Bridging the Structural Gap Between Encoding and Decoding for Data-To-Text Generation

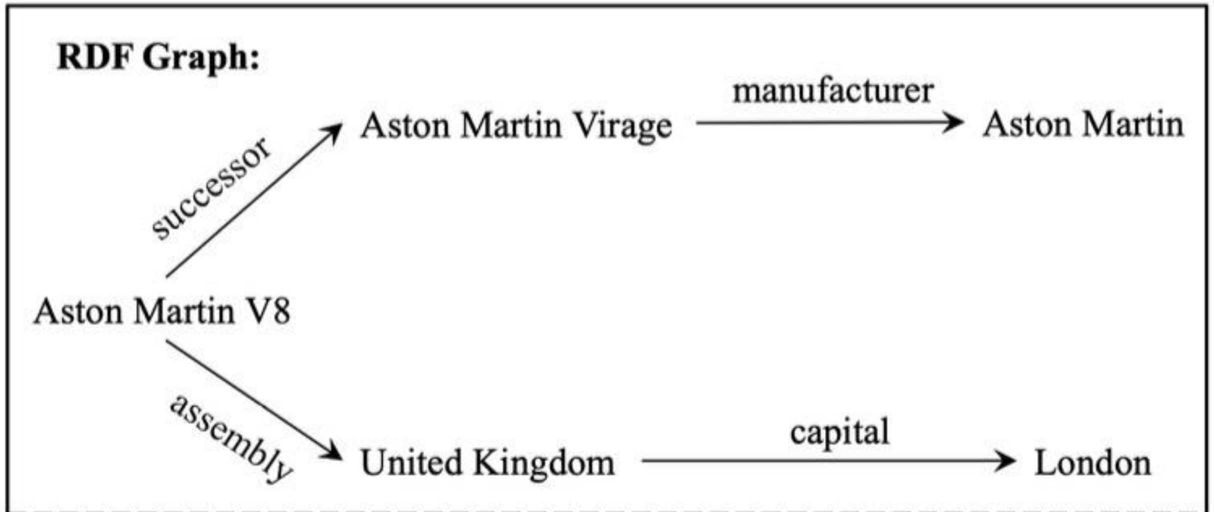
Chao Zhao[†], **Marilyn Walker**[‡] **and Snigdha Chaturvedi**[†]
[†] Department of Computer Science, University of North Carolina at Chapel Hill
[‡] Natural Language and Dialog Systems Lab, University of California, Santa Cruz
{zhaochao, snigdha}@cs.unc.edu mawalker@ucsc.edu

Introduction

- During generation, the structure of the data as well as the content inside the structure jointly determine the generated text.

- these architectures broaden the structural gap between the encoder and decoder. That is, while the encoder receives the input data as a graph, the decoder has to create the output text as a linear chain structure.

- we present a dual encoding model that is not only aware of the input graph structure but also incorporates a content planning stage.
- To encode the structural information in the input graph, we use a GCN based graph encoder
- The plan is then encoded by an LSTM based sequential encoder.



Text: The Aston Martin V8 is assembled in the U.K which has the capital city of London. It was succeeded by the Aston Martin Virage which was made by the company Aston Martin.

Figure 1: Illustration of the WebNLG challenge: the source data is an RDF graph and the target output is a text description of the graph.

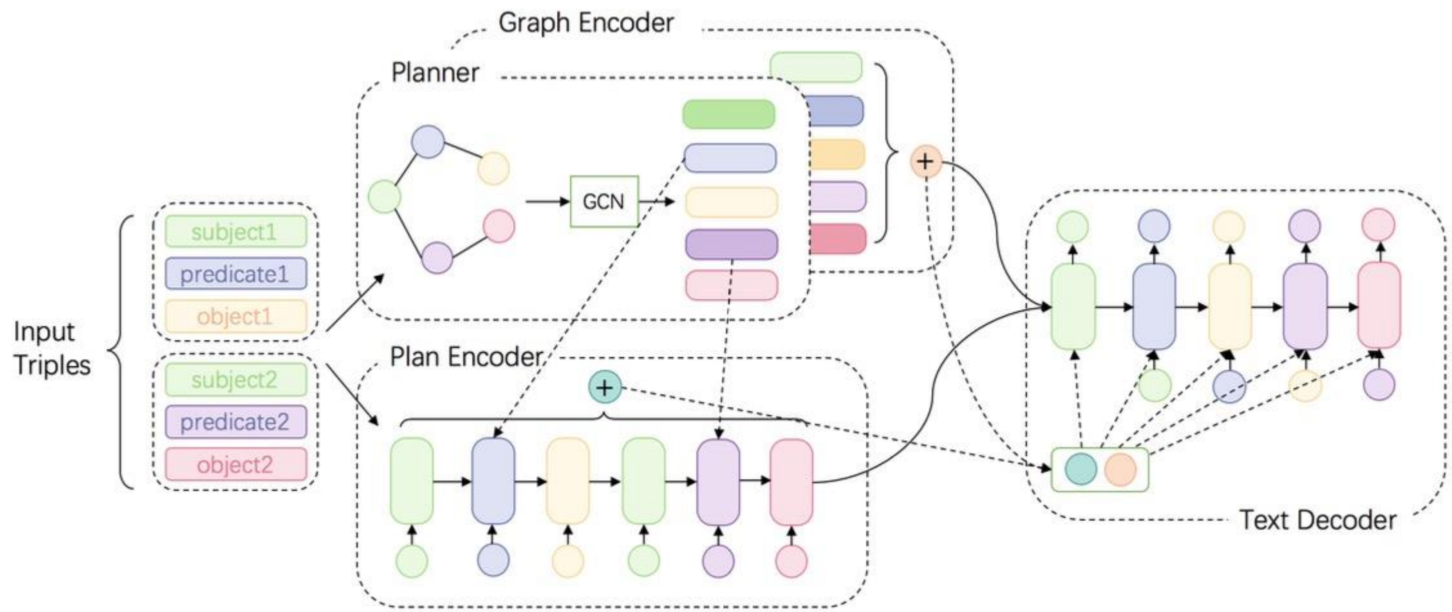


Figure 2: The architecture of the proposed DUALENC model. The input triples are converted as a graph and then fed to two GCN encoders for plan and text generation (Planner and Graph Encoder, top center). The plan is then encoded by an LSTM network (Plan Encoder, bottom center). Finally an LSTM decoder combines the hidden states from both the encoders to generate the text (Text Decoder, middle right).

Graph Representation and Encoding

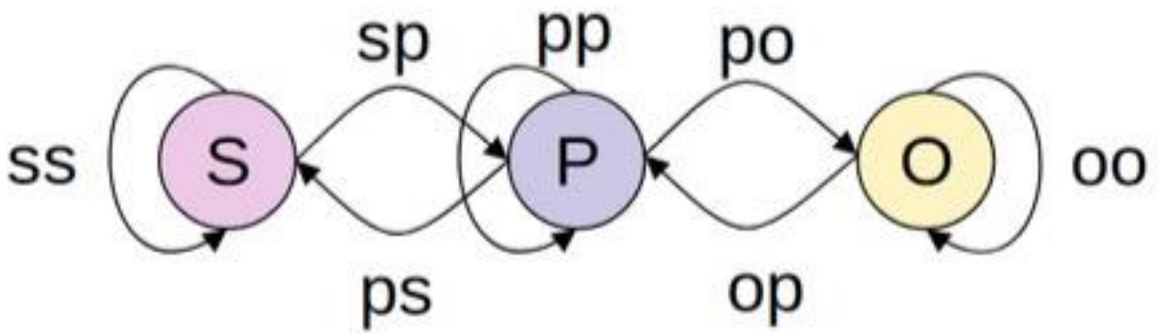


Figure 3: The graph obtained from an RDF triple.

two entities from different triples that have the same mentions will be regarded as the same node

two predicates with the same mentions will be regarded as separate nodes.

Planning Creation and Encoding

- In the planning stage, we determine the *content plan* or order of triples (identified by their predicates) for text realization.

- For example, the content plan for the text in Figure 1 is: “assembly → capital → successor → manufacturer

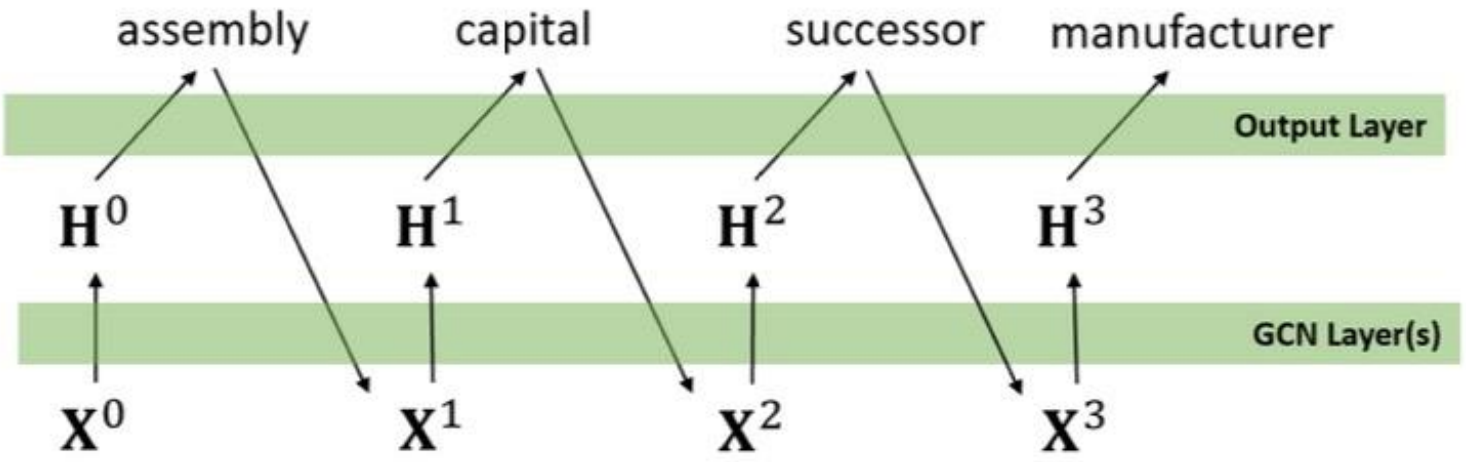


Figure 4: The sequential decision-making process of the planning stage.

$$P(r_i) = \mathrm{softmax}(\mathbf{h}_{r_i}^T \mathbf{W} \bar{\mathbf{h}}_{\mathcal{R}})$$

- After determining an order of input predicates, we complete the plan’s triples by adding the corresponding subjects and objects.

- <S> Aston Martin V8 <P> assembly <O> United King- dom <S> United Kingdom <P> capital <O> London <S> Aston Martin V8 <P> successor <O> Aston Mar- tin Virage <S> Aston Martin Virage <P> manufacturer <O> Aston Martin

-

Planning results

	Accuracy			BLEU-2		
	SEEN	UNSEEN	ALL	SEEN	UNSEEN	ALL
Random	0.28	0.34	0.31	54.1	62.1	57.9
Structure-random	0.32	0.38	0.34	56.6	62.9	59.5
Transformer (Ferreira et al., 2019)	0.56	0.09	0.34	74.3	20.9	49.3
GRU (Ferreira et al., 2019)	0.56	0.10	0.35	75.8	25.4	52.2
Step-By-Step II (Moryossef et al., 2019b)	0.45	0.44	0.44	67.7	67.3	67.5
Step-By-Step (Moryossef et al., 2019a)	0.49	0.44	0.47	73.2	68.0	70.8
GCN	**0.63**	**0.61**	**0.62**	**80.8**	**79.3**	**80.1**

Table 1: Planning results of three test sets evaluated by accuracy and BLEU-2.

	BLEU (↑)			METEOR (↑)			TER (↓)		
	SEEN	UNSEEN	ALL	SEEN	UNSEEN	ALL	SEEN	UNSEEN	ALL
TILB-SMT	54.29	29.88	44.28	0.42	0.33	0.38	0.47	0.61	0.53
ADAPT	60.59	10.53	31.06	0.44	0.19	0.31	0.37	1.40	0.84
MELBOURNE	54.52	33.27	45.13	0.41	0.33	0.37	0.40	0.55	0.47
GTR-LSTM (2018)	54.00	29.20	37.10	0.37	0.28	0.31	0.45	0.60	0.55
GCN-EC (2018)	55.90	-	-	0.39	-	-	0.41	-	-
GRU (2019)	56.09	25.12	42.73	0.42	0.22	0.33	0.39	0.64	0.51
Transformer (2019)	56.28	23.04	42.41	0.42	0.21	0.32	0.39	0.63	0.50
Step-By-Step (2019a)	53.30	34.41	47.24	0.44	0.34	0.39	0.47	0.56	0.51
PLANENC	**64.42**	**38.23**	**52.78**	0.45	0.37	0.41	**0.33**	**0.53**	**0.42**
DUALENC	**63.45**	36.73	51.42	**0.46**	**0.37**	**0.41**	**0.34**	0.55	**0.44**

Table 2: Generation results evaluated by BLEU, METEOR, and TER. We compare our methods with different generation systems (SMT, Sequential NMT, Graph NMT, Pipeline). Both of our methods outperform all the baselines on all three measures. We highlight both results if there is no significant difference.

	Absolute(%)		Pairwise(%)			
	CVGE	FAITH	CVGE	FAITH	FLCY	ALL
MELBOURNE	83.0	75.2	-35.0	-42.5	-38.8	-68.8
STEP	**96.1**	**89.3**	**5.0**	**-3.7**	-45.0	-55.0
E2E-TRANS	85.5	78.0	-21.2	-32.5	-21.2	-46.3
GCN	79.8	76.8	-48.7	-50.0	-26.3	-67.5
PLANENC	**92.3**	**88.2**	**-7.5**	**-12.5**	**-7.5**	-21.2
DUALENC	**94.5**	**91.8**	–	–	–	–

Table 4: Results of human evaluation. DUALENC outperforms most of the baselines on all measures.

Tripleset	(William Anders | birthPlace | British Hong Kong), (William Anders | was a crew member of | Apollo 8), (Apollo 8 | crewMembers | Frank Borman), (Apollo 8 | backup pilot | Buzz Aldrin), (Apollo 8 | operator | NASA), (William Anders | dateOfRetirement | 1969-09-01)
MELBOURNE	william anders (born in british hong kong) was a crew member of **apollo 8' s apollo 8 8** mission along with buzz aldrin as backup pilot and **buzz aldrin on 1969-09-01 .** **[Frank Borman, NASA]**
Step-by-Step	william anders was a crew member of apollo 8 operated by nasa. apollo 8' s backup pilot was buzz aldrin and **frank borman**. william anders was born in british hong kong. william anders retired on september 01st, 1969.
PLANENC	william anders was born in british hong kong and was a crew member of nasa' s apollo 8. frank borman was a crew members of apollo 8 and **he** retired on september 1st, 1969 . **[Buzz Aldrin]**
DUALENC	william anders was born in british hong kong and served as a crew member of nasa' s apollo 8 along with frank borman and backup pilot buzz aldrin. he retired on september 1st, 1969 .
Reference	william anders was born in british hong kong and served as a crew member on apollo 8 along with frank borman. nasa operated apollo 8, where buzz aldrin was a back up pilot. anders retired on sept 1, 1969 .

Table 5: Sample texts generated by our methods and baselines, compared with a human-provided reference. We highlight in different color the **[missing]**, **unfaithful**, and **unfluent** parts of each text. Only the results of our DUALENC correctly mention all the input triples.