論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Towards Faithful Logical Natural Language Generation from Structured Data
著者(和文)	LIU AO
Author(English)	Ao Liu
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12583号, 授与年月日:2023年9月22日, 学位の種別:課程博士, 審査員:岡崎 直観,篠田 浩一,徳永 健伸,林 晋平,宮崎 純
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12583号, Conferred date:2023/9/22, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

（博士課程）
Doctoral Program

論 文 要 旨

THESIS SUMMARY

系・コース： Department of, Graduate major in	情報工学 系 知能情報 コース	申請学位（専攻分野）： Academic Degree Requested	博士 （工学） Doctor of
学生氏名： Student's Name	LIU AO	審査員主査： Chief Examiner	岡崎直観　教授

要旨（英文 800 語程度）
Thesis Summary（approx.800 English Words ）

Natural Language Generation (NLG) is the process of generating natural language outputs from certain inputs, which is a vital component of human-machine interfaces, such as modern Artificial Intelligence (AI) assistants. A subfield of NLG is data-to-text, which aims to enhance the accessibility of complex data for humans by generating understandable natural language descriptions from structured tables. Traditionally, most methodologies have focused on the generation of surface-level descriptions of superficial facts explicitly present in structured data. However, a recent trend, known as logical natural language generation (logical NLG), is aiming to describe inferred facts in the data, which requires logical reasoning over structure data.

While deep learning models, particularly large-scale pretrained language models, have shown great promise in terms of fluency due to the wealth of contextual knowledge learned from large-scale corpora, they fail to address a critical aspect of logical NLG: faithfulness. Faithfulness, in this context, refers to the factual consistency of the generated content with the given data. Logical NLG poses challenges for maintaining faithfulness as it requires logical reasoning during content planning and non-trivial surface realization to transform an inferred fact into natural language. The joint learning of reasoning and generation is challenging for end-to-end neural models. The scarcity of logical NLG data exacerbates this difficulty.

This dissertation proposes two novel approaches to enhance the faithfulness of logical NLG from structured tables by leveraging logical forms (LFs) as additional resources to address the primary challenges of logical NLG. The first methodology treats LFs as explicit control features for generating the target textual description. The LF control allows the model to primarily concentrate on surface realization, as the LF provides a complete logical fact for description, thereby reducing the need for logical reasoning over the table. However, training a model with explicit LF control requires large-scale parallel data of LFs and texts, which is limited due to the labor-intensive human annotation required. To address the issue of data scarcity, this proposed approach aims to augment parallel LF-to-text data based on the abundant information in tables. The method consists of two stages for data augmentation. The first part generates additional LFs and texts from existing tables using pretrained table-to-text and table-to-LF models, incorporating a pre-defined logic type as a prompt to foster topic-diversified augmented data. Because these LFs and texts are independently augmented and not aligned, the second part employs a semi-supervised dual learning framework for creating pseudo-parallel data from the new LFs and texts, thereby improving the training of the NLG model. Experimental results have proven the effectiveness of this approach.

The second approach, in contrast, addresses table-to-text generation without explicit LF control, thus requiring the model to engage in logical reasoning. This approach only augments LFs as pretraining resources for task transfer learning, proposing a pretraining task for generating correct LFs from a table. The model is first pretrained on this task and then finetuned on the logical NLG task. LFs represent unambiguous semantics, enabling models to learn more reliable logical reasoning from LF generation and transfer the knowledge to downstream NLG tasks. Furthermore, a rule-based sampling method is proposed to collect a large amount of LFs for pretraining to alleviate the data scarcity problem. To better assess logical NLG, we introduce a controlled logical NLG benchmark with annotated highlighted table cells serving as control features. Several pretrained language models were selected as the basis to evaluate our method, and experimental results have demonstrated that the pretraining improves the models' logical reasoning abilities over tables and enhances their faithfulness on two logical NLG benchmarks.

We also conducted a comparative analysis of the difficulties associated with using different model interfaces and control features of logical NLG. Experimental results showed that using explicit LFs eased the generation and led to the best performance of the same model. The no-control interface with only a table as the input is the most challenging, where the model suffers from uncontrollability and low fidelity. The interface with highlighted cells offers moderate controllability and difficulty. The two proposed approaches offer different perspectives on the utilization of LFs, thus adopting different interfaces. The first approach adopts the interface with explicit LFs, focusing on the surface realization part of logical NLG. The second approach, on the other hand, utilizes LFs implicitly and adopts the interface with highlighted cells or only tables, tackling a more challenging task setting of logical NLG. Unlike the first approach, it jointly solves logical reasoning and surface realization, with a focus on the former step.