Modeling Information Extraction End-to-end

Hao Fei

Sea AI Lab

Sep 13 , 2022

Content

一、 **Modeling NLP tasks**

二、**Modeling Information Extraction End-to-end**

一、 Modeling NLP tasks

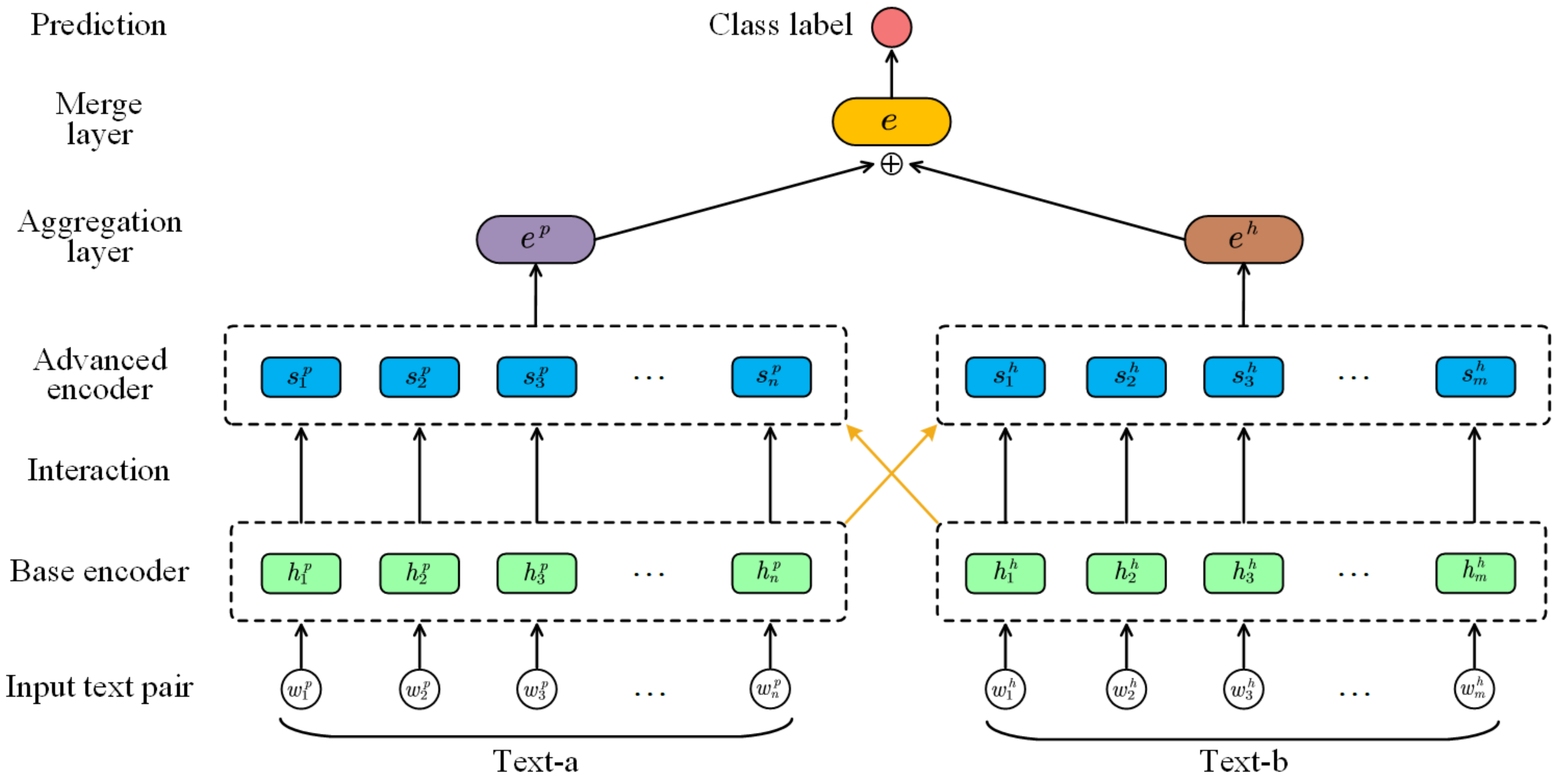
- **Classification**
  - Sentence-level classification ★
  - Sentence-pair classification
  - Span-level classification
  - Token-level classification
    - Input-output Synchronous token-level classification (aka. sequence labeling) ★
    - Input-output Asynchronous token-level classification ★
- **Clustering**
  - Topic modeling
  - …
- **Regression**
  - Multi-label classification
  - …

一、 Modeling NLP tasks

- Sentence-pair classification

☐ Representative NLP tasks：

1. Recognition of Text Entailment (RTE)
2. Natural language inference
3. Paraphrase Identification
…

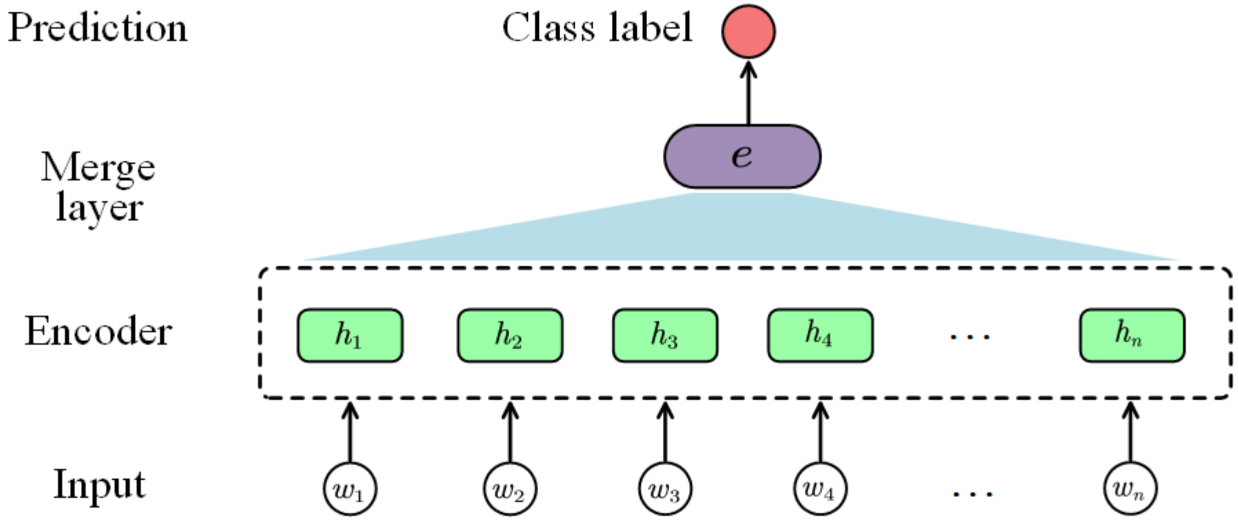


一、 Modeling NLP tasks

➢ Sentence-level classification

☐ Representative NLP tasks：

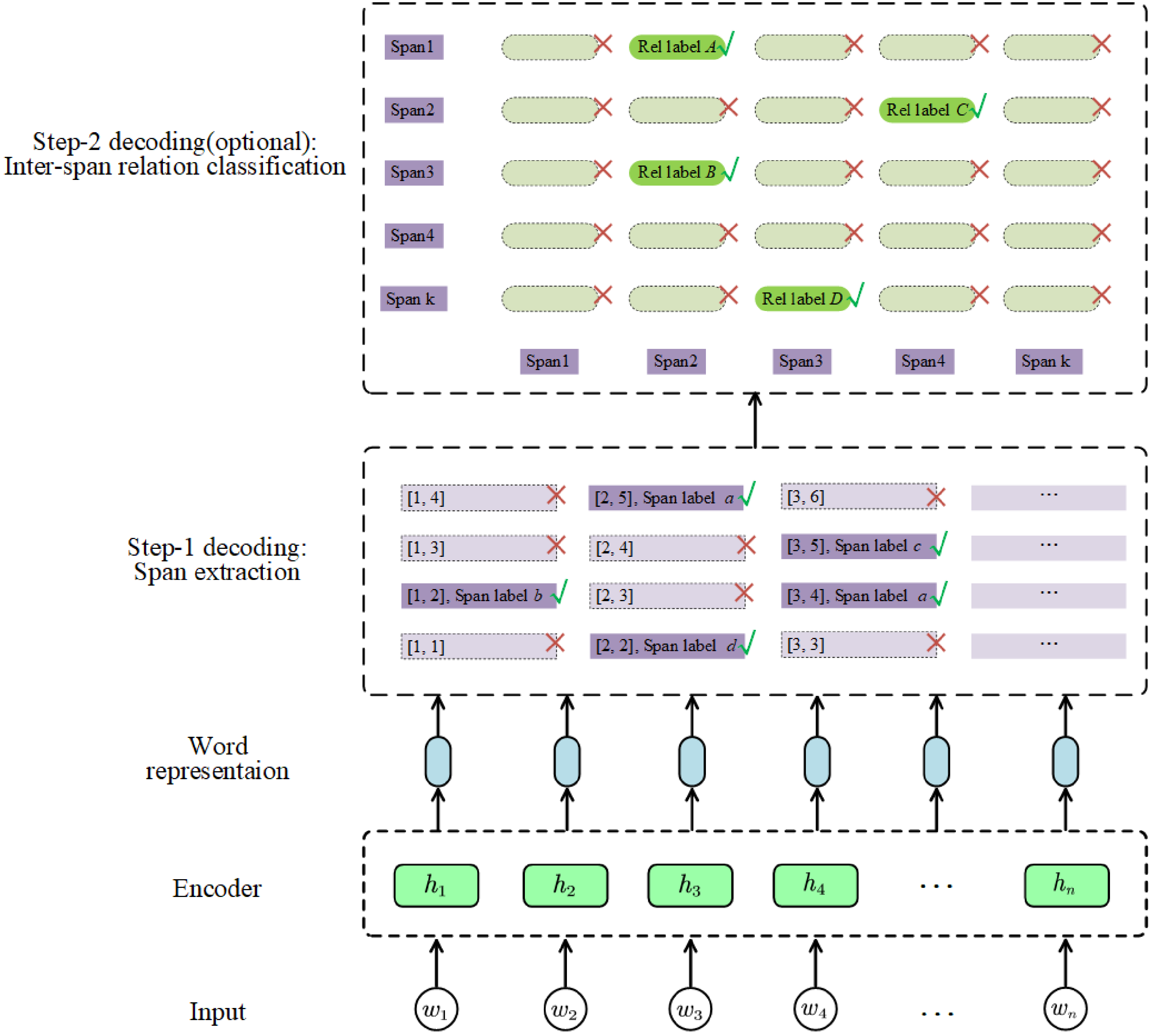
1. *Relation classification*
2. *Topic classification*
3. *Sentiment classification*
4. *Question type classification*
5. *Intention classification*
6. *Emotion classification*
7. *Aggressive language classification*
8. *…*



一、 Modeling NLP tasks

➢ Span-level classification

- ☐ **Step-I：span extraction**
- ☐ **Step-II：span-relation classification**



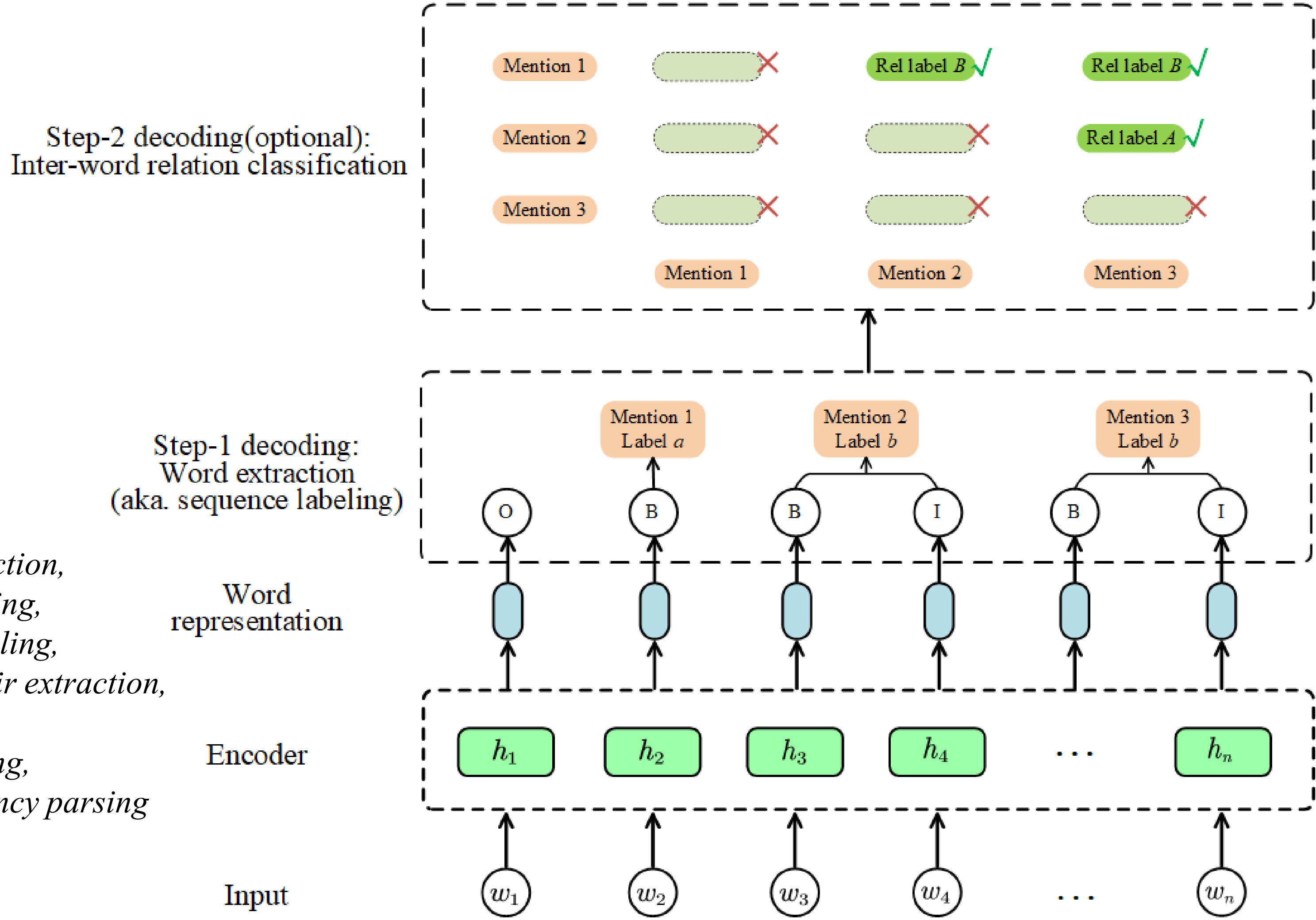
☐ Representative NLP tasks：

1. *Machine reading comprehension*
2. *Extractive automatic summarization*
3. *Nested NER*
4. *Constituency parsing*
5. *Nested RE*
6. *Coreference/anaphora resolution*
7. …

一、 Modeling NLP tasks

- Token-level classification
  - **Input-output Synchronous token-level classification (aka. sequence labeling)**
  - Input-output Asynchronous token-

- ***Step-I：sequence labeling***
- ***Step-II：mention-relation classification***



Representative NLP tasks：

Step 1:

1. *chunk analysis,*
2. *part of speech tagging,*
3. *named entity recognition,*
4. *Chinese word segmentation,*
5. *fine-grained emotion analysis,*
6. *stance extraction,*
7. *autoregressive language modeling,*
8. *...*

Step 2:

1. *relationship extraction,*
2. *opinion role labeling,*
3. *semantic role labeling,*
4. *opinion-aspect pair extraction,*
5. *event extraction,*
6. *dependency parsing,*
7. *semantic dependency parsing*
8. *...*

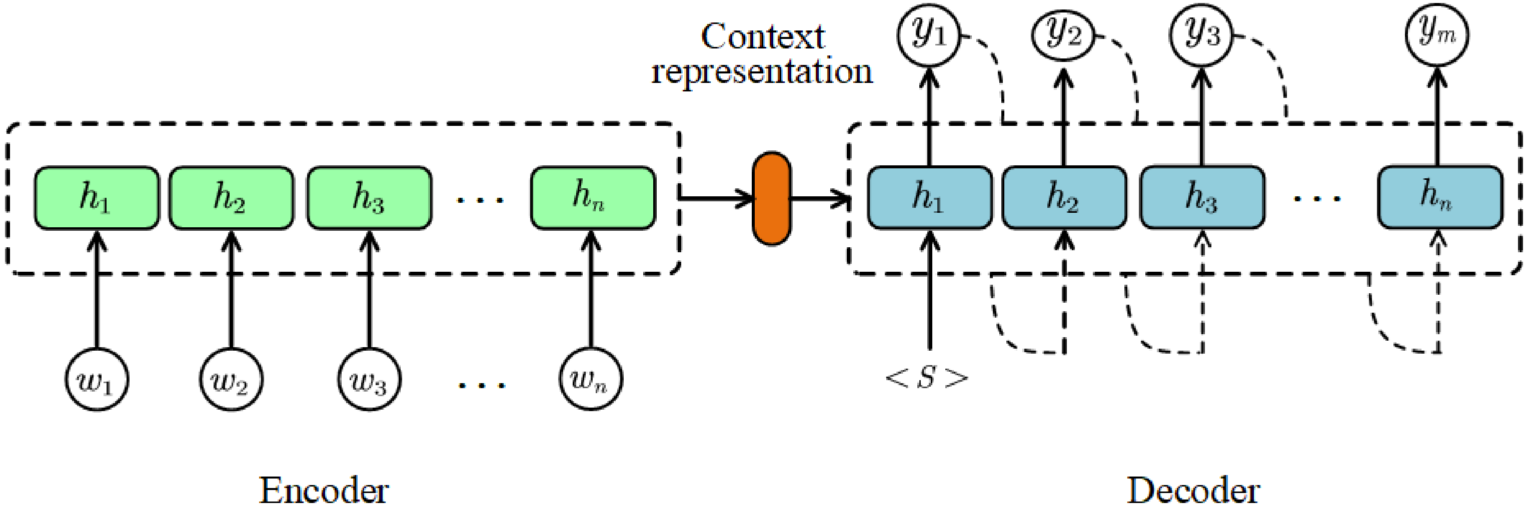
一、 Modeling NLP tasks

- Token-level classification
  - Input-output Synchronous token-level classification (aka. sequence labeling)
  - Input-output Asynchronous token-level classification

    Aka.
    - *Sequence-to-Sequence framework*
    - *Encoder-Decoder framework*
    - *End-to-end framework*

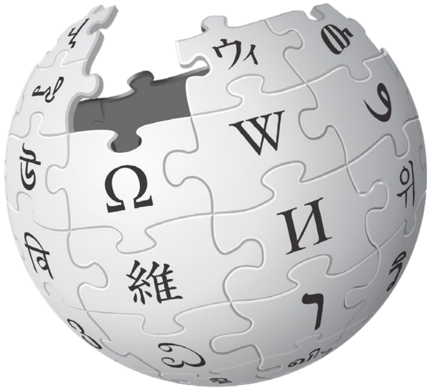
☐ Representative NLP tasks：

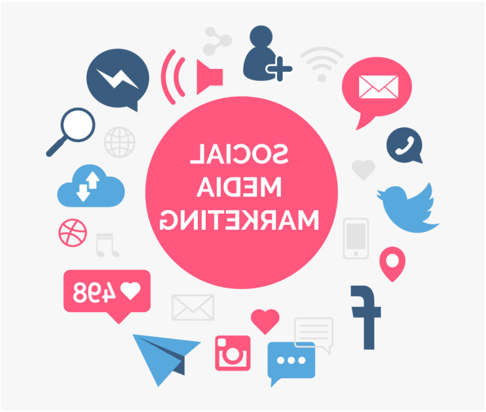
1. *neural machine translation,*
2. *automatic summarization,*
3. *dialogue system,*
4. *autoregressive language modeling,*
5. *machine reading comprehension,*
6. …



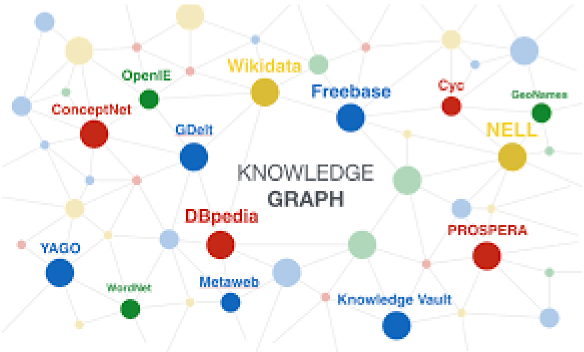
二、Modeling Information Extraction End-to-end

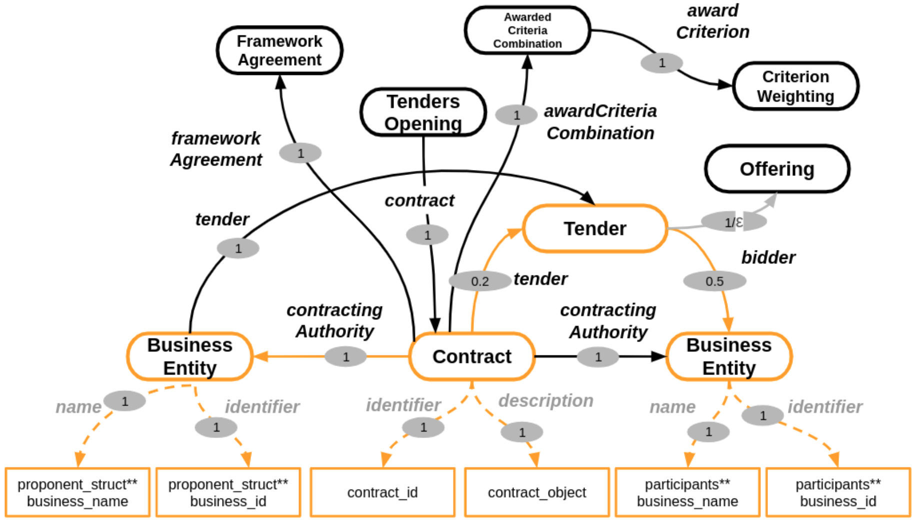
Plain Text



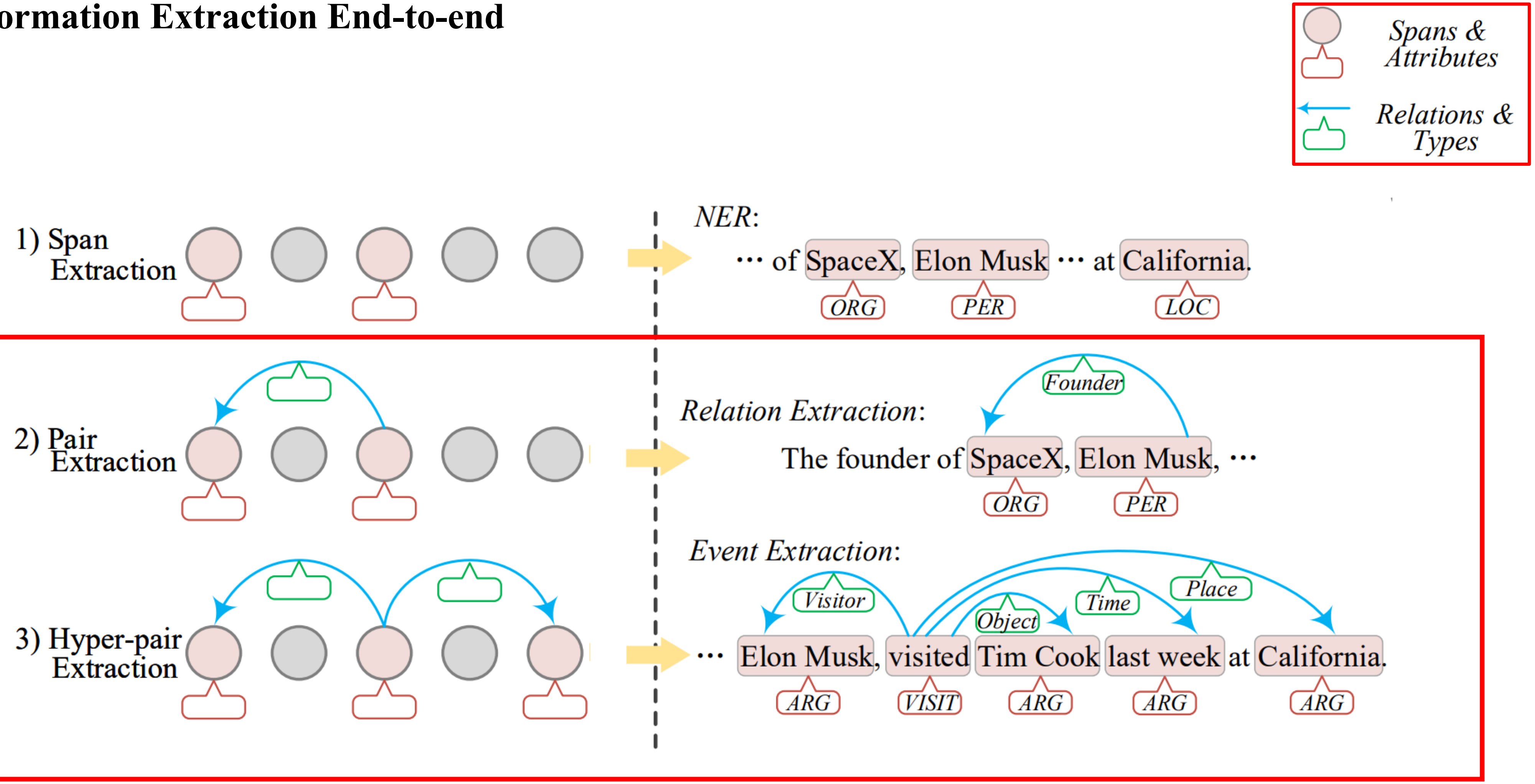


Structured Knowledge/Representation





二、Modeling Information Extraction End-to-end



(a) Information extraction task prototypes

(b) Representative task examples

Structural learning

二、Modeling Information Extraction End-to-end

◆ Structure Prediction

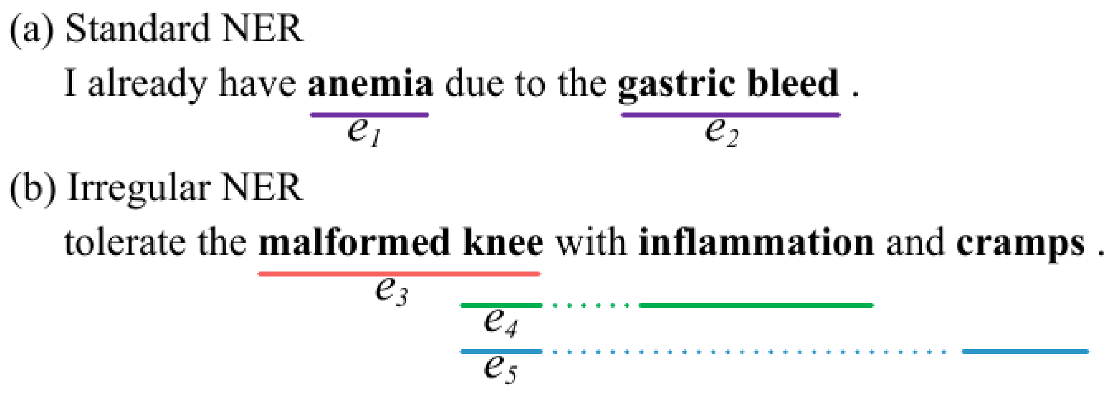
- Well-defined, simple tasks:
  - *dependency parsing*
  - *constituency parsing*
  - *relation extraction (RE)*
  - *aspect-opinion pair extraction (AOP)*
  - *emotion-cause pair extraction (ECPE)*
  - *aspect-based sentiment triplet extraction (ASTE)*
  - *aspect-based sentiment quadruple extraction (ASQE)*
  - *event extraction (EE)*
  - *semantic role labeling (SRL)*
  - *opinion role labeling (ORL)*

- Complex tasks:
  - *overlapped NER/RE/EE*
  - *nested NER*
  - *discontinuous NER*

  - *combinatory categorial grammar (CCG)*
  - *semantic dependency parsing*
  - *broad-coverage semantic parsing*
  - *meaning representation parsing (MRP)*

二、Modeling Information Extraction End-to-end

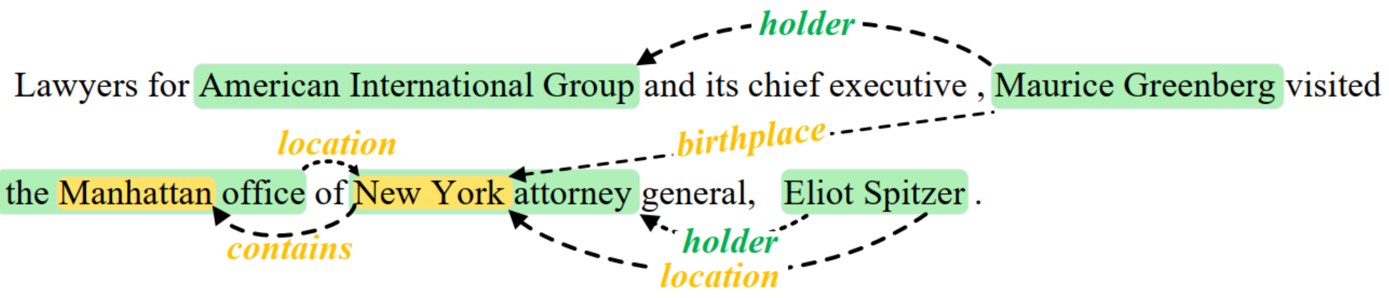
◆ Overlapped NER/RE/EE



- e_1 and e_2 are regular entity mentions.
- e_3 overlaps with two discontinuous mentions e_4 and e_5 at the token ***knee***.

二、Modeling Information Extraction End-to-end

◆ Overlapped NER/RE/EE



- The entity ‘Manhattan’ nests with ‘the Manhattan office’, and the mention ‘New York’ nests with ‘New York attorney’.
- The triplets (Eliot Spitzer, holder, New York attorney) and (Eliot Spitzer, location, New York) overlap on entity ‘Eliot Spitzer’.

二、Modeling Information Extraction End-to-end

◆ Overlapped NER/RE/EE

■ Normal Triplet

two flat entities exist standalone, forming a relational triplet with no other triplet overlapping.

■ Single Triplet Overlap

- Single Normal-Entity Triplet

  one of the entities co-exists in other triplet(s).

  the triplet *(Trump, PresidentOf, United States)* shares the entity '*Trump*' with *(Trump, LiveIn, WhiteHouse)*

- Single Overlaping-Entity Triplet

  Based on Single Normal-Entity Triplet, the shared entities is further nested with the one in other triplet.

  In triplet (Donald Trump, PresidentOf, the United States), the entity 'Donald Trump' nests with the one 'Trump' in (Trump, LiveIn, WhiteHouse)

■ Pair Triplet Overlap

- Pair Normal-Entity Triplet

  both of the two entities in one triplet co-exists in other triplet(s).

  (Trump, PresidentOf, United States) collides with the another one (Trump, Governance, United States)

- Pair Overlaping-Entity Triplet

  both two entities are overlapped anywhere else, further with at least one entity nesting with other.

  In (Trump, PresidentOf, United States) and (Donald Trump, Governance, United States), the entity 'Trump' nests with 'Donald Trump'.

二、Modeling Information Extraction End-to-end

◆ Traditional solution: Pipeline handling

➢ **Step1**: token/span/clique extracting

text → token/span/clique (w/ type)

➢ **Step1**: relation detecting/grouping

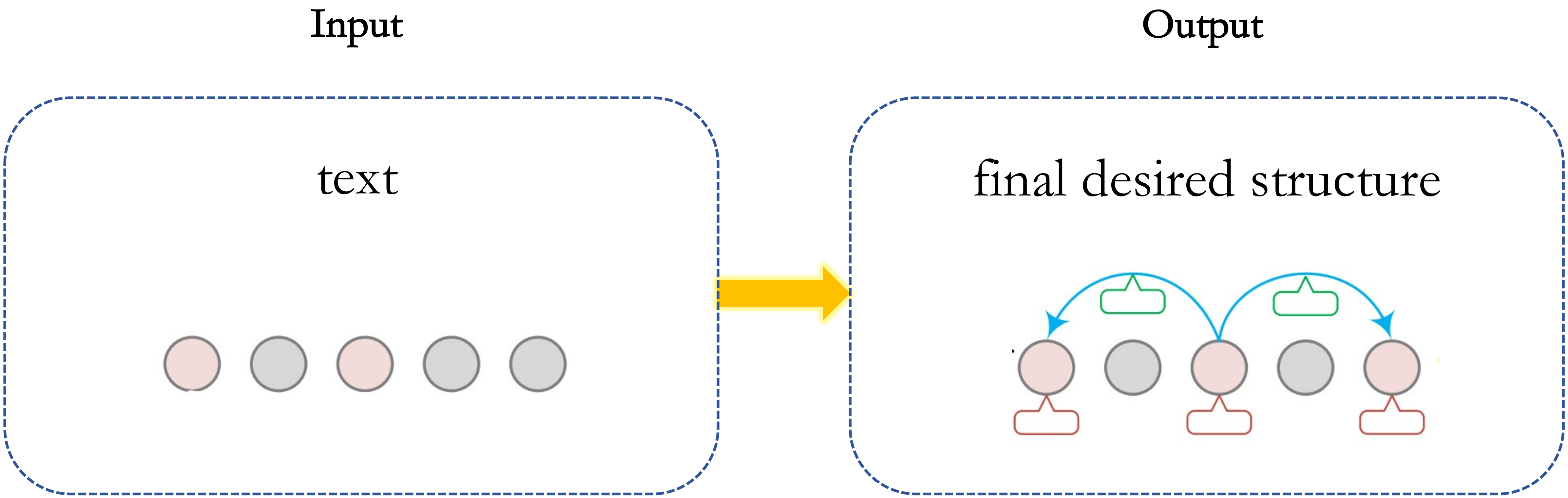
text + token/span/clique → inter-relation (w/ type)

➢ **Step3** (post-processing): formatting into final desired structure

token/span/clique
inter-relation
→ final structure

二、Modeling Information Extraction End-to-end

- Recent solution: End-to-end handling



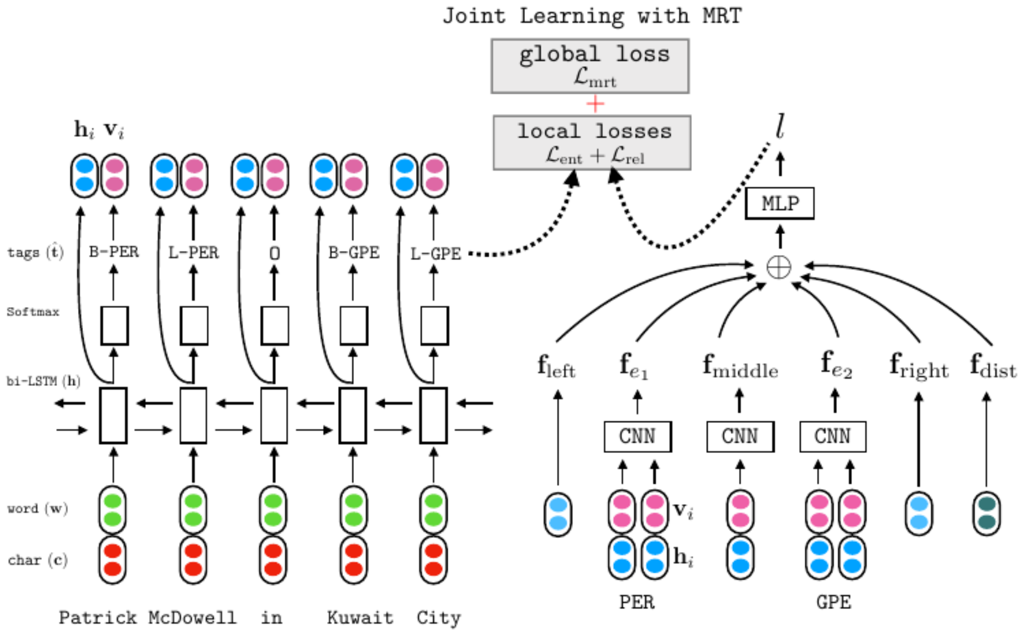
二、Modeling Information Extraction End-to-end

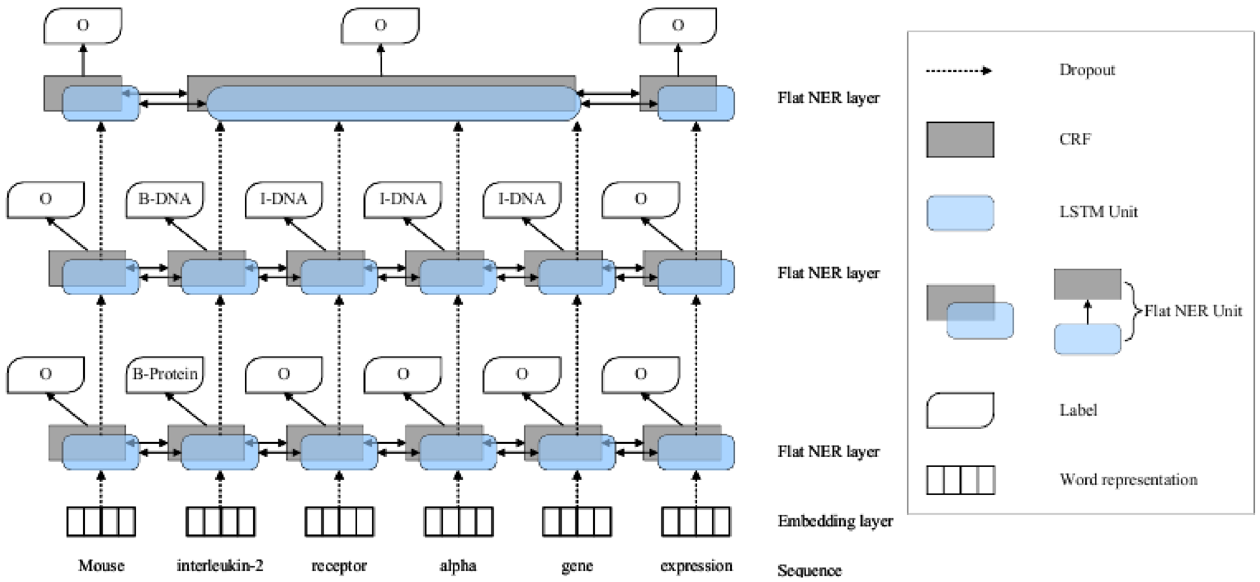
◆ End-to-end modeling

➤ Type-A: Parameter Sharing

➤ Multi-task learning

➤ Stacking layer framework





- *A Neural Layered Model for Nested Named Entity Recognition. NAACL-HLT 2018: 1446-1459*
- *Extracting Entities and Relations with Joint Minimum Risk Training. EMNLP 2018: 2256-2265*

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Type-B: Joint Decoding

- Transition model
- Span-graph model
- Hypergraph model
- Table-filling/Grid-tagging model
- Seq2seq (encoder-decoder) model
- Transforming into MRC-QA
- ……

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

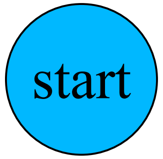
- ✓ *The process of a finite state automata.*
- ✓ *Transition process from initial state to terminal state.*
- ✓ *The transition framework consists of two core elements:* ***Action*** *and* ***State****.*

- *Yue Zhang and Stephen Clark. 2011. Syntactic Processing Using the Generalized Perceptron and Beam Search. In Computational Linguistics, 37(1), March.*

二、Modeling Information Extraction End-to-end

- ◆ End-to-end modeling
  - ➢ Transition model

Automata



- ***State***

  Corresponds to partial results during decoding

  start state, end state, S_i

- ***Action***

  The operations that can be applied for state transition

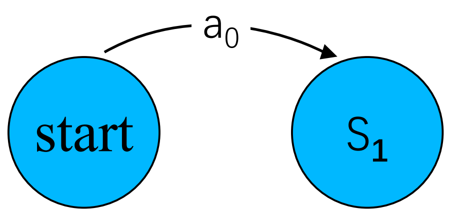
  Construct output incrementally

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

Automata



- ***State***

  Corresponds to partial results during decoding

  start state, end state, S_i

- ***Action***

  The operations that can be applied for state transition

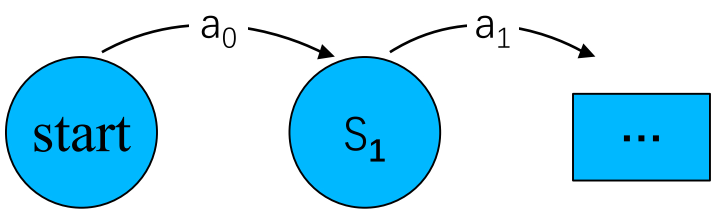
  Construct output incrementally

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

Automata



- ***State***

  Corresponds to partial results during decoding

  start state, end state, S_i

- ***Action***

  The operations that can be applied for state transition

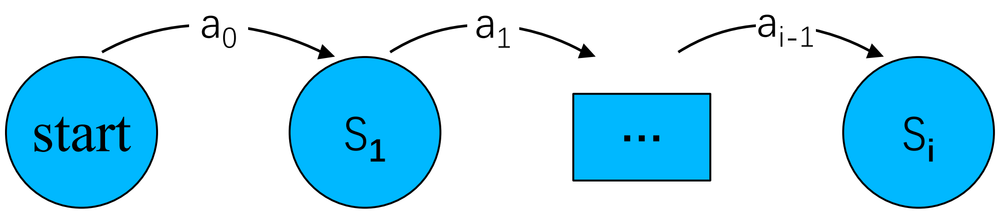
  Construct output incrementally

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

Automata



- ***State***

  Corresponds to partial results during decoding

  start state, end state, S_i

- ***Action***

  The operations that can be applied for state transition

  Construct output incrementally

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

- ***State***

  Corresponds to partial results during decoding

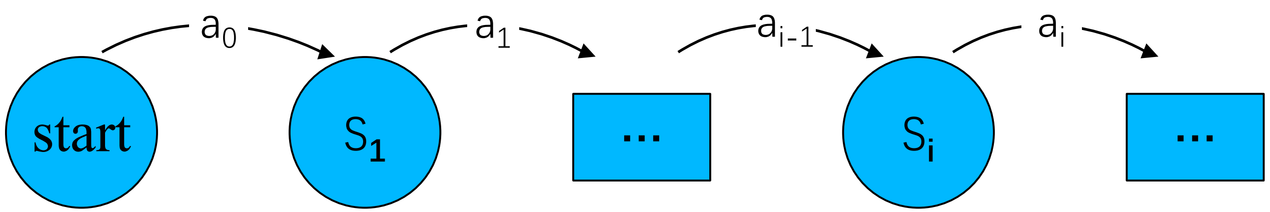
  start state, end state, S_i

- ***Action***

  The operations that can be applied for state transition

  Construct output incrementally

Automata



二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

- *State*

  Corresponds to partial results during decoding

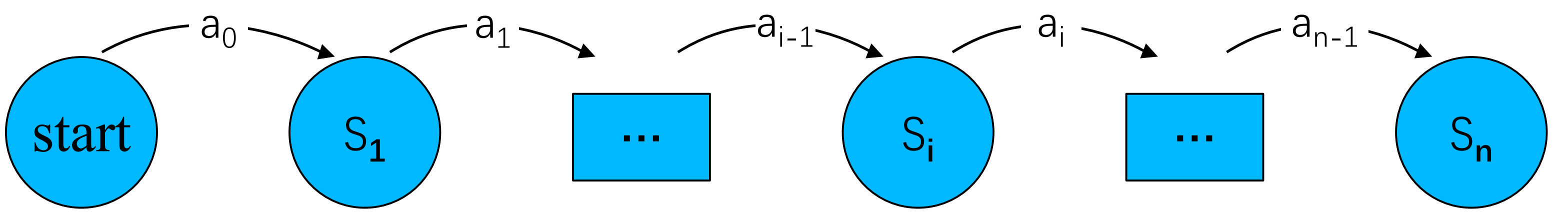
  start state, end state, S_i

- *Action*

  The operations that can be applied for state transition

  Construct output incrementally

Automata



二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

- ***State***

  Corresponds to partial results during decoding

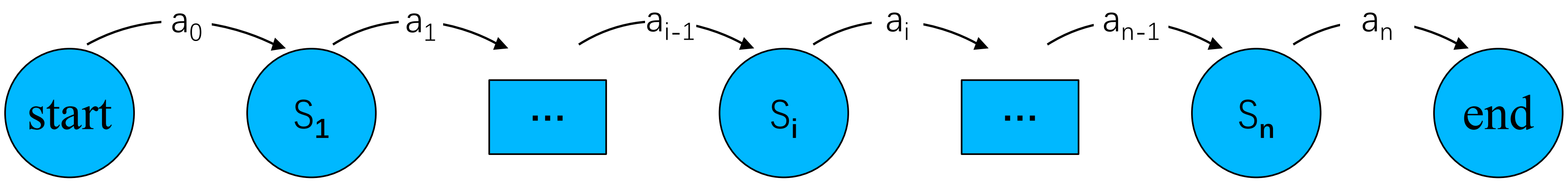
  start state, end state, S_i

- ***Action***

  The operations that can be applied for state transition

  Construct output incrementally

Automata



二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

He does it here

➢ An Example

- S-SHIFT
- R-REDUCE
- AL-ARC-LEFT
- AR-ARC-RIGHT

二、Modeling Information Extraction End-to-end

- End-to-end modeling
  - Transition model

He does it here $\xrightarrow{S}$ He does it here

- An Example
  - S-SHIFT
  - R-REDUCE
  - AL-ARC-LEFT
  - AR-ARC-RIGHT

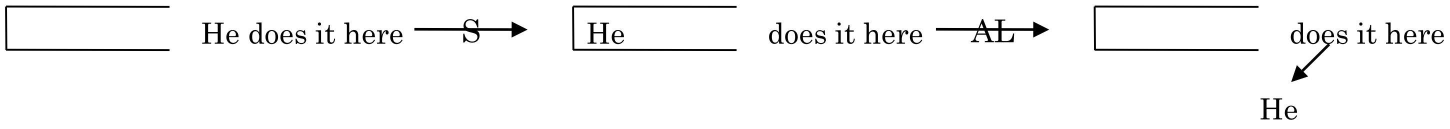
二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

➢ An Example

- S-SHIFT
- R-REDUCE
- AL-ARC-LEFT
- AR-ARC-RIGHT

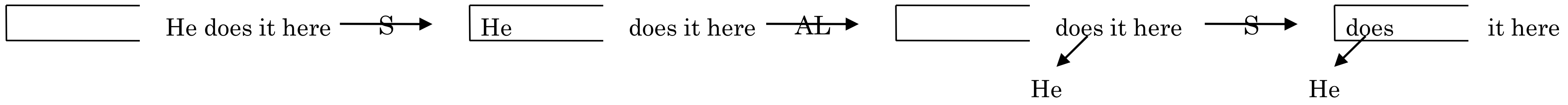


二、Modeling Information Extraction End-to-end

- ◆ End-to-end modeling
  - ➢ Transition model

➢ An Example

- S-SHIFT
- R-REDUCE
- AL-ARC-LEFT
- AR-ARC-RIGHT



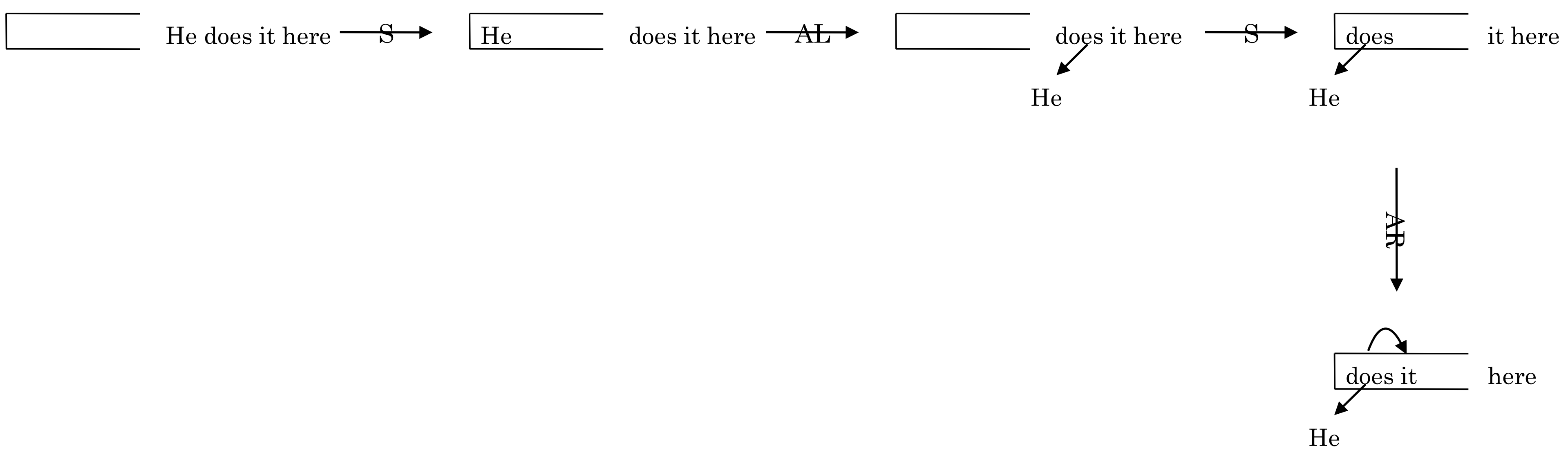
二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

➢ An Example

- S-SHIFT
- R-REDUCE
- AL-ARC-LEFT
- AR-ARC-RIGHT



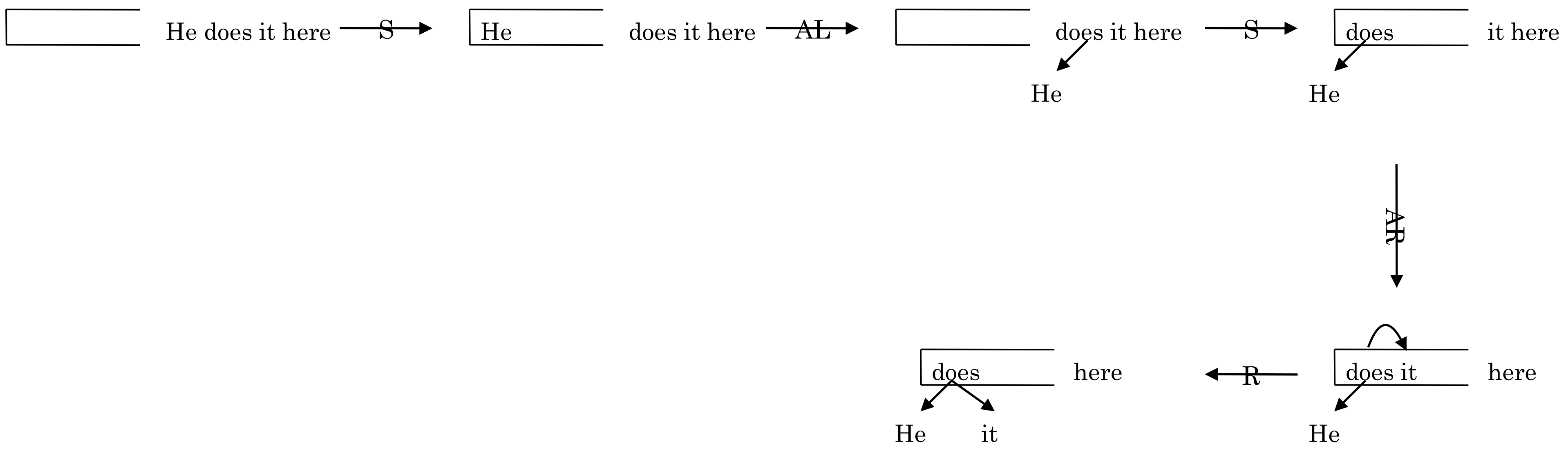
二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

➢ An Example

- S-SHIFT
- R-REDUCE
- AL-ARC-LEFT
- AR-ARC-RIGHT



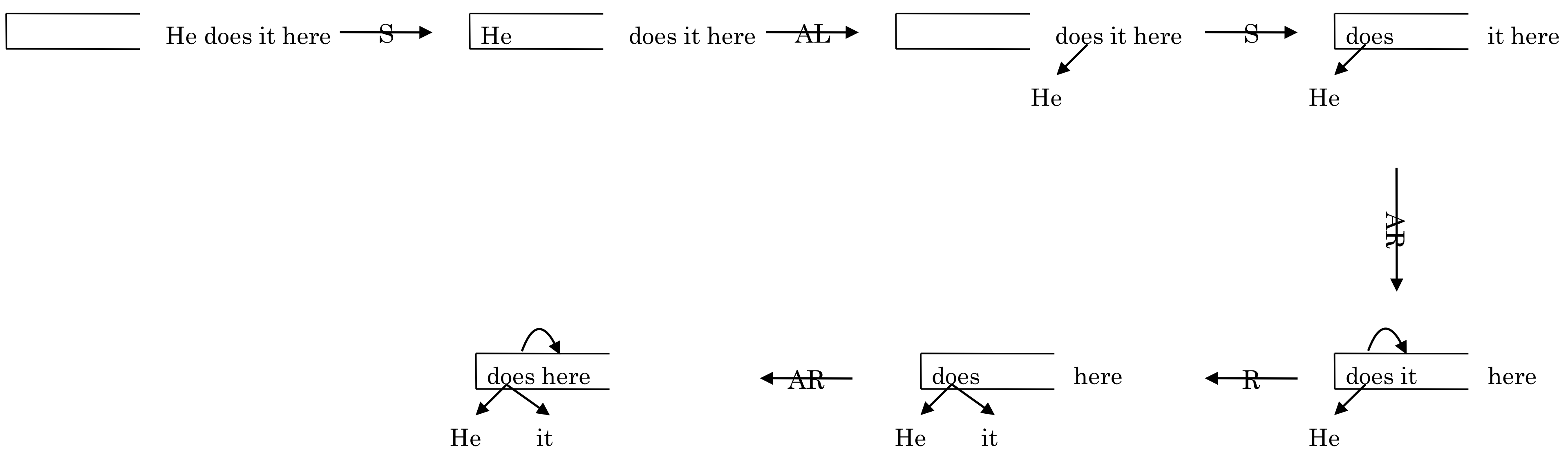
二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

➢ An Example

- S-SHIFT
- R-REDUCE
- AL-ARC-LEFT
- AR-ARC-RIGHT



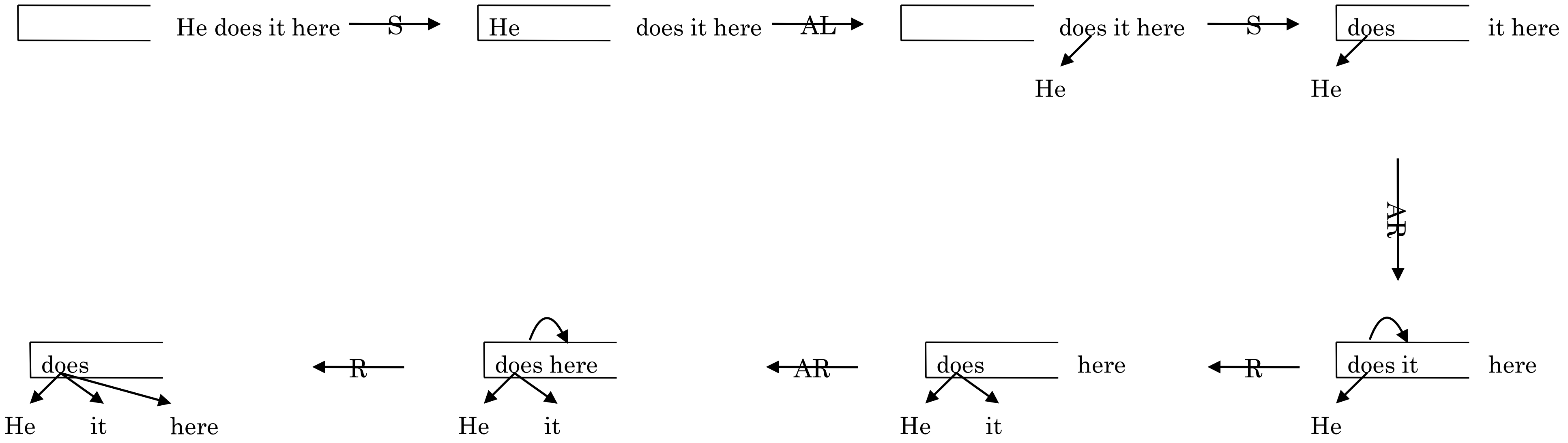
二、Modeling Information Extraction End-to-end

- ◆ End-to-end modeling

  - ➢ Transition model

➢ An Example

- S-SHIFT
- R-REDUCE
- AL-ARC-LEFT
- AR-ARC-RIGHT



二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transition model

Sentence : He$_1$ says$_2$ the$_3$ agency$_4$ seriously$_5$ needs$_6$ money$_7$ to$_8$ develop$_9$

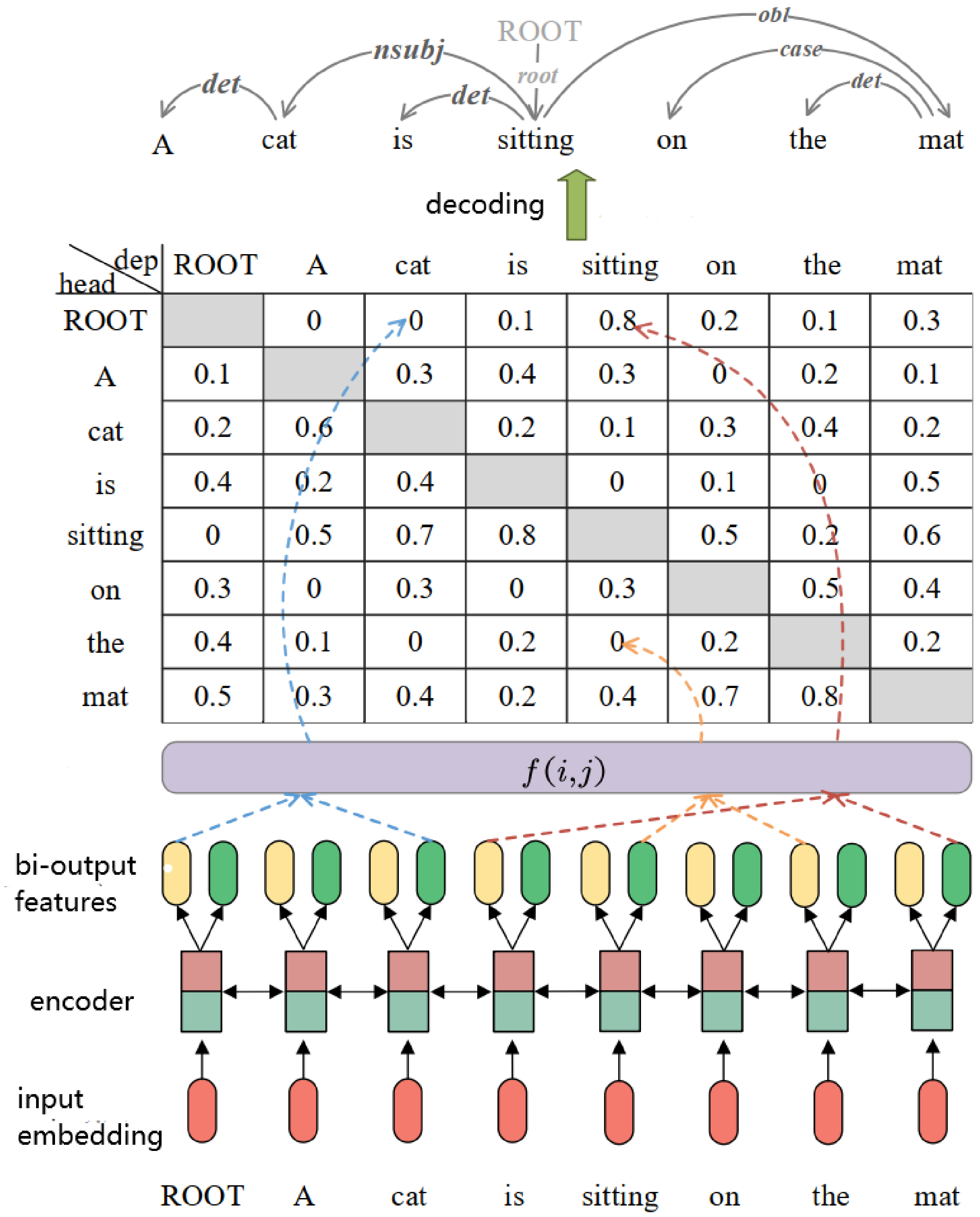
Step	Action	σ^o	α^o	λ	σ^r	α^r	β	Ptr	Y
0	-	[]	[]	Null	[]	[]	[1,···,9]		
1	R-START	[]	[]	$(1,1)^r$	[]	[]	[1,···,9]	[1,···,9]	
2	SHIFT	[]	[]	Null	[(1,1)]	[]	[2,···,9]		
3	O-START	[]	[]	$(2,2)^o$	[(1,1)]	[]	[2,···,9]	[2,···,9]	
4	ARC	[]	[]	$(2,2)^o$	[]	[(1,1)]	[2,···,9]		$Y \cup \{\langle(2,2)^o,(1,1)^r(hd)\rangle\}$
5	SHIFT	[(2,2)]	[]	Null	[(1,1)]	[]	[3,···,9]		
6	R-START	[(2,2)]	[]	$(3,4)^r$	[(1,1)]	[]	[3,···,9]	[3,4,···,9]	
7	ARC	[]	[(2,2)]	$(3,4)^r$	[(1,1)]	[]	[3,···,9]		$Y \cup \{\langle(2,2)^o,(3,4)^r(tg)\rangle\}$
8	SHIFT	[(2,2)]	[]	Null	[(1,1),(3,4)]	[]	[4,···,9]		
9	NO-START	[(2,2)]	[]	Null	[(1,1),(3,4)]	[]	[5,···,9]		
10	O-START	[(2,2)]	[]	$(5,6)^o$	[(1,1),(3,4)]	[]	[5,···,9]	[5,6,···,9]	
11	ARC	[(2,2)]	[]	$(5,6)^o$	[(1,1)]	[(3,4]	[5,···,9]		$Y \cup \{\langle(5,6)^o,(3,4)^r(hd)\rangle\}$
12	NO-ARC	[(2,2)]	[]	$(5,6)^o$	[]	[(1,1),(3,4]	[5,···,9]		
13	SHIFT	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4]	[]	[6,···,9]		
14	NO-START	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4]	[]	[7,8,9]		
15	R-START	[(2,2),(5,6)]	[]	$(7,9)^r$	[(1,1),(3,4]	[]	[7,8,9]	[7,8,9]	
16	ARC	[(2,2)]	[(5,6)]	$(7,9)^r$	[(1,1),(3,4]	[]	[7,8,9]		$Y \cup \{\langle(5,6)^o,(7,9)^r(tg)\rangle\}$
17	NO-ARC	[]	[(2,2),(5,6)]	$(7,9)^r$	[(1,1),(3,4]	[]	[7,8,9]		
18	SHIFT	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4,(7,9)]	[]	[8,9]		
19	NO-START	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4,((7,9)]	[]	[9]		
20	NO-START	[(2,2),(5,6)]	[]	Null	[(1,1),(3,4,(7,9)]	[]	[]		

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Span-graph model

- *Parsing task is regarded as the process of building a tree.*
- *Searching a weighted graph to find the subgraph with the highest score.*

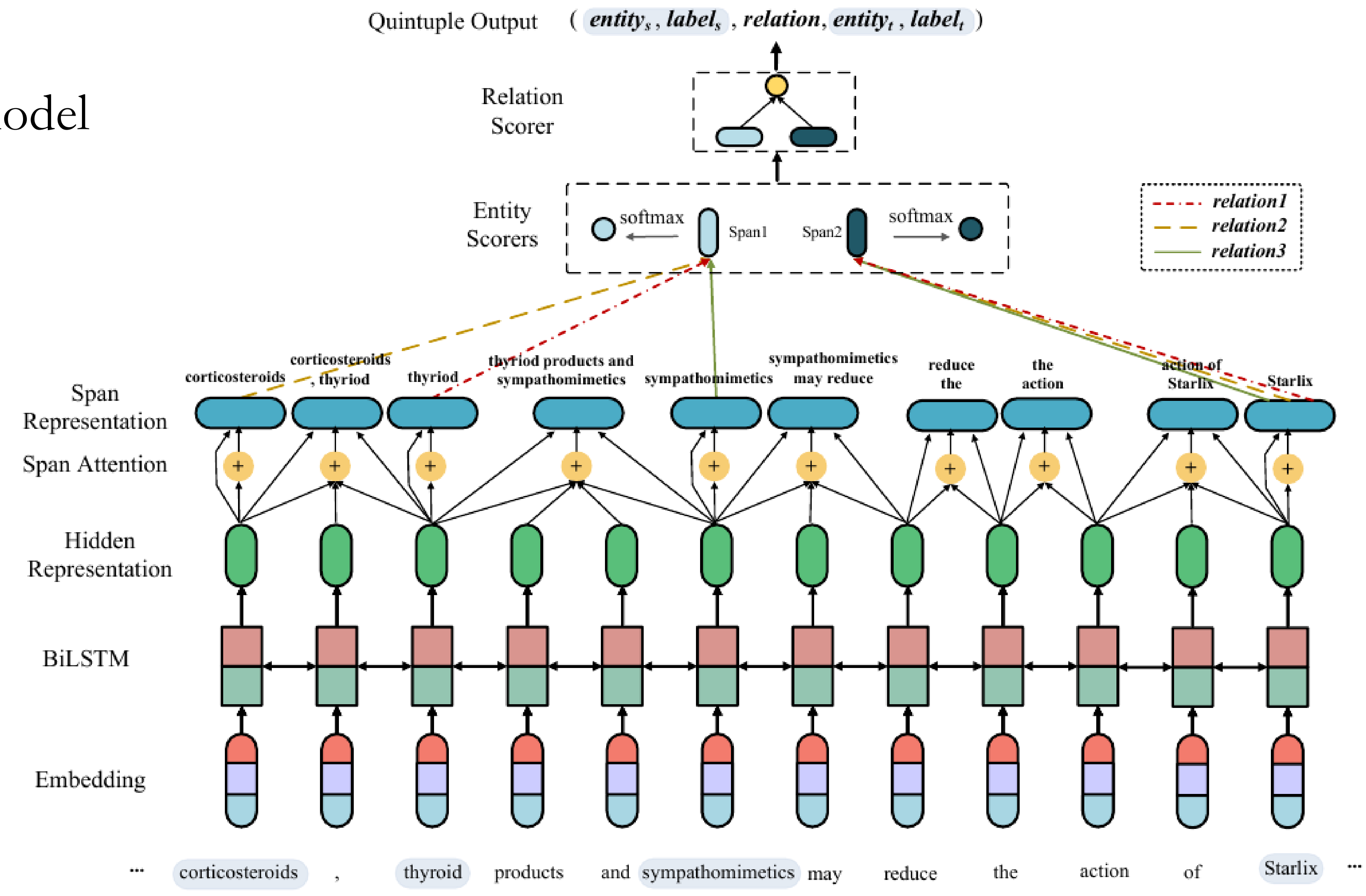


head \ dep	ROOT	A	cat	is	sitting	on	the	mat
ROOT		0	0	0.1	0.8	0.2	0.1	0.3
A	0.1		0.3	0.4	0.3	0	0.2	0.1
cat	0.2	0.6		0.2	0.1	0.3	0.4	0.2
is	0.4	0.2	0.4		0	0.1	0	0.5
sitting	0	0.5	0.7	0.8		0.5	0.2	0.6
on	0.3	0	0.3	0	0.3		0.5	0.4
the	0.4	0.1	0	0.2	0	0.2		0.2
mat	0.5	0.3	0.4	0.2	0.4	0.7	0.8	

- *Timothy Dozat, Christopher D. Manning. 2017.*
  Deep Biaffine Attention for Neural Dependency Parsing. ICLR.

二、Modeling Information Extraction End-to-end

- ◆ End-to-end modeling
  - ➢ Span-graph model

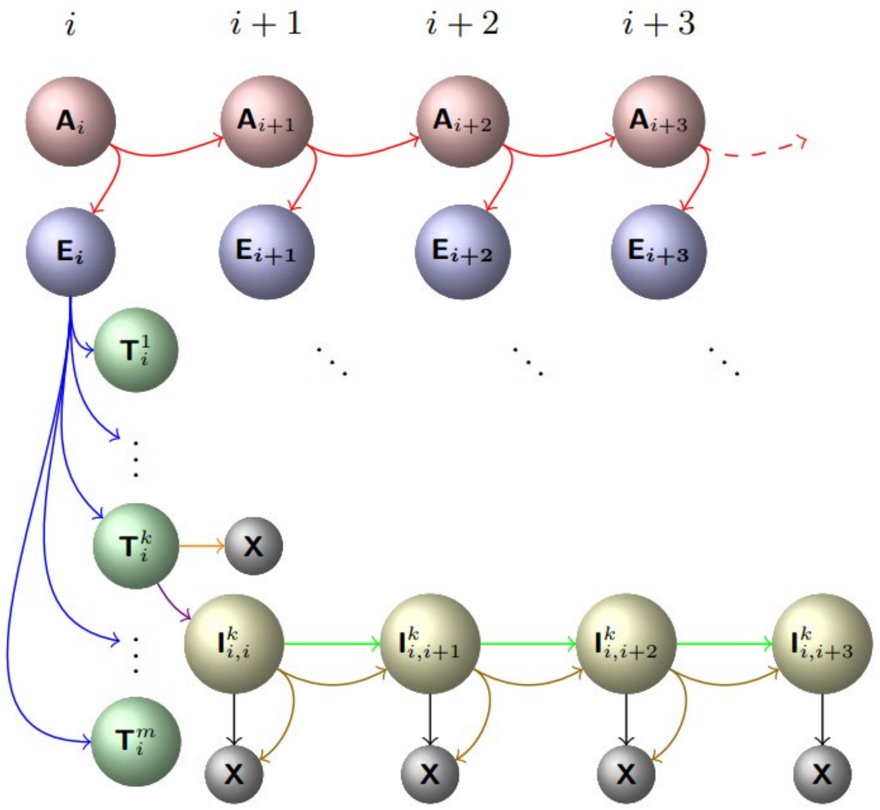


- *Span-Based Joint Entity and Relation Extraction with Transformer Pre-Training. ECAI 2020: 2006-2013*

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Hypergraph model



✓ Standard graph

an edge only connects two vertices.

✓ Hypergraph

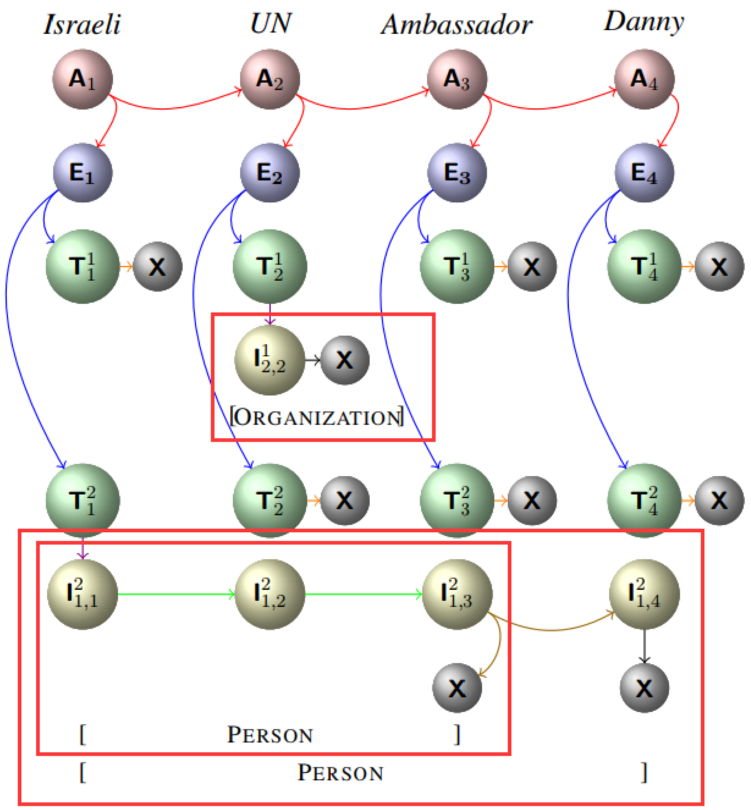
a hypergraph is a generalization of a graph, where an edge can connect ***<u>any number of vertices</u>****.*

- *Joint Mention Extraction and Classification with Mention Hypergraphs. EMNLP 2015: 857-867*
- *Labeling Gaps Between Words: Recognizing Overlapping Mentions with Mention Separators. EMNLP 2017: 2608-2618*
- *Nested Named Entity Recognition Revisited. NAACL-HLT 2018: 861-871*
- *Neural Segmental Hypergraphs for Overlapping Mention Recognition. EMNLP 2018: 204-214*

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Hypergraph model

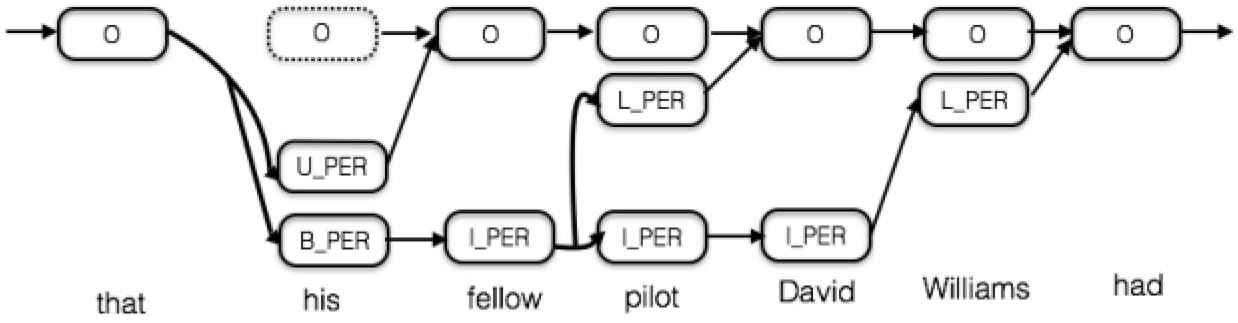


- $\mathbf{A}_i$ encodes all such mentions that start with the i-th or a later word
- $\mathbf{E}_i$ encodes all mentions that start exactly with the i-th word
- $\mathbf{T}_i^k$ represents all mentions of type k starting with the i-th word
- $\mathbf{I}_i^k$ represents all mentions of type k that contain the j-th word and start with the i-th word
- $\mathbf{X}$ marks the end of a mention

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Hypergraph model



- One hyperedge presents a separate valid labeling of mention.

二、Modeling Information Extraction End-to-end

- End-to-end modeling
  - Table-filling/Grid-tagging model

- **Sequence labeling scheme**

[O,I,O,O,I]

✓ 1-D sequential tagging
✓ Extracting flat mention

- **Table-filling scheme**

$$\begin{bmatrix} O & O & O & O & O \\ O & O & O & O & N \\ O & O & O & O & O \\ O & O & O & O & O \\ O & P & O & O & O \end{bmatrix}$$

✓ 2-D grid tagging
✓ Extracting complex mention
✓ Representing relation

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Table-filling/Grid-tagging model

Tags	Meanings
A	two words of word-pair (w_i, w_j) belong to the same aspect term.
O	two words of word-pair (w_i, w_j) belong to the same opinion term.
P	two words of word-pair (w_i, w_j) respectively belong to an aspect term and an opinion term, and they form opinion pair relation.
N	no above three relations for word-pair (w_i, w_j).

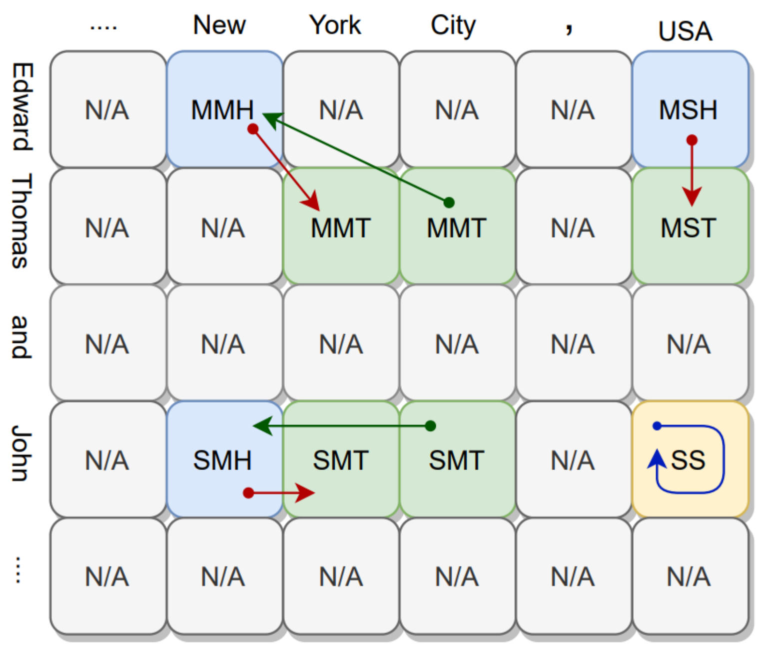
	The	hot	dogs	are	top	notch	but	average	coffee
The	N	N	N	N	N	N	N	N	N
hot		A	A	N	Pos	Pos	N	N	N
dogs			A	N	Pos	Pos	N	N	N
are				N	N	N	N	N	N
top					O	O	N	N	N
notch						O	N	N	N
but							N	N	N
average								O	Neu
coffee									A

- *A Novel Global Feature-Oriented Relational Triple Extraction Model based on Table Filling. EMNLP (1) 2021: 2646-2656*
- *Grid Tagging Scheme for End-to-End Fine-grained Opinion Extraction. EMNLP (Findings) 2020: 2576-2585*

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Table-filling/Grid-tagging model



✓ Advantages:

- Able to model many complex structure
- Highly parallel computation
- Strong representation capability
- Easy intra-feature modeling/encoding
- Easy re-production

- *A Novel Global Feature-Oriented Relational Triple Extraction Model based on Table Filling. EMNLP (1) 2021: 2646-2656*
- *Grid Tagging Scheme for End-to-End Fine-grained Opinion Extraction. EMNLP (Findings) 2020: 2576-2585*

二、Modeling Information Extraction End-to-end

- End-to-end modeling
  - Seq2seq (encoder-decoder) model

✓ Advantages:

- Linearizing everything, sequence in sequence out
- Taking better advantage of GLM

- Other methods:

text → structure

- Seq2seq methods:

text → sequence generation → structure

(structure linearization)

二、Modeling Information Extraction End-to-end

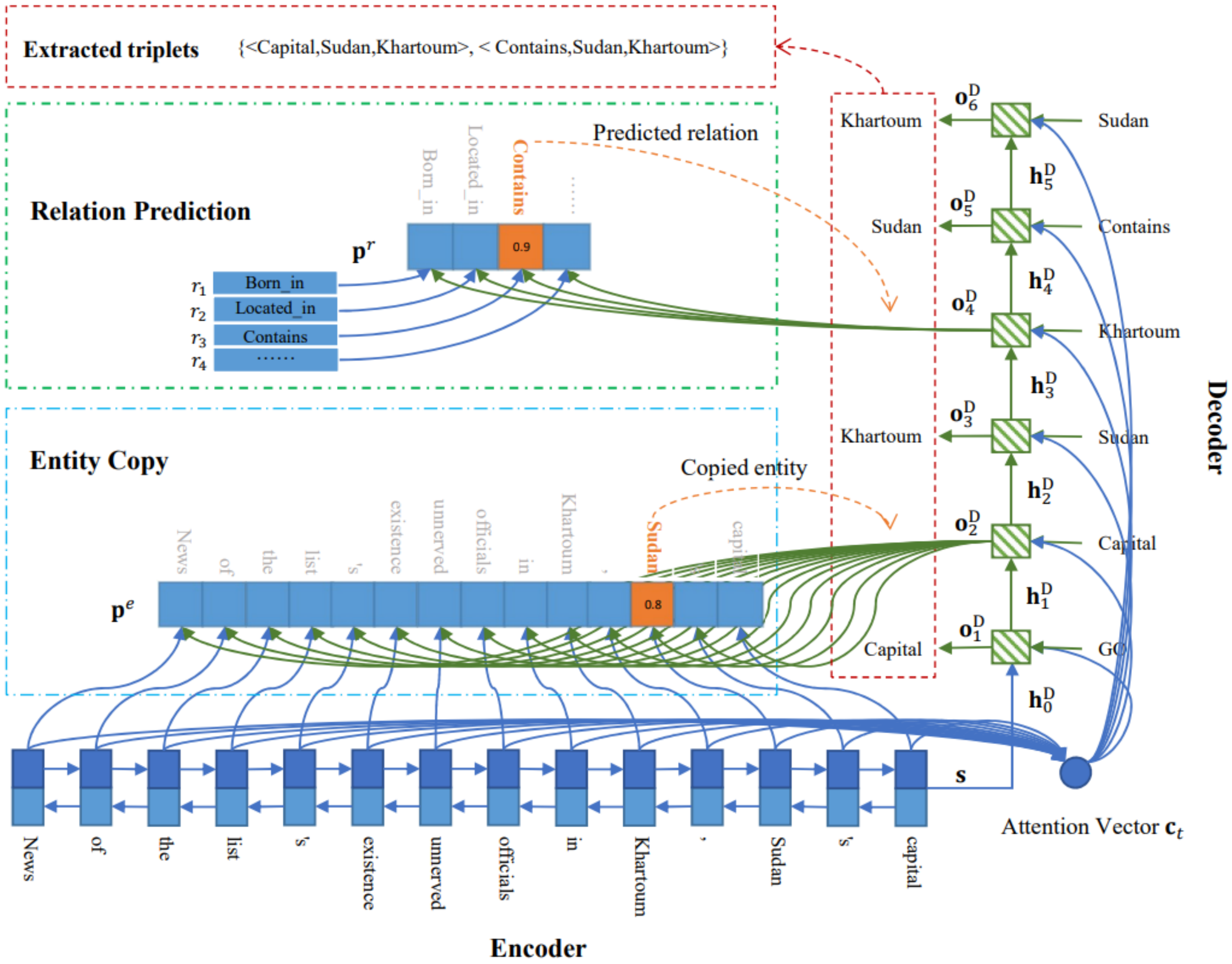
- End-to-end modeling
  - Seq2seq (encoder-decoder) model
    - Copy mechanism
    - Pointer Net
    - Generative LM
      - + Prompt learning

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Seq2seq (encoder-decoder) model

- ☐ Copy mechanism



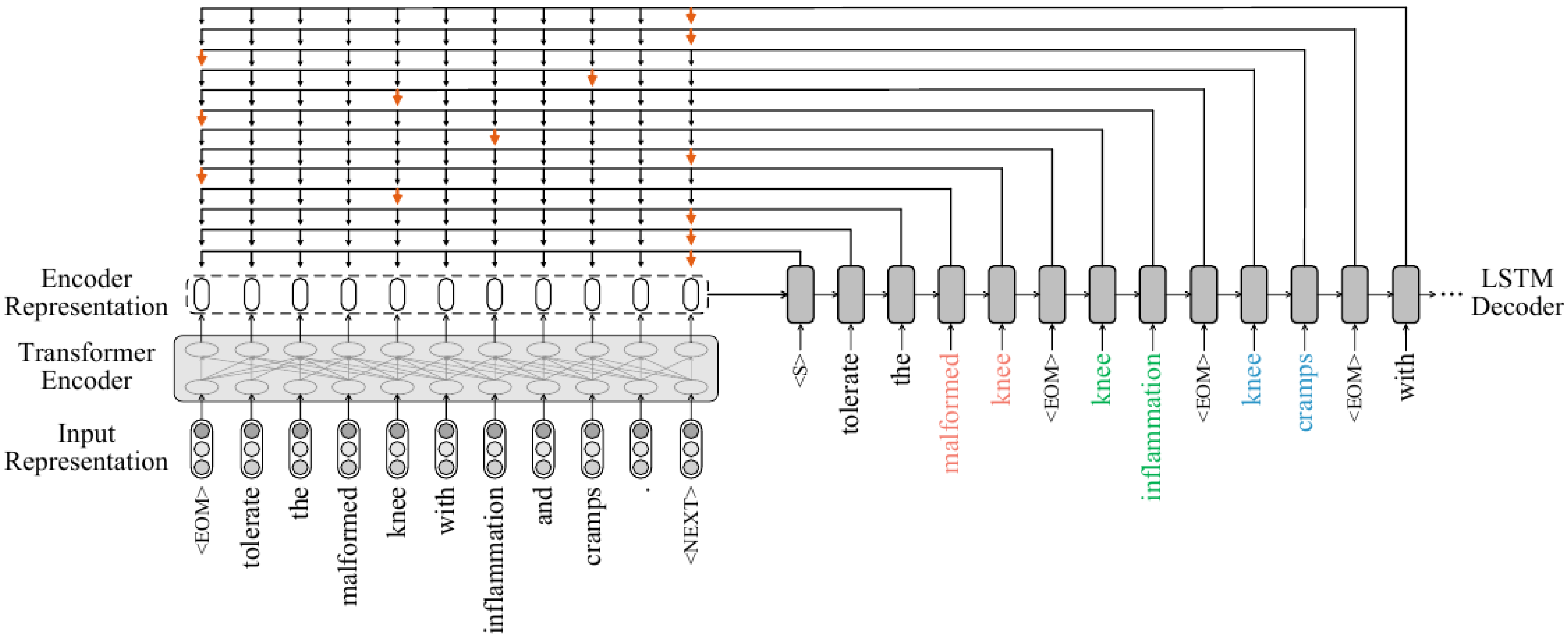
- *Extracting Relational Facts by an End-to-End Neural Model with Copy Mechanism. ACL (1) 2018: 506-514*

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Seq2seq (encoder-decoder) model

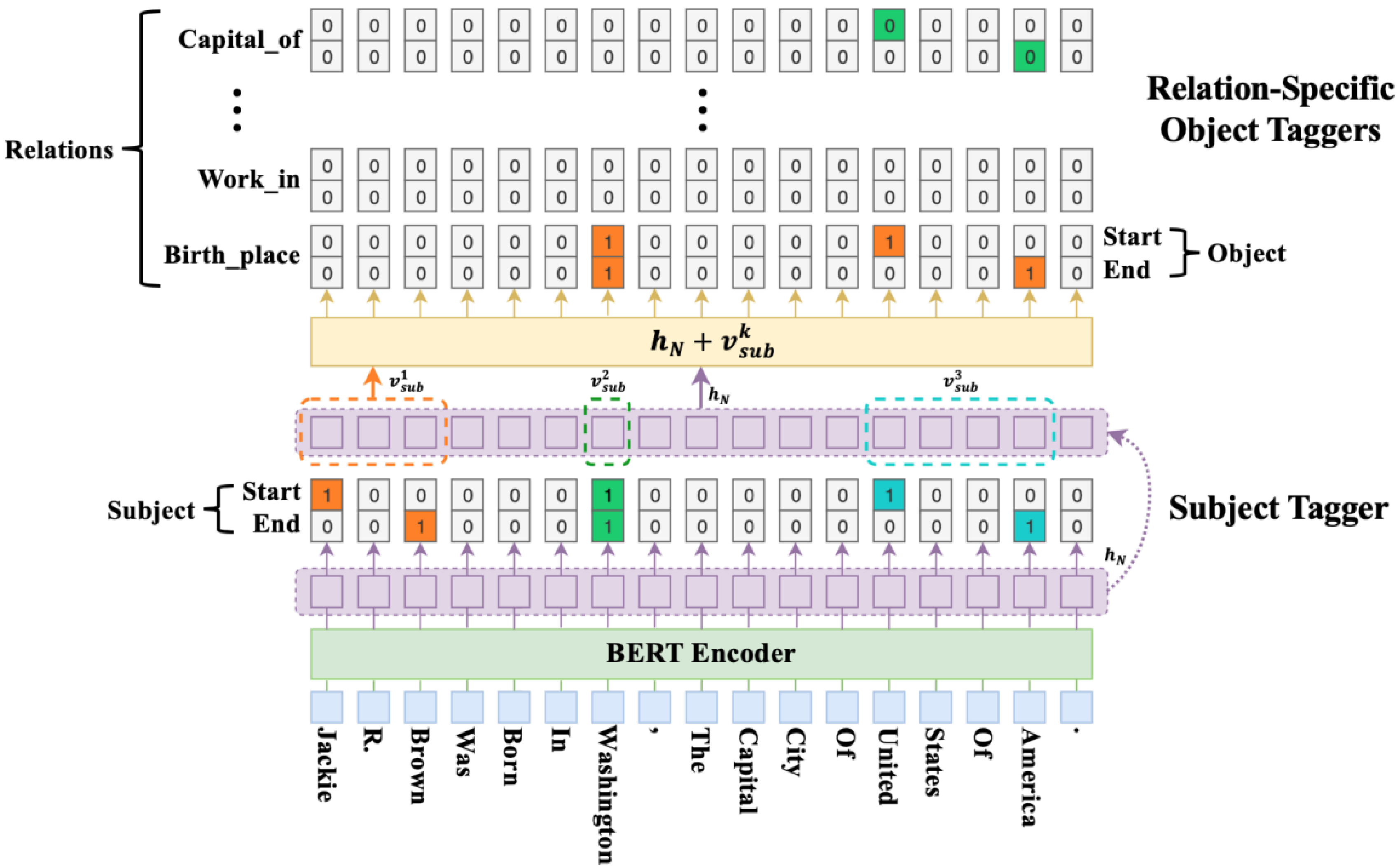
☐ Pointer Net



- *Rethinking Boundaries: End-To-End Recognition of Discontinuous Mentions with Pointer Networks. AAAI 2021: 12785-12793*

二、Modeling Information Extraction End-to-end

- End-to-end modeling
  - Seq2seq (encode
    - Pointer Net



- *A Novel Cascade Binary Tagging Framework for Relational Triple Extraction. ACL 2020: 1476-1488*

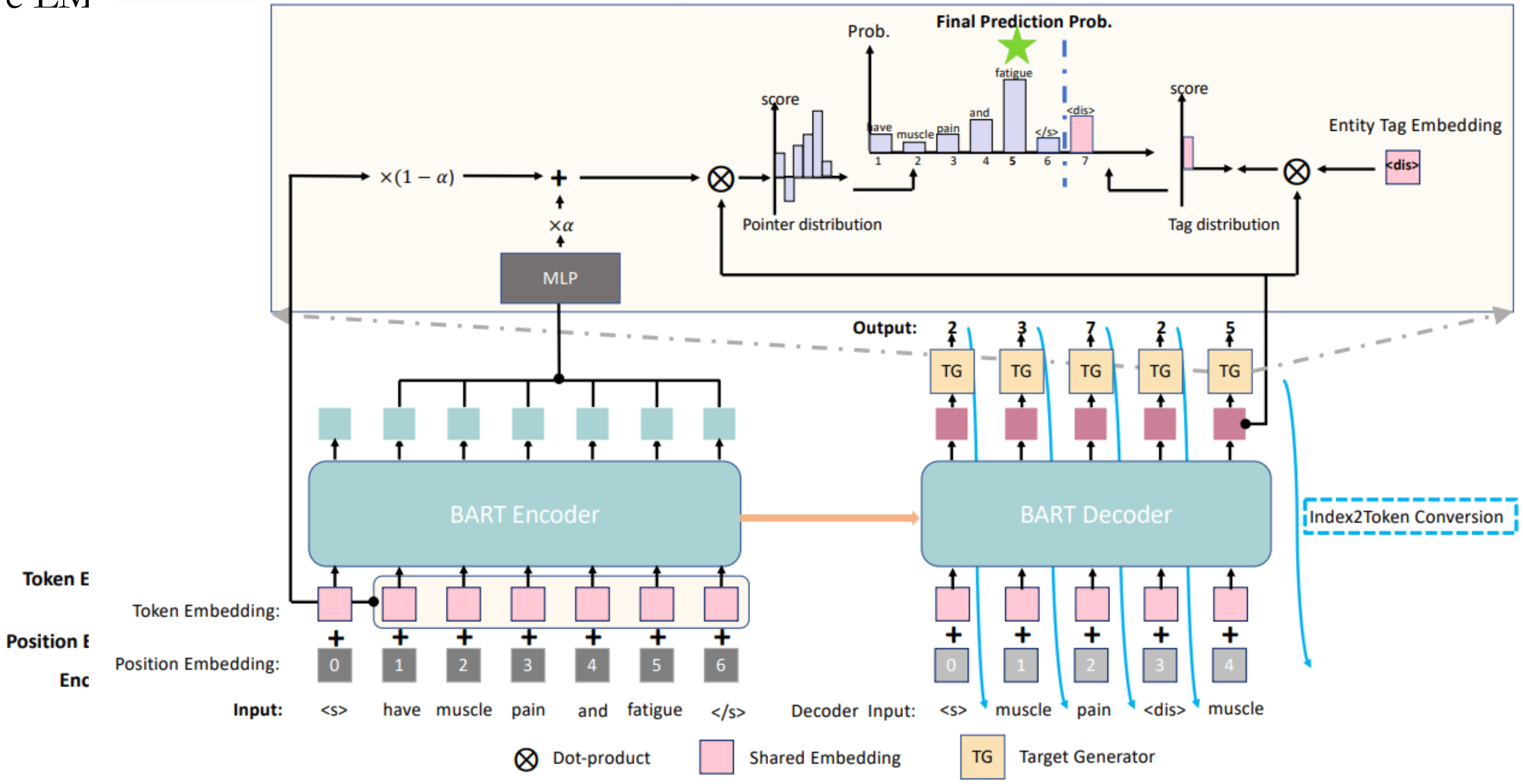
二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Seq2seq (encoder-decoder) model

☐ Generative LM

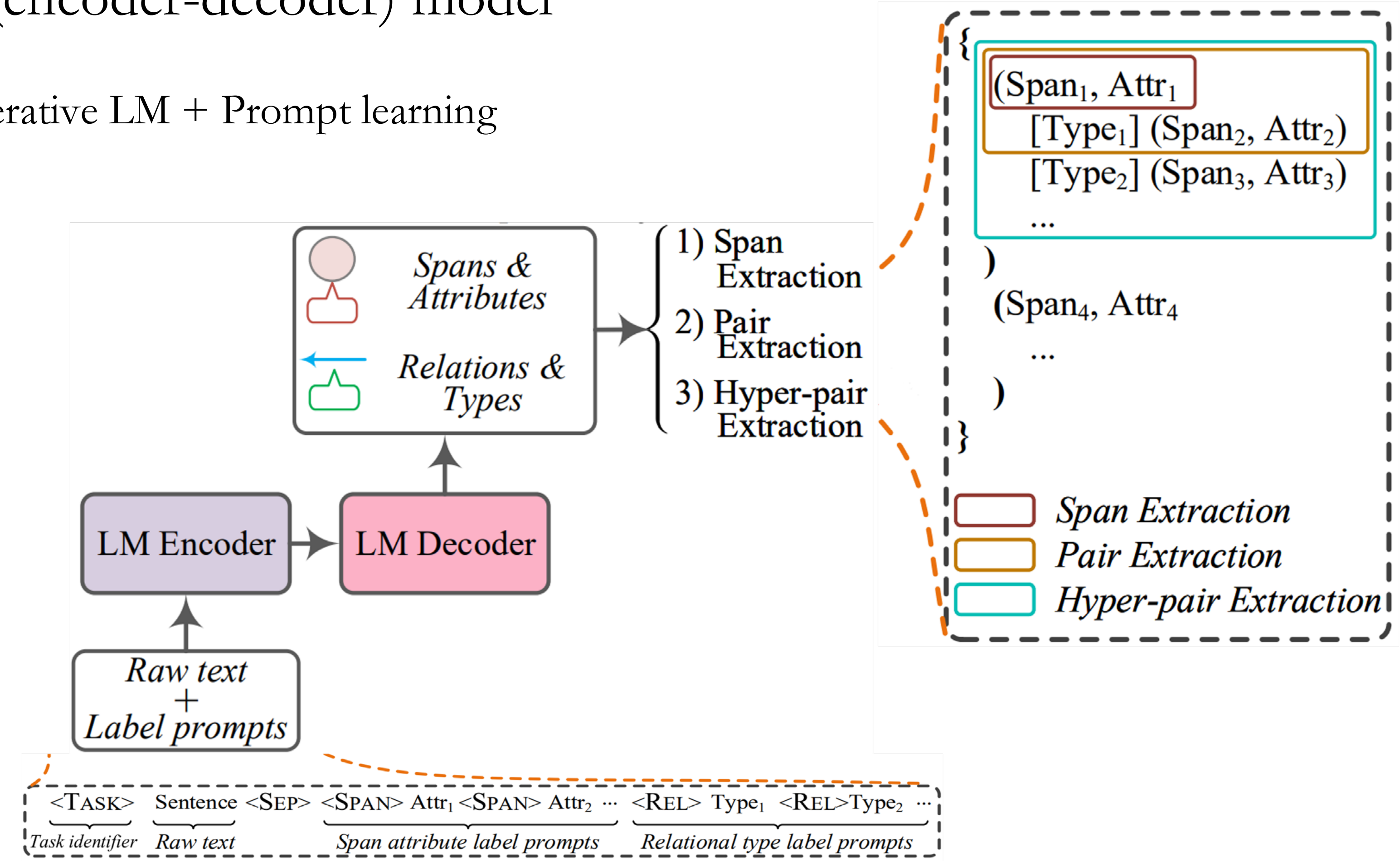
Input: <s> have muscle pain and fatigue </s>
Output: 2 3 7 2 5 6



- *A Unified Generative Framework for Various NER Subtasks. ACL/IJCNLP (1) 2021: 5808-5822*
- *A Unified Generative Framework for Aspect-based Sentiment Analysis. ACL/IJCNLP (1) 2021: 2416-2429*

二、Modeling Information Extraction End-to-end

- End-to-end modeling
  - Seq2seq (encoder-decoder) model
    - Generative LM + Prompt learning



- *Unified Structure Generation for Universal Information Extraction. ACL (1) 2022: 5755-5772*

二、Modeling Information Extraction End-to-end

- End-to-end modeling
  - Seq2seq (encoder-decoder) model
    - Generative LM + Prompt learning

Input

Next week' s trial of Mee is expected to attract widespread media attention .

Output

{ (trial , trial hearing [defendant] (Mee , argument) [time] (Next week , argument)) }

- *Unified Structure Generation for Universal Information Extraction. ACL (1) 2022: 5755-5772*

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transforming into MRC-QA

◆ Core idea:

✓ *Re-formatting the raw structure parsing job as in a machine reading comprehension & QA task, based on the pointer network.*

✓ *With MRC framework, treating the given input text and structure labels and manually constructed prompt queries/questions as semantic prior information, for better task prediction.*

- *A Unified **MRC** Framework for Named Entity Recognition. ACL 2020: 5849-5859*
- *An **MRC** Framework for Semantic Role Labeling. CoRR abs/2109.06660 (2021)*
- *A Joint Training **Dual-MRC** Framework for Aspect Based Sentiment Analysis. AAAI 2021: 13543-13551*
- *Dependency Parsing as **MRC-based** Span-Span Prediction. ACL (1) 2022: 2427-2437*
- *MRC4BioER: Joint extraction of biomedical entities and relations in **MRC** framework. J. Biomed. Informatics 125: 103956 (2022)*

二、Modeling Information Extraction End-to-end

◆ End-to-end modeling

➢ Transforming into MRC-QA

◆ Constructing query templates.

(1) ABSA as MRC

Original training example:

- **input text**: The ambience was nice , but service was not so great.
- **annotations**: (ambience, nice, positive), (service, no so great, negative)

⇩

Converted training example 1:

- **query-1**: Find the *aspect terms* in the text.
- **answer-1**: ambience, service
- **query-2**: Find the *sentiment polarity* and *opinion terms* for ambience in the text.
- **answer-2**: (nice, positive)

Converted training example 2:

- **query-1**: Find the *aspect terms* in the text.
- **answer-1**: ambience, service
- **query-2**: Find the *sentiment polarity* and *opinion terms* for service in the text.
- **answer-2**: (not so great, negative)

(1) SRL as MRC

Input Sentence

The stock has been < p> beaten </p> down for two days.

Multiple-Choice MRC for Predicate Disambiguation

Question: What is the sense of predicate “beaten”?

A. (Cause) pulsating motion that often makes sound

B. push, cause motion

C. win over some competitor

Answer: B

Extractive MRC for Argument Labeling

Question for A0: What are the arguments with meaning "causer of motion"?

Answer: No Answer

Question for A1: What are the arguments with meaning "thing moving"?

Answer: the stock

Question for A2: What are the arguments with meaning "direction, destination"?

Answer: down

Question for TMP: What are the time modifiers of predicate “beaten”?

Answer: for two days

- *An **MRC** Framework for Semantic Role Labeling. CoRR abs/2109.06660 (2021)*
- *A Joint Training **Dual-MRC** Framework for Aspect Based Sentiment Analysis. AAAI 2021: 13543-13551*

二、Modeling Information Extraction End-to-end

- Trends for end-to-end modeling: What to do next?

Unifying & Sharing

- Two key challenges in IE:
  - more accurate boundary detection of mention spans.
  - more intelligent relation assignment between mentions.

二、Modeling Information Extraction End-to-end

◆ Trends for end-to-end modeling: What to do next?

➢ Some unsolved challenges:

- Unnormalized information extraction/structure parsing
  - IE in social media text, casual/colloquial expressions
  - Financial IE, numeric mentions, numeric-text mixed mentions
- Linguistic challenges
  - Coreference
  - Word ambiguity
- Multimodal IE
  - Text + Image
  - Text + Image + Audio

二、Modeling Information Extraction End-to-end

◆ Trends for end-to-end modeling: What to do next?

➢ Joint prediction of a homogeneous type of tasks

◆ Core idea:

Jointly modeling many tasks in one same topic with one unified framework.

◆ Feasibility:

Tasks in homogeneous type essentially share same/common features.

◆ Advantages:

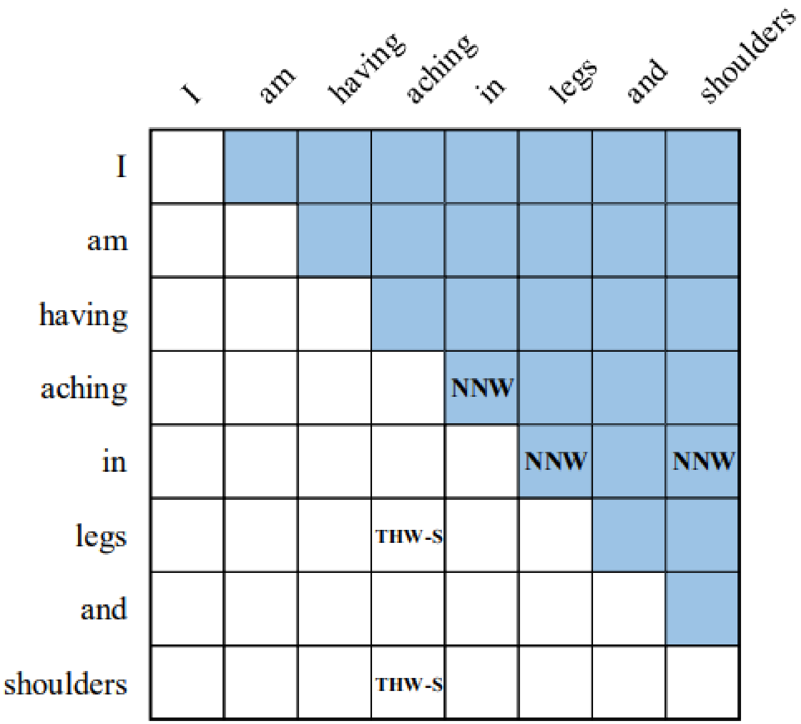
Unified modeling, one model for many tasks,
Better feature reuse, collaboration,
Stronger capability on few-shot learning.
…

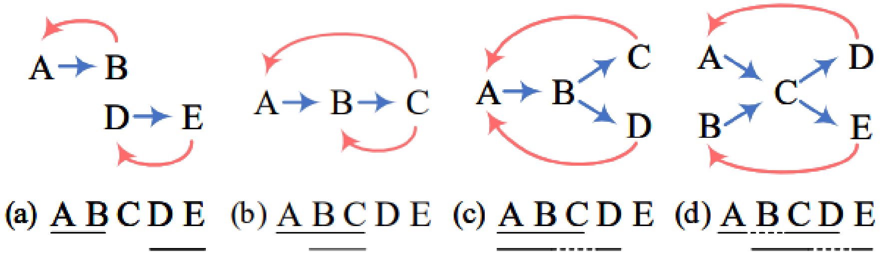
二、Modeling Information Extraction End-to-end

◆ Trends for end-to-end modeling: What to do next?

➢ Joint prediction of a homogeneous type of tasks

□ Unified end-to-end NER





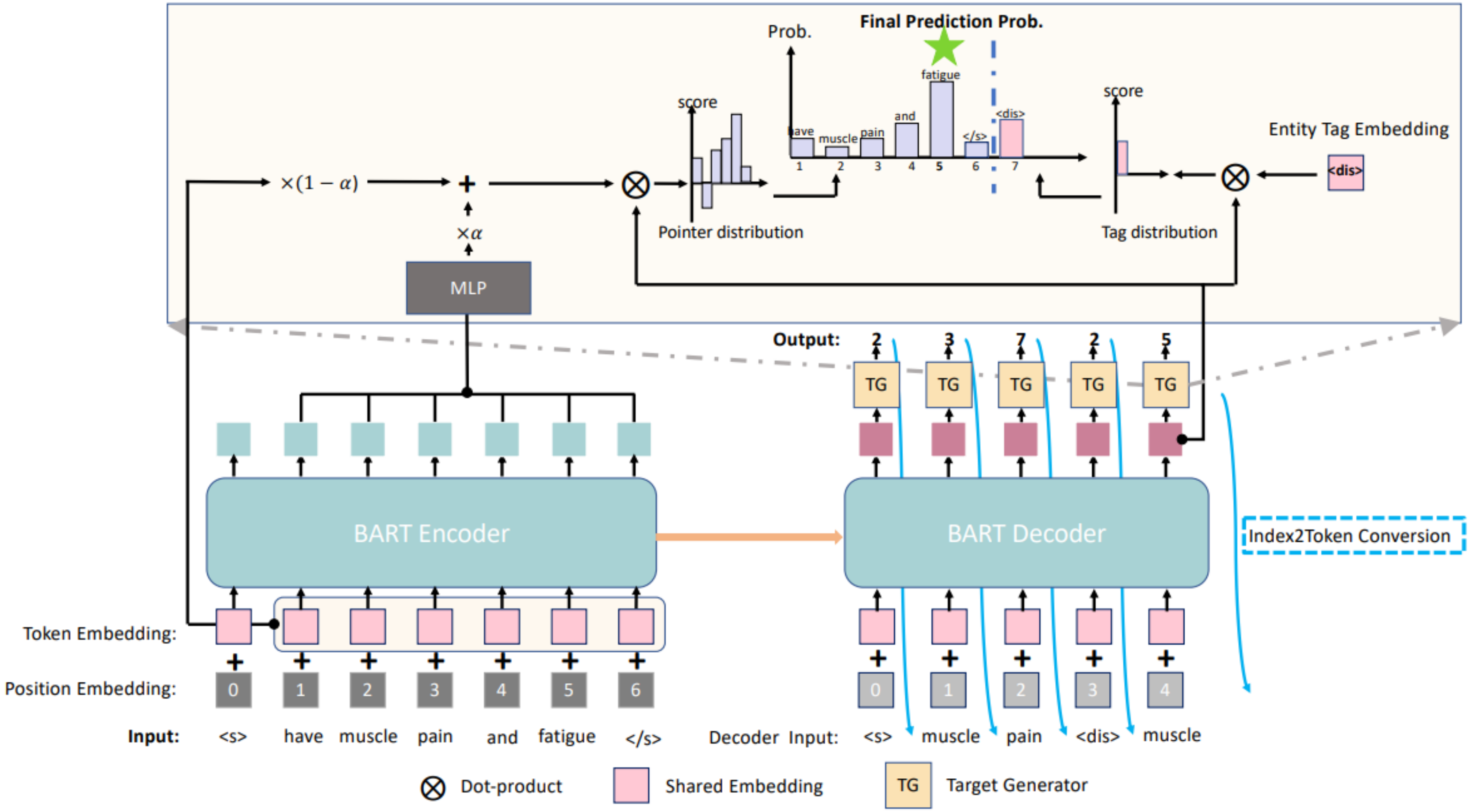
二、Modeling Information Extraction End-to-end

◆ Trends for end-to-end modeling: What to do next?

➢ Joint prediction of a homogeneous type of tasks

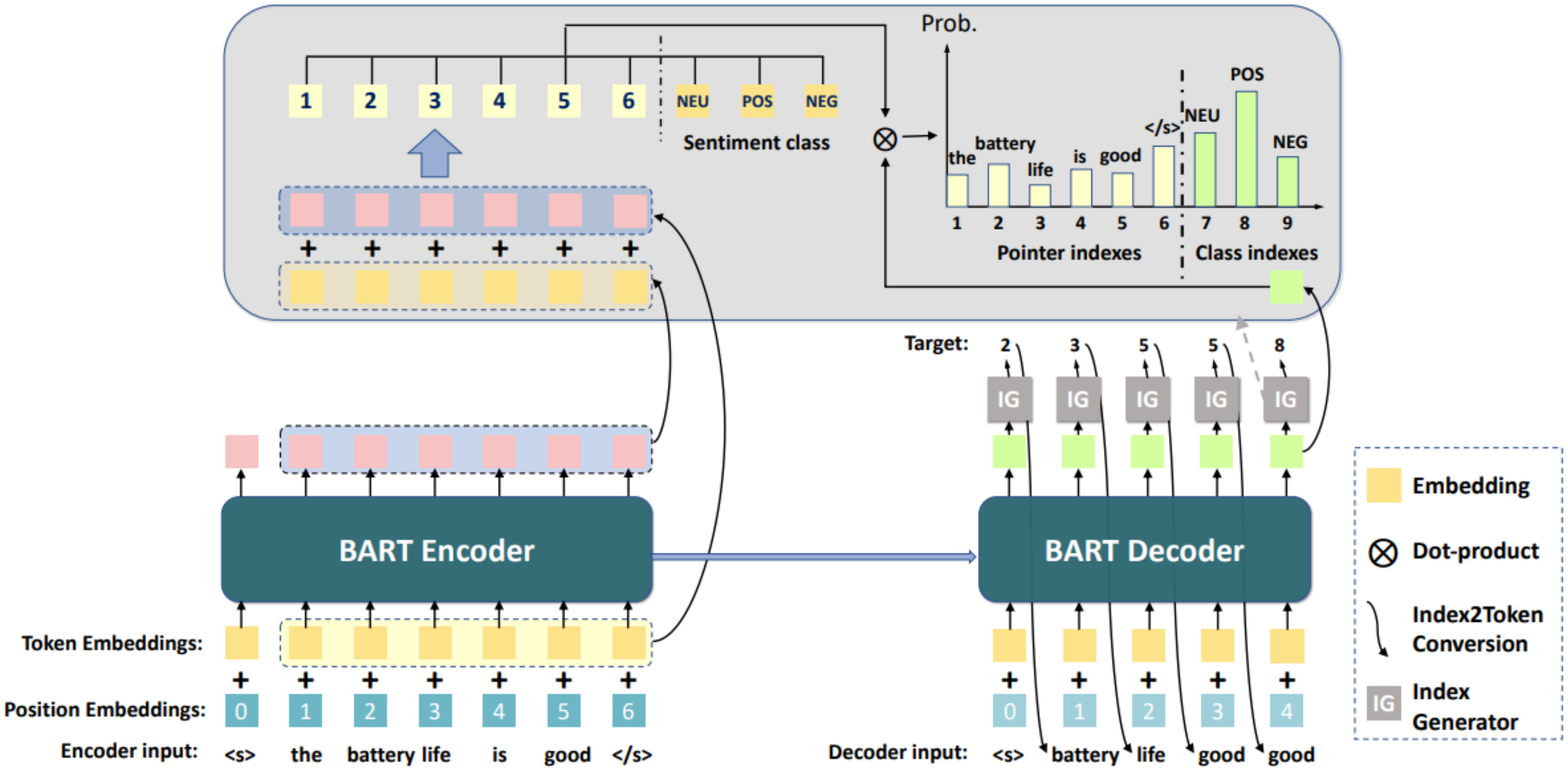
- ☐ Unified end-to-end NER

Input: <s> have muscle pain and fatigue </s>
Output: 2 3 7 2 5 6



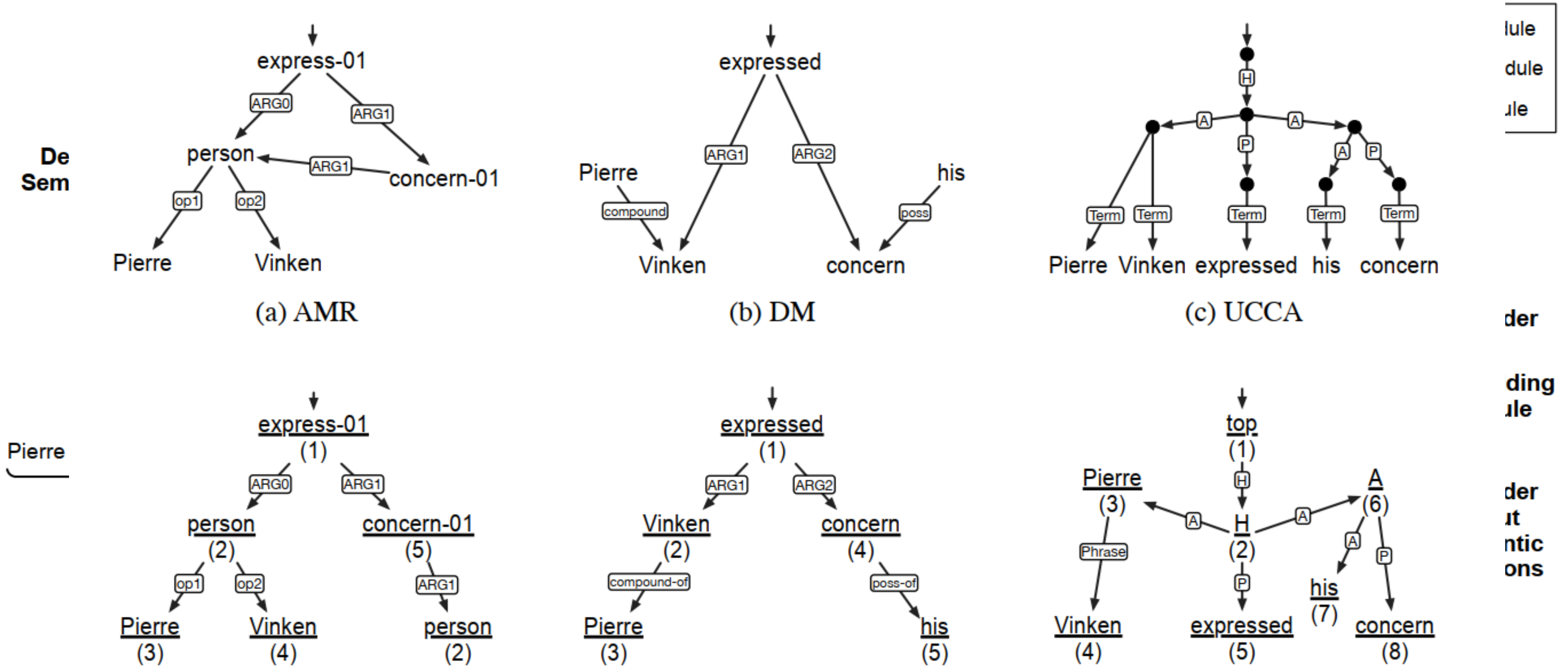
二、Modeling Information Extraction End-to-end

- ◆ Trends for end-to-end modeling: What to do next?
  - ➢ Joint prediction of a homogeneous type of tasks
    - ❑ Unified end-to-end ABSA



二、Modeling Information Extraction End-to-end

- Trends for end-to-end modeling: What to do next?
  - Joint prediction of a homogeneous type of tasks
    - Unified syntax/semantic parsing



(a) AMR (b) DM (c) UCCA

(d) AMR arborescence (e) DM arborescence (f) UCCA arborescence

二、Modeling Information Extraction End-to-end

- Trends for end-to-end modeling: What to do next?
  - Universal extraction
    - Core idea:

      *Jointly modeling <u>**ALL** IE task</u> with one unified framework.*

    - Feasibility:

      *All IE tasks essentially depend mostly on **boundary detection** & relation assignment.*

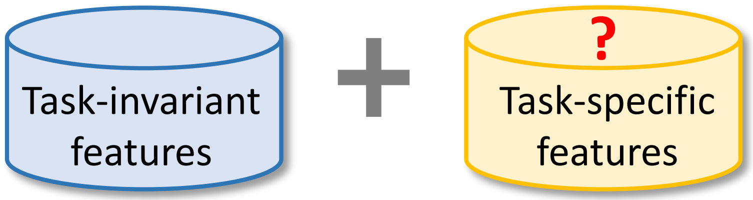
    - Advantages:

      *Universal modeling, one model for **ALL** tasks, especially for real-world production,*
      Best feature reuse, collaboration,
      With PLM, lower dependence on in-demand annotated training data,
      Stronger capability on few-shot learning (cross-task, cross-domain),
      …

      <u>Easier to receive big impact in research community.</u>

二、Modeling Information Extraction End-to-end

- ◆ Trends for end-to-end modeling: What to do next?
  - ➢ Universal extraction
    - ◆ Key requirements & challenge:
      - *Different tasks in different type/genre rely much on learning distinct & unique features.*



      - *So, how to properly coordinate the feature learning and best satisfy all tasks' specific feature?*

二、Modeling Information Extraction End-to-end

- Trends for end-to-end modeling: What to do next?
  - Universal extraction
    - TODO
      - *Modeling UIE with better & sophisticated pretraining language models.*
      - *More feasible modeling scheme of UIE, e.g., text-to-table.*
      - *Minimizing the gap between different feature spaces of different tasks, e.g., constructing more sophisticated optimization algorithm. (Machine learning)*
      - *With better and plausible universal feature corporation:*
        - *External knowledge graph*
        - *Syntactic features*
      - *Cross-lingual universal structure learning/information extraction.*
      - *Multimodal universal information extraction.*

Thank you