TEXT2EVENT: Controllable Sequence-to-Structure Generation for End-to-end Event Extraction

Yaojie Lu, Hongyu Lin, Jin Xu, Xianpei Han, Jialong Tang, Annan Li, Le Sun, Meng Liao, Shaoyi Chen

ACL-IJCNLP 2021

Introduction

- Most previous work extracts event records by performing different subtasks.
- This work proposes a new generation-based model to generate events and arguments in an end-to-end manner.

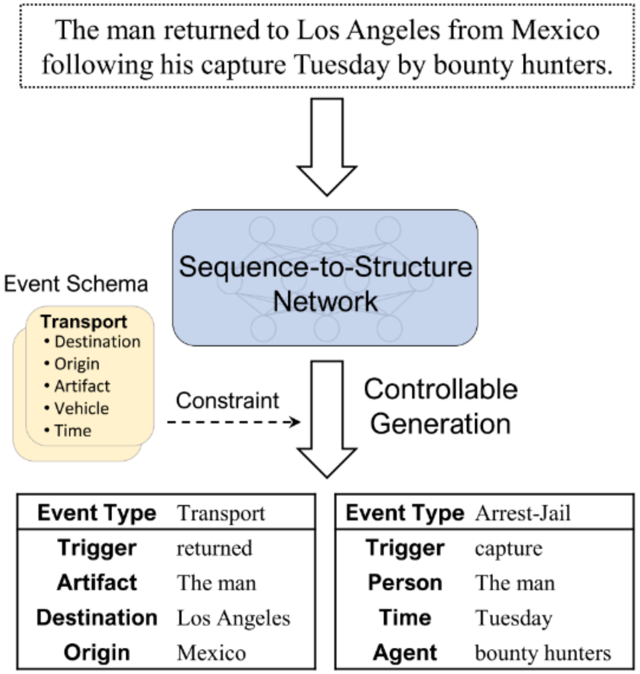
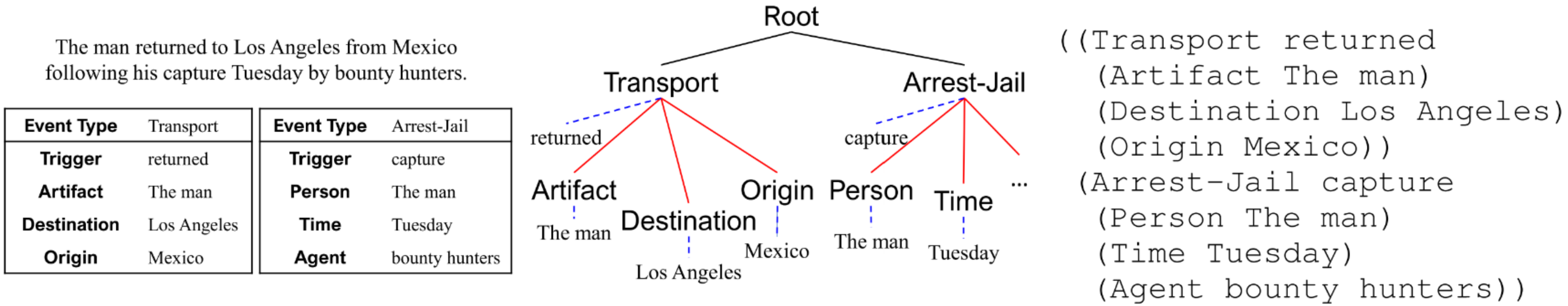


Figure 1: The framework of TEXT2EVENT. Here, TEXT2EVENT takes raw text as input and generates a *Transport* event and an *Arrest-Jail* event.

Event Extraction as Structure Generation

- Text2event uses T5 large for the encoder-decoder component in their model.



(a) Record format. (b) Tree format. (c) Linearized format.

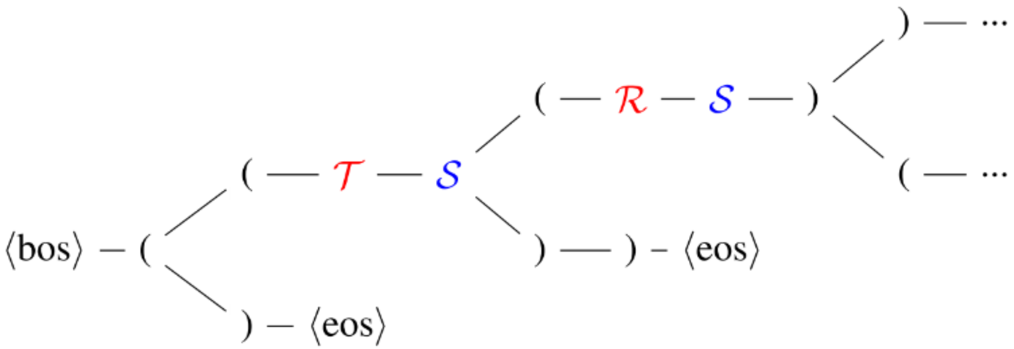
Training

- The output sequence is not in natural language, especially with the frequent presence of brackets “(“ and “)”.
- Proposed solution: curriculum learning:
  - Pretrain T5 model using substructures
  - Train the full model using full structure later.
- Substructure learning:
  (Transport returned) (Artifact The man) (Arrest-Jail capture) …
- Full structure learning:

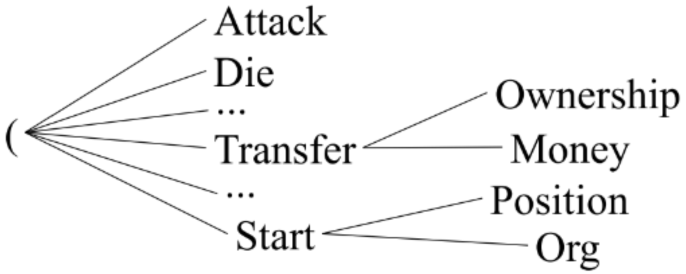
```
((Transport returned
  (Artifact The man)
  (Destination Los Angeles)
  (Origin Mexico))
 (Arrest-Jail capture
  (Person The man)
  (Time Tuesday)
  (Agent bounty hunters))
```

Constrained Decoding

- Employ a trie-based constrained decoding algorithm *(Chen et al., 2020a; Cao et al., 2021)*



(a) The trie of event structure.



(b) The trie of event type $\mathcal{T}$.

Results

Models	Trig-C			Arg-C			PLM
	P	R	F1	P	R	F1	
Models using Token Annotation + Entity Annotation							
Joint3EE (Nguyen and Nguyen, 2019)	68.0	71.8	69.8	52.1	52.1	52.1	-
DYGIE++ (Wadden et al., 2019)	-	-	69.7	-	-	48.8	BERT-large
GAIL (Zhang et al., 2019b)	74.8	69.4	72.0	61.6	45.7	52.4	ELMo
$\text{OneIE}_{\text{w/o Global}}$ (Lin et al., 2020)	-	-	73.5	-	-	53.9	BERT-large
OneIE (Lin et al., 2020)	-	-	74.7	-	-	56.8	BERT-large
Models using Token Annotation							
EEQA (Du and Cardie, 2020)	71.1	73.7	72.4	56.8	50.2	53.3	2×BERT-base
MQAEE (Li et al., 2020)	-	-	71.7	-	-	53.4	3×BERT-large
Generation-based Baselines using Token Annotation							
TANL (Paolini et al., 2021)	-	-	68.4	-	-	47.6	T5-base
Multi-Task TANL (Paolini et al., 2021)	-	-	68.5	-	-	48.5	T5-base
Our Model using Parallel Text-Record Annotation							
TEXT2EVENT	67.5	71.2	69.2	46.7	53.4	49.8	T5-base
TEXT2EVENT	69.6	74.4	71.9	52.5	55.2	53.8	T5-large

Table 2: Experiment results on ACE05-EN. Trig-C indicates trigger identification and classification. Arg-C indicates argument identification and classification. PLM represents the pre-trained language models used by each model.

Datasets	Trig-C			Arg-C		
	P	R	F1	P	R	F1
SOTA (Token + Entity Annotation)						
ACE05-EN^{+}	-	-	72.8	-	-	54.8
ERE-EN^{*}	56.9	58.7	57.8	51.9	47.8	49.8
TEXT2EVENT (Parallel Text-Record Annotation)						
ACE05-EN^{+}	71.2	72.5	71.8	54.0	54.8	54.4
ERE-EN	59.2	59.6	59.4	49.4	47.2	48.3

Table 3: Experiment results on ACE05-EN^{+} and ERE-EN. SOTA indicates the state-of-the-art system – OneIE. * The result of SOTA for ERE-EN is reproduced by the official release code because of the slightly different dataset statistic result on ERE-EN.