Autoregressive Structured Prediction with Language Models

Tianyu Liu[ς]   **Yuchen Eleanor Jiang**[ς]
Nicholas Monath[γ]   **Ryan Cotterell**[ς]   **Mrinmaya Sachan**[ς]

[ς]ETH Zürich   [γ]Google Research

{tianyu.liu, yuchen.jiang}@inf.ethz.ch

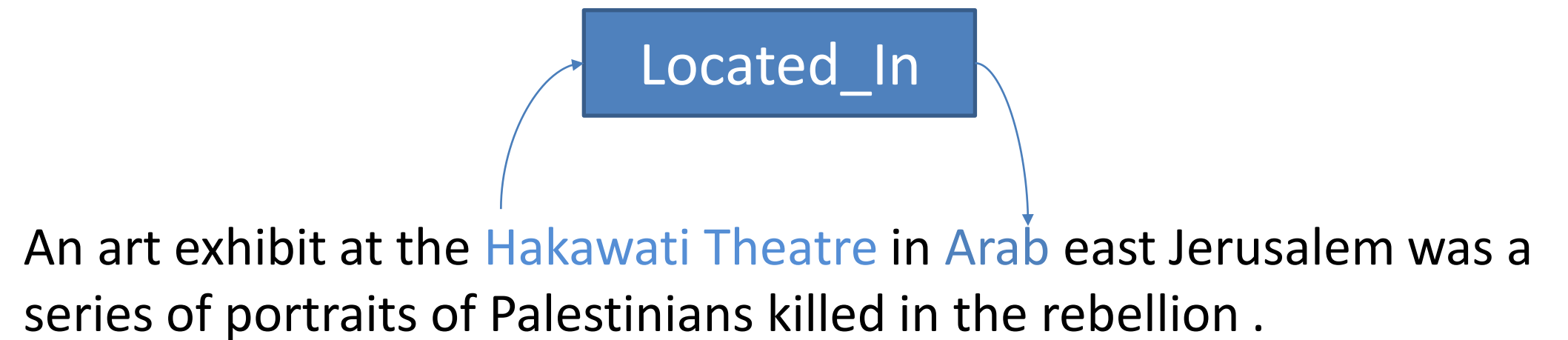nmonath@google.com   {ryan.cotterell, mrinmaya.sachan}@inf.ethz.ch

EMNLP2022

汇报人：李雨倩

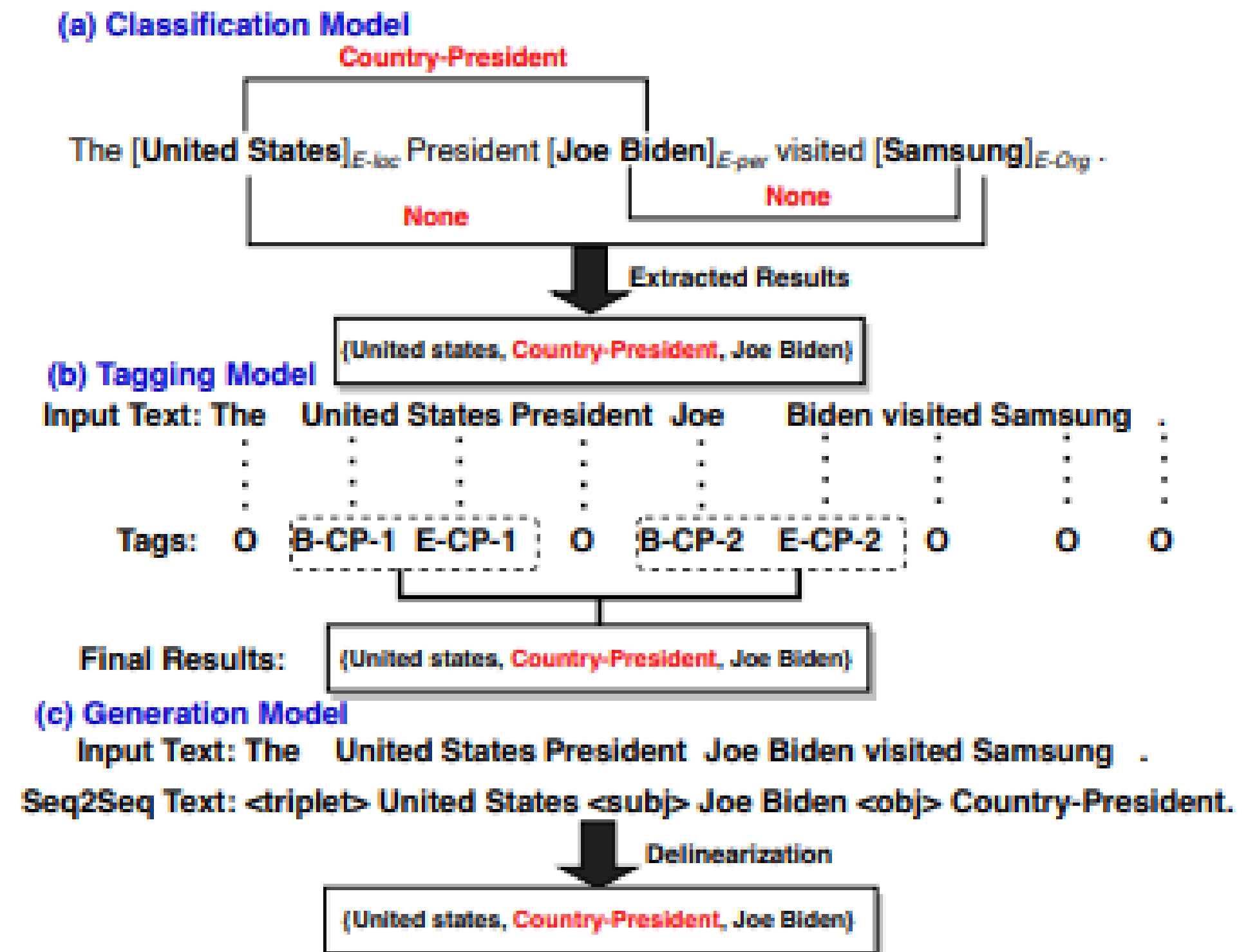# Task

Structured Prediction:

- Named Entity Recognition

- Relation Extraction

- Coreference Resolution

Located_In

An art exhibit at the Hakawati Theatre in Arab east Jerusalem was a series of portraits of Palestinians killed in the rebellion .

Methods

- **Classification Model**

- **Tagging Model**

- **Generation Model**

  开发一个通用架构来同时解决不同的任务

(a) Classification Model
Country-President

The [United States]$_{E\text{-}loc}$ President [**Joe Biden**]$_{E\text{-}per}$ visited [**Samsung**]$_{E\text{-}Org}$.

None   None

Extracted Results

(United states, Country-President, Joe Biden)

(b) Tagging Model
Input Text: The   United States President   Joe   Biden visited Samsung   .

Tags: O   B-CP-1 E-CP-1   O   B-CP-2   E-CP-2   O   O   O

Final Results: (United states, Country-President, Joe Biden)

(c) Generation Model
Input Text: The   United States President Joe Biden visited Samsung   .

Seq2Seq Text: <triplet> United States <subj> Joe Biden <obj> Country-President.

Delinearization

(United states, Country-President, Joe Biden)

Motivation

- String(flatten the target structure)

Intent = AddToPlaylist

Sentence	Add	Kent	James	to	the	Disney	soundtrack
Slot labels	O	B-artist	I-artist	O	O	B-playlist	O

⇕

((AddToPlaylist)) Add [Kent James | artist] to the [Disney | playlist] soundtrack.

⟹ Action

Approach

Action sequence: $y = y_1 y_2 \ldots y_N$

Action space: $\mathcal{Y}_n = (\mathcal{A} \times \mathcal{B}_n \times \mathcal{Z}_n)$
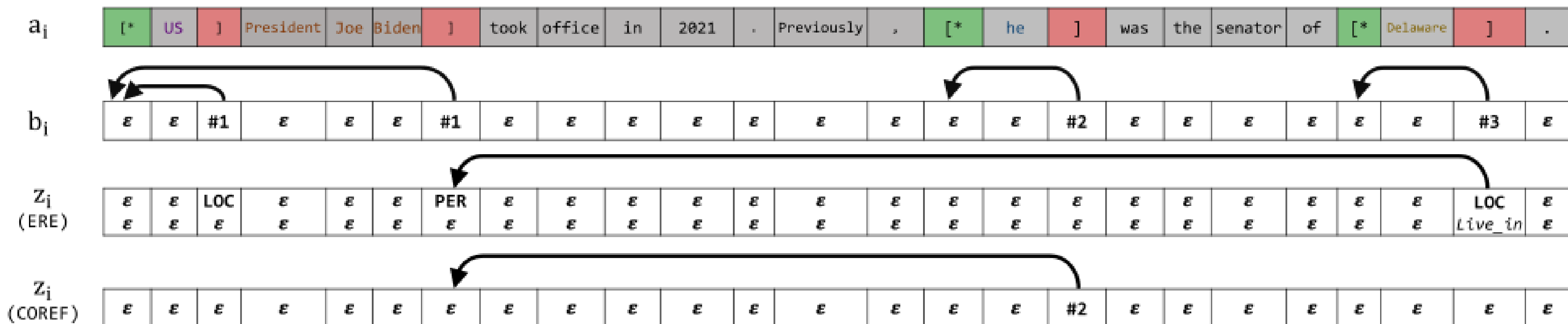
Approach

Structure-Building Actions: $\mathcal{A} = \{[\,,]\,,COPY\}$

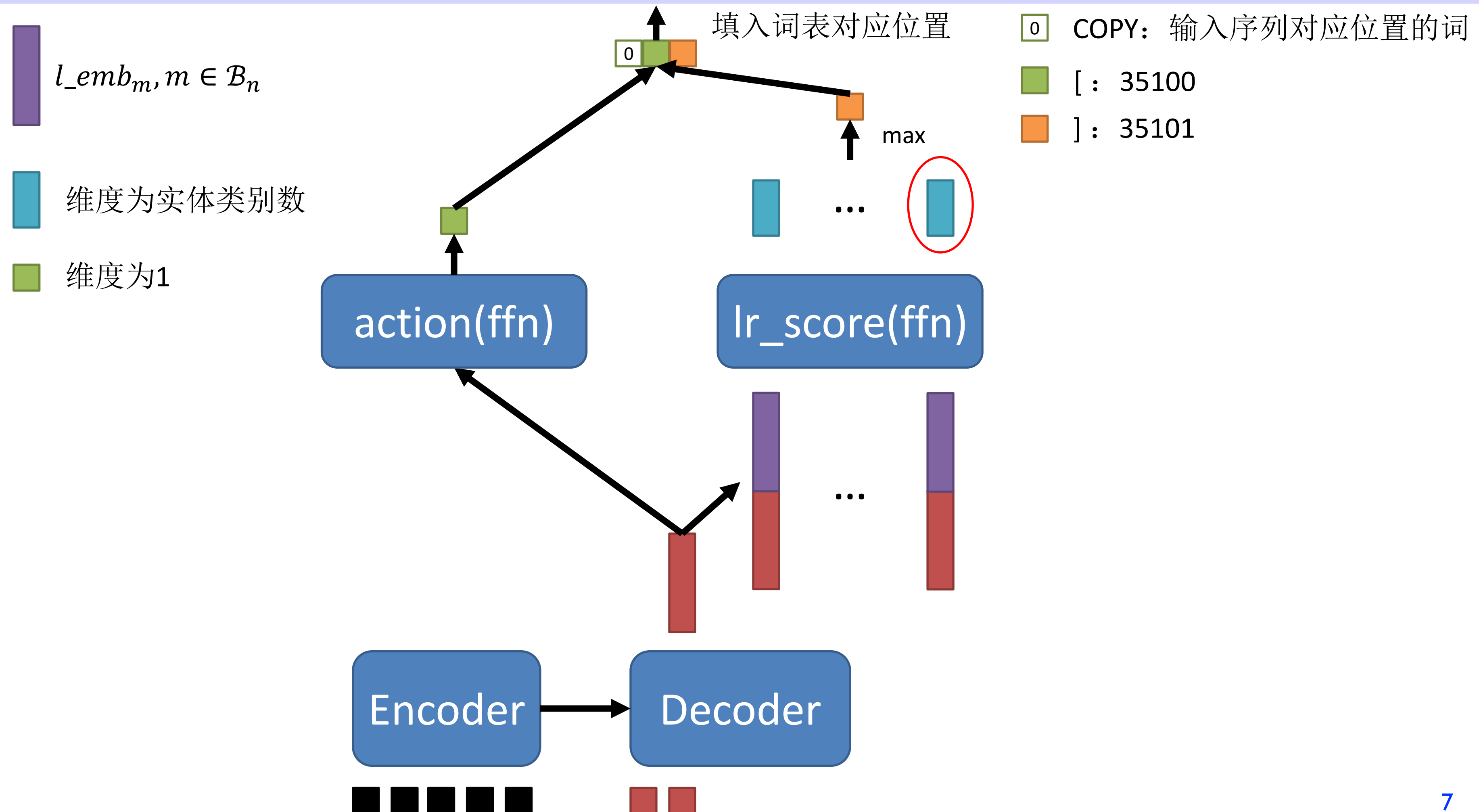Bracket-Pairing Actions $\mathcal{B}_n = \{m|m < n \wedge a_m = [\,\}$

Span-Labeling Actions $\mathcal{Z}_n = \{m|m < n \wedge a_m =]\,\} \times \mathcal{L}$

Named Entity Recognition: Inference

$l\_emb_m, m \in \mathcal{B}_n$

维度为实体类别数

维度为1

填入词表对应位置

COPY：输入序列对应位置的词

[：35100

]：35101

max

action(ffn)

lr_score(ffn)

Encoder

Decoder

0	a	l

$$y = \max(\quad) + logsumexp$$

COPY:

0	-inf	-inf

$$\hat{y}_{COPY} = 0$$

[:

-inf	a	-inf

$$\hat{y}_{[} = a$$

]:

-inf	-inf	l

$$\hat{y}_{]} = l$$

Named Entity Recognition: Result

	Prec.	Rec.	F1
Ma and Hovy (2016)	91.4	91.1	91.2
Devlin et al.+BERT_L	-	-	92.8
Ye et al.+ROBERTA_L	-	-	94.0
Athiwaratkun et al.	-	-	91.5
Paolini et al.+T5_B	-	-	91.7
ASP+T5_B	91.4	92.2	91.8
ASP+T5_L	92.1	93.4	92.8
ASP+T5_3B	93.8	94.4	**94.1**

Bi-LSTMs+CNN+CRF ← Ma and Hovy (2016)

Packed Levitated Marker ← Ye et al.+ROBERTA

T5+ flattened-string ← Athiwaratkun et al.

T5+ flattened-string ← Paolini et al.+T5

Table 1: Test F1 scores of named entity recognition on the CoNLL-03 test set.

CoNLL-03数据集：文档级别的命名实体识别的数据集，包含946篇训练文档，216篇验证文档，231篇测试文档

End-to-End Relation Extraction: Result

	Ent	Rel
Eberts and Ulges (2020)	88.9	71.5
Zhao et al. (2020)	88.9	71.9
Wang and Lu+ALBERT$_{XXL}$	90.1	73.8
Paolini et al.+T5$_B$	89.4	71.4
ASP+T5$_B$	89.5	73.2
ASP+T0$_{3B}$	**90.3**	**76.3**

Table 2: **Micro** F1 scores of entity extraction and relation extraction on the CoNLL-04 joint entity relation extraction test set.

CoNLL-04数据集：句子级别的关系抽取的数据集，标注了4种实体，5种关系；其中，训练集包含922个句子，验证集包含231个句子，测试集包含288个句子

Coreference Resolution: Result

计算F1的不同策略

	MUC	B^3	$CEAF_{\phi_4}$	Avg. F1
Lee et al. (2017)	75.8	65.0	60.8	67.2
Joshi et al. (2020)	85.3	78.1	75.3	79.6
Joshi et al.$+T5_B$ [†]	79.8	70.2	66.8	72.3
Joshi et al.$+T5_L$ [†]	81.4	73.1	73.1	74.9
Urbizu et al.	64.9	66.5	65.3	65.6
Paolini et al.$+T5_B$	81.0	69.0	68.4	72.8
Dobrovolskii	86.3	79.9	76.6	81.0
ASP$+T5_B$	82.3	75.1	72.5	76.6
ASP$+T5_L$	84.7	77.7	75.2	79.3
ASP$+T0_{3B}$	86.9	81.5	78.4	82.3
ASP$+FLAN-T5_{XXL}$	87.2	81.7	78.6	**82.5**

Table 4: Results on the CoNLL-12 English test set. Avg. F1 denotes the average F1 of MUC, B^3, and $CEAF_{\phi_4}$. Models marked with [†] are our re-implementation. Other results are taken from their original papers. The full results are in Tab. 5.

CoNLL-12数据集：文档级别的指代消解的数据集，包含2802篇训练文档，343篇验证文档，348篇测试文档