

GROVE: A Retrieval-augmented Complex Story Generation Framework with A Forest of Evidence

Zhihua Wen, Zhiliang Tian*, Wei Wu, Yuxin Yang, Yanqi Shi,
Zhen Huang, Dongsheng Li*

College of Computer, National University of Defense Technology, Hunan, China
{zhwen, tianzhiliang, weiwu_2568,
yangyuxin21a, yqshi, huangzhen, dsli}@nudt.edu.cn

Abstract

Conditional story generation is significant in human-machine interaction, particularly in producing stories with complex plots. While Large language models (LLMs) perform well on multiple NLP tasks, including story generation, it is challenging to generate stories with both complex and creative plots. Existing methods often rely on detailed prompts to guide LLMs to meet target conditions, which inadvertently restrict the creative potential of the generated stories. We argue that leveraging information from exemplary human-written stories facilitates generating more diverse plotlines. Delving deeper into story details helps build complex and credible plots. In this paper, we propose a retrieval-augmented story generation framework with a forest of evidence (GROVE) to enhance stories' complexity. We build a retrieval repository for target conditions to produce few-shot examples to prompt LLMs. Additionally, we design an "asking-why" prompting scheme that extracts a forest of evidence, providing compensation for the ambiguities that may occur in the generated story. This iterative process uncovers underlying story backgrounds. Finally, we select the most fitting chains of evidence from the evidence forest and integrate them into the generated story, thereby enhancing the narrative's complexity and credibility. Experimental results and numerous examples verify the effectiveness of our method.

1 Introduction

Conditional automatic storytelling, generating a story that satisfies specific target conditions, has gained significant attention in the natural language processing community (Kumar, 2023). Generating stories with complex plots is particularly crucial as it creates engaging stories of human-level quality for various applications, such as AI novelists and AI playwrights (Alhussain and Azmi, 2021).

Story generation is an active research area where existing studies approach it from two directions: enhancing controllability and incorporating commonsense knowledge (Alabdulkarim et al., 2021). To satisfy target constraints, researchers enhance the controllability of generation models (Zhou et al., 2023a). Rashkin et al. (2020) follow an outline of the plots to generate stories. Wang et al. (2022b) propose a BART-based (Lewis et al., 2020) model to generate stories according to the fine-grained personalized guidance. Additionally, to produce fluent and coherent storylines, researchers investigate incorporating commonsense knowledge into generation (Wang et al., 2020a; Guan et al., 2020; Zhang et al., 2023). Peng et al. (2022) introduce commonsense inference into GPT-2-based (Radford et al., 2019) model to improve narrative coherence. Qin and Zhao (2022) combine knowledge retrieval, knowledge selection, and story generation together to make the generated story more reasonable. The above studies focus on improving controllability and logical coherence but rarely explore the generation of stories with complex plots.

Large Language Models (LLMs) learn commonsense knowledge from massive texts and develop strong abilities to follow human instructions (Ouyang et al., 2022; OpenAI, 2023; Taori et al., 2023). Thus, LLM-based prompt learning generates fluent and coherent stories with high controllability (Lu et al., 2023; Xie et al., 2023; Yao et al., 2023). Lu et al. (2023) prompt GPT-3 (Brown et al., 2020) with combinations of multiple target conditions. Xie et al. (2023) demonstrate that by using prompts, GPT-3 generates higher-quality stories than other state-of-the-art (SOTA) models. Typically, LLMs generate stories based on a given prompt (i.e. text spans or a few sentences) and the outputs are continuations of the given texts. However, a recurring issue emerges in the LLM-based prompting approaches to generate complex stories: there is a struggle to balance

*Corresponding Authors.