Event Transition Planning for Open-ended Text Generation

香 港 大 學
THE UNIVERSITY OF HONG KONG

Qintong Li[1], Piji Li[2], Wei Bi[2], Zhaochun Ren[3], Yuxuan Lai[1], Lingpeng Kong[1,4]

[1]The University of Hong Kong, [2]Tencent AI Lab, [3]Shandong University, [4]Shanghai Artificial Intelligence Laboratory

qtleo@outlook.com

Tencent AI Lab

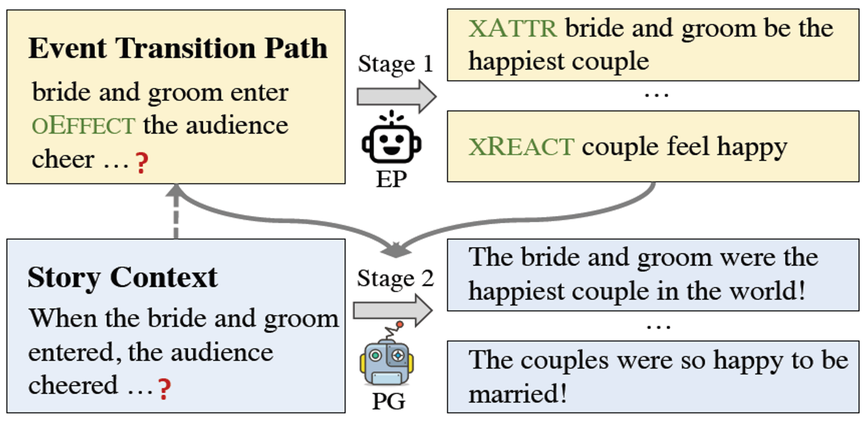
上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

Background and Motivation

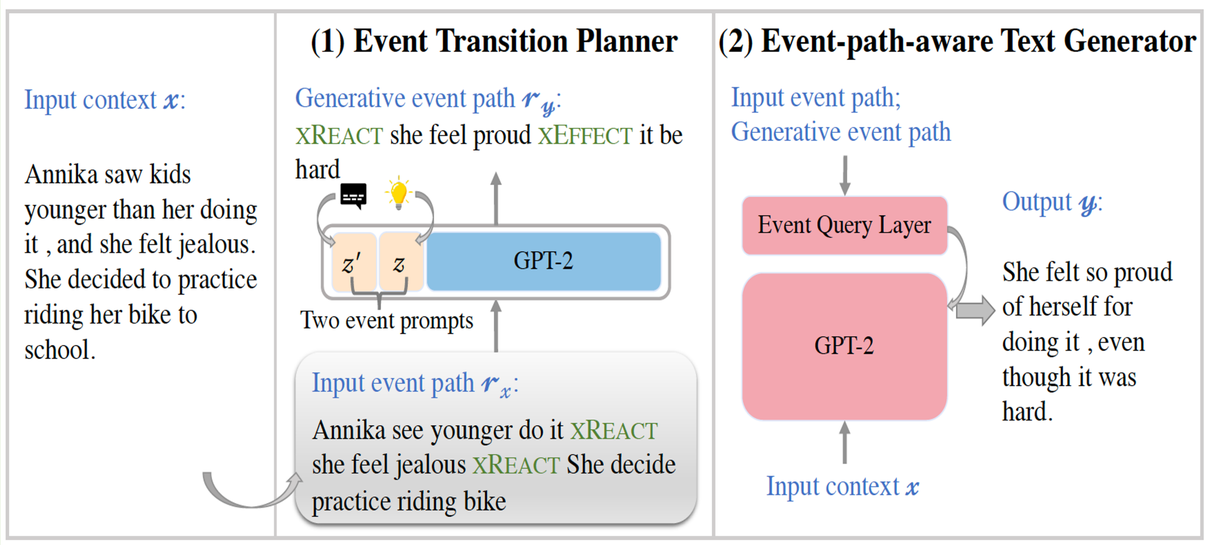
Open-ended text generation tasks, such as dialogue generation and story completion, require models to generate a coherent continuation given limited preceding context. The open-ended nature of these tasks brings new challenges to the neural auto-regressive text generators nowadays. Despite these neural models are good at producing human-like text, it is difficult for them to arrange causalities and relations between given facts and possible ensuing events.

To bridge this gap, we propose a novel two-stage method which explicitly arranges the ensuing events in open-ended text generation. Our approach can be understood as a specially-trained coarse-to-fine algorithm, where an event transition planner provides a ``coarse'' plot skeleton and a text generator in the second stage refines the skeleton.

- Specifically, in stage one, **an event transition planner (EP)** outlines a transition path of events starting from the ones extracted from the input context.
- In stage two, this path is used to ensure a relevant and sound continuation from an **event-path-aware text generator (PG)**.



Our Method



Event Transition Planner

The planer should produce high-quality and diverse paths that can generalize well to the unseen events at test time. For this challenge, we fine-tune a GPT-2 [1] on a large amount of event paths extracted from commonsense graphs ATOMIC [2], as well as from the training set of the specific task, aiming to extrapolate to event sequences that never appeared in these sources with the help of general knowledge stored in the large pre-trained model.

1. We prefix-tune[3] a GPT-2 on a large amount of event paths extracted from commonsense graphs ATOMIC **[z of Planner]**.
2. Then we prefix-tune on training set of the specific task **[z' of Planner]**.

Event-path-aware Text Generator

The auto-regressive text generator need to work effectively under the supervision of the even transition path. We thus design an event query layer to absorb information from the planned paths and use the query layer to guide the text generation process.

1. Another GPT-2 is fine-tuned on specific downstream dataset. **[Transformer parameters of Generator]**
2. Work effectively under the supervision of the even transition path. **[Event query layer of Generator]**

Experiments

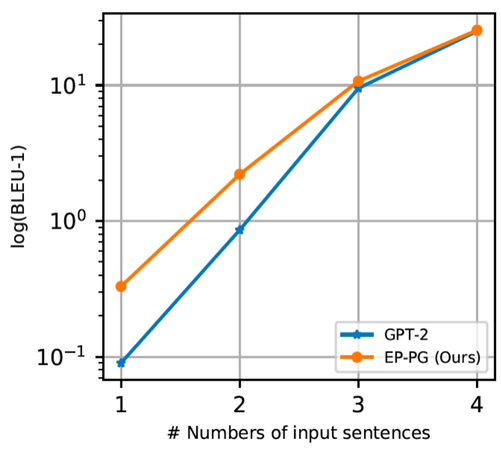
RQ1: How to develop a better event transition planner?

Tasks	Methods	BELU-1	BLEU-2	BLEU-4	DIST-1	DIST-2
Dialogue Generation	GPT-2	23.43	11.50	3.31	1.57	4.18
	PLANGENERATION (Ours)	**26.52**	**12.38**	3.29	**1.88**	**5.52**
	w/o PROMPT	23.58	11.85	**3.58**	1.80	5.13
	w/o TUNING ON ATOMIC	19.82	7.90	1.81	1.16	2.54
	PLANRETRIEVAL	0.75	0.14	0.00	13.05	39.52
Story Completion	GPT-2	15.98	7.19	1.08	5.53	17.44
	PLANGENERATION (Ours)	**19.51**	**9.01**	**1.35**	**5.83**	**17.48**
	w/o PROMPT	13.64	6.14	1.12	4.71	15.77
	w/o TUNING ON ATOMIC	12.74	4.61	0.47	6.08	12.27
	PLANRETRIEVAL	1.28	0.15	0.00	11.88	37.70

RQ2: Whether the integration of event transition paths enhances the open-ended text generation?

Tasks	Models	BLEU-1	BLEU-2	BLEU-4	DIST-1	DIST-2
Dialogue Generation	GPT-2	16.07	6.41	2.13	2.06	7.70
	GPT-2-CS-FT (Guan et al.)	16.43	6.83	2.31	2.16	8.28
	R-EP-PG	16.68	6.71	2.27	**2.21**	**8.44**
	EP-PG (Ours)	**16.74**	**6.94**	**2.39**	2.19	8.25
Story Completion	GPT-2	25.03	9.58	2.70	8.38	31.33
	GPT-2-CS-FT (Guan et al.)	25.09	9.64	2.72	8.07	30.68
	R-EP-PG	24.72	9.27	2.63	7.01	26.49
	EP-PG (Ours)	**25.47**	**9.71**	**2.74**	**8.99**	**34.48**

RQ3: How do the event transition paths benefit text generation?



Story Context:
Alex was in training to be a police officer.
He was not in the best shape.
Alex failed the physical assessment.
Alex started working out.

Golden Event Path:
XEFFECT he take the test again XEFFECT he pass
Retrieved Event Path:
wants to be best police officer XWANT tells person to stop
Generated Event Path:
XEFFECT Alex able get good shape XEFFECT Alex able pass physical test

Reference:
He took the test again and passed .
GPT-2:
Alex was able to get a good job.
GPT-2-CS-FT:
Alex made the squad.
R-EP-PG:
Alex was able to become a police officer.
EP-PG:
Alex was able to pass the physical exam.

Conclusion

- We design a coarse-to-fine framework:
  - a special-trained **event transition planner** to explicitly arrange the ensuing events;
  - an **event-path-aware text generator** to exploit the event transition guidance for language generation.
- We investigate two open-ended text generation tasks, i.e., story completion and dialogue generation.
- Explicit arrangement of event transition path facilitates models to generate more coherent and diverse text in open-ended scenery.
- Our method could be extended to any other language models and open-ended generation tasks.

Code

Paper

References

[1] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners. OpenAI blog, 1(8):9.

[2] Maarten Sap, Ronan Le Bras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A. Smith, and Yejin Choi. 2019. ATOMIC: an atlas of machine commonsense for if-then reasoning. In AAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019, pages 3027–3035. AAAI Press.

[3] Xiang Lisa Li and Percy Liang. 2021. Prefix-tuning: Optimizing continuous prompts for generation. In ACL 2021, pages 4582–4597, Online. Association for Computational Linguistics.