2Seq are a powerful paradigm, which is more general and flexible.