

# Effect Generation Based on (Event-Level) Causal Reasoning

# Introduction

- Causal reasoning은 현재 action을 보고 일어날 법한 미래 scenarios를 예측하는 task이다.
- 지금까지의 causal reasoning 방법은 word-level에서 진행되었다.
  - > 이러면 context가 없어서 이해하기가 어렵다.
  - > 본 논문은 word level이 아닌 event level에서 진행한다.
- Observed causal event를 그대로 사용하지 않고 event causality network에 포함시켜 그래프를 만들어서 처리하였다.
  - > 기존 방법에 존재한 sparsity라는 단점을 해결

# Event Causality Network Construction

2단계로 이루어진다.

1. Event Eventification
2. Events Structuralization

# Event Eventification

- Event를 특정한 representation으로 변환시키는 것이 목적
- 먼저, 충분한 cause-effect sentence pairs를 얻는다.
- 원인-결과를 나타내는 접속사로 연결된 sentence pairs를 추출한다.
  - > 'because', 'as a result'
- 이렇게 얻은 sentence pairs로부터 causal event pairs를 얻는다.
  - > dependency analysis를 사용

Event는 다음과 같은 representation으로 표현한다.

$$(s, v, o, m)$$

S = 주어의 head noun

V = 동사

O = 형용사나 direct object의 head noun

M = 전치사의 목적어나 indirect object의 head noun

# Event Structuralization

- 앞서 얻은 **causal event pairs**를 **event causality network**로 변환하는 과정
  - > 의미상 비슷한 events를 cluster화하는 것이다.
  - > **Event abstractions**를 이용해서 의미상 비슷한 지를 평가한다.

# Event Abstractions

- Event의 s, v, o, m을 각각 linguistic resources에 맞춰 generalize한 것을 의미
  - > v는 VerbNet에, 나머지는 WordNet에 맞춘다.

# Event Structuralization

- Node는 event이고 edge의 weight은 다음 규칙으로 부여된다.

- 1) Dataset으로부터 extract된 event pair  $(e_i, e_j)$  의 weight

$$w_{ij} = 1$$

- 2)  $(e_i, e_j)$  가 semantic-similarity가 있는  $(e_i, e_k)$  와 causal relation이 있는  $(e_k, e_j)$ 로부터 추론이 될 때 weight

$$w_{ij} = \text{sim}(e_i, e_k)$$

- $\text{sim}(e_i, e_k)$ 은 WordNet에서 계산한 path-similarity 값으로 semantic-similarity score이기도 하다.

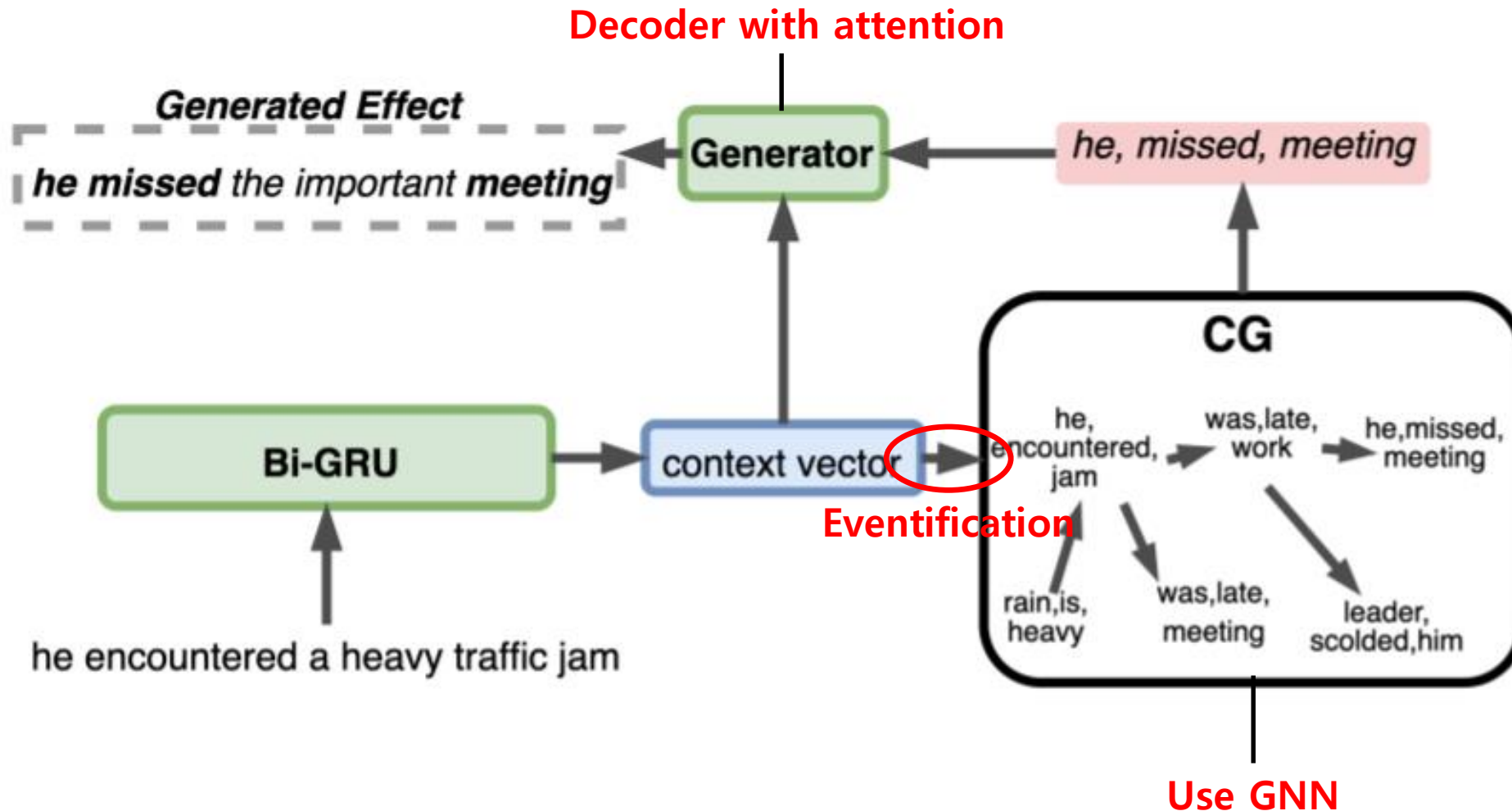


# Effect Generation

- 목표는

- 1) Input cause에 따라 알맞은 **effect event**를 예측하고
- 2) 그 effect event를 **effect sentence**로 **rewrite**하는 것이다.

# Effect Generation



# GNN

$$\mathbf{z}_i^l = \mathbf{W}^l \mathbf{e}_i^l$$

$$\mathbf{e}_i^{l+1} = \sigma \left( \sum_{j=1}^{N_{CG}} \frac{\exp(w_{ij}(\mathbf{z}_i^{lT} \cdot \mathbf{z}_j^l))}{\sum_k \exp(w_{kj}(\mathbf{z}_k^{lT} \cdot \mathbf{z}_j^l))} \mathbf{z}_j^l \right)$$

$\mathbf{e}_i^l$  = event  $e_i$  의  $l$ 번째 layer vector

$\mathbf{W}^l$  = parameter

$\mathbf{e}_i^l$  에서  $\mathbf{e}_i^{l+1}$  를 구하는 식

# Experiments

- Datasets

- 1) English Wikipedia train/validation/test = 8 : 1 : 1

- 2) COPA Benchmark validation/test = 1 : 1

- Use training set to construct the event causality network

# Results

Model	EnWiki				COPA			
	BLEU-4	Distinct-1/2	AbsMat	Plau	BLEU-4	Distinct-1/2	AbsMat	Plau
GPT2	0.69	5.57/16.82	0.3	0.08	1.35	22.61/44.25	0.2	0.02
BART	1.28	8.23/24.83	1.7	0.11	1.22	22.37/43.71	0.5	0.04
CausalBERT	0.74	5.33/22.23	8.5	0.12	0.92	22.39/52.56	3.7	0.06
CopyNet	2.85	10.63/39.82	16.4	0.17	1.18	32.74/75.17	2.6	0.04
EGCER(ours)	4.90	13.99/43.58	26.4	0.27	1.74	48.08/83.97	5.3	0.07

# Generation Example

<b>Input cause</b>	he encountered a heavy traffic jam.
GPT2	the lighthouse was closed over three weeks.
BART	he was delayed for over an hour.
CopyNet	he missed missed the meeting.
CausalBert	causing him to miss bus.
EGCER	he missed the important meeting.

# Ablation Study

Models	BLEU-4	Distinct-1/2	AbsMat	Plau
<i>Full model</i>	4.90	13.99/43.58	26.4	0.27
<i>w/o weights</i>	4.37	14.10/42.86	23.3	0.24
<i>w/o 2nd layer</i>	3.89	13.15/41.56	20.6	0.21
<i>w/o GNN</i>	2.89	13.00/42.02	18.3	0.19

# Conclusion

- **Event-level** causal reasoning 기반의 effect generation 방법을 제안하였다.
  - > 실제로 성능이 뛰어남을 증명하였다.