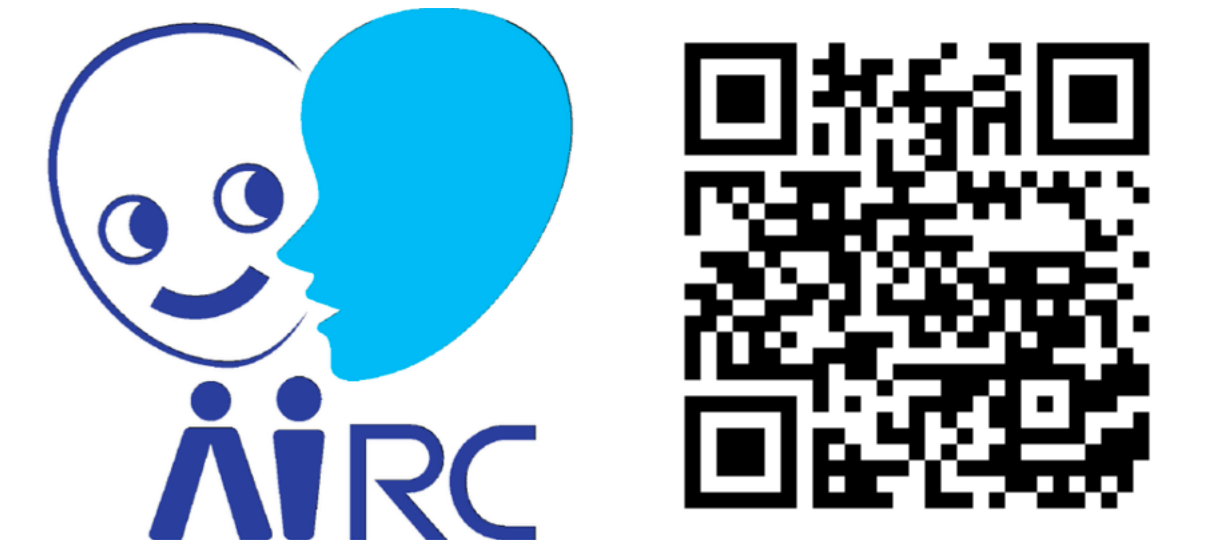


Learning to Select, Track, and Generate for Data-to-Text



Hayate Iso† Yui Uehara♦ Tatsuya Ishigaki♦♦ Hiroshi Noji♦ Eiji Aramaki♦
Ichiro Kobayashi♦♦ Yusuke Miyao♦♦ Naoaki Okazaki♦♦ Hiroya Takamura♦♦

†NAIST ♦Artificial Intelligence Research Center, AIST ♦TITech ♥Ochanomizu Univ. ♣UTokyo



Summary

- ✓ Tracking entity states while summarizing table information improves adequacy while preserving fluency
- ✓ Incorporating the writer information is also helpful for selecting appropriate data and writing better summary

Data-to-Text Generation

Task: Given a set of records, the system should produce a **fluent** and an **adequate** summary.

Problem:

- ▶ Template system could generate the **adequate** text but not **fluent**
- ▶ NN based system could generate the **fluent** text but not **adequate**

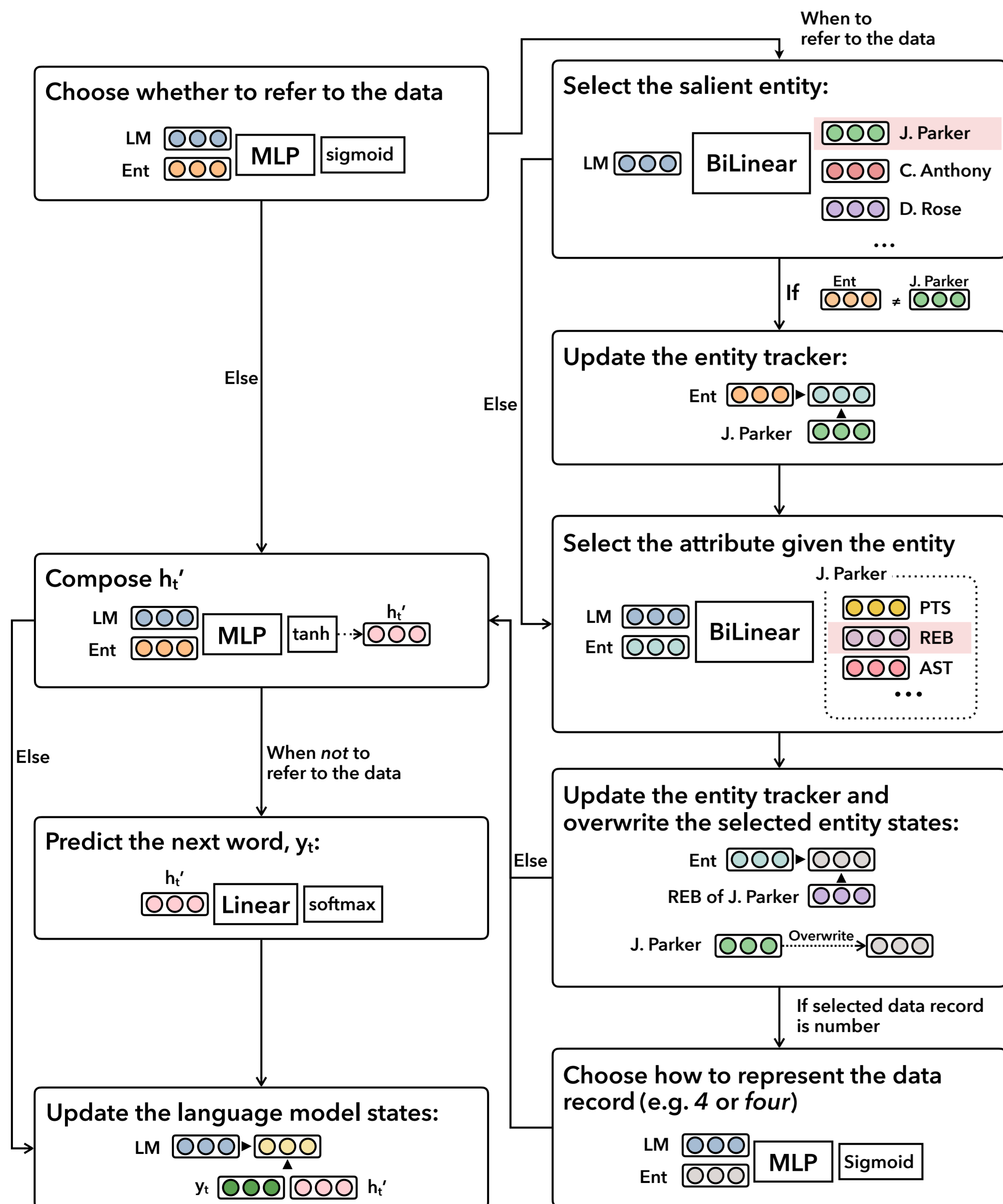
	Adequacy	Fluency
Template	✓	✗
NN	✗	✓
Ours	✓	✓

- ▶ For document-level generation, the system has an influence of the writer's bias.

Entity Tracking

Idea: Equipping an NN-based generation system with the **entity-tracking module** enables capturing of the **saliency** and the **coherence** while preserving the **fluency**.

Method: Inspired by EntityNLM [1], our model dynamically updates each entity state and the tracking module.



Annotation: To train the model in a fully supervised way, we obtain the following annotations using the information extraction system developed by Wiseman et al. [2].

t	199	200	201	202	203	204	205	206	207	208	209
Y_t	Jabari	Parker	contributed	15	points	.	four	rebounds	.	three	assists
Z_t	1	1	0	1	0	0	1	0	0	1	0
E_t	JABARI PARKER	JABARI PARKER		JABARI PARKER			JABARI PARKER			JABARI PARKER	
A_t	FIRST NAME	LAST NAME		PLAYER PTS			PLAYER REB			PLAYER AST	
N_t				0			1			1	

TEAM	H/V	WIN	LOSS	PTS	REB	AST	FG_PCT	FG3_PCT	...
KNICKS	H	16	19	104	46	26	45	46	...
BUCKS	V	18	16	105	42	20	47	32	...

PLAYER	H/V	PTS	REB	AST	BLK	STL	MIN	CITY	...
CARMELO ANTHONY	H	30	11	7	0	2	37	NEW YORK	...
DERRICK ROSE	H	15	3	4	0	1	33	NEW YORK	...
COURTNEY LEE	H	11	2	3	1	1	38	NEW YORK	...
GIANNIS ANTETOKOUNMPO	V	27	13	4	3	1	39	MILWAUKEE	...
GREG MONROE	V	18	9	4	1	3	31	MILWAUKEE	...
JABARI PARKER	V	15	4	3	0	1	37	MILWAUKEE	...
MALCOLM BROGDON	V	12	6	8	0	0	38	MILWAUKEE	...
MIRZA TELETOVIC	V	13	1	0	0	0	21	MILWAUKEE	...
JOHN HENSON	V	2	2	0	0	0	14	MILWAUKEE	...

The Milwaukee Bucks defeated the New York Knicks, 105-104, at Madison Square Garden on Wednesday. The Knicks (16-19) checked in to Wednesday's contest looking to snap a five-game losing streak and heading into the fourth quarter, they looked like they were well on their way to that goal. ... Antetokounmpo led the Bucks with 27 points, 13 rebounds, four assists, a steal and three blocks, his second consecutive double-double. Greg Monroe actually checked in as the second-leading scorer and did so in his customary bench role, posting 18 points, along with nine boards, four assists, three steals and a block. Jabari Parker contributed 15 points, four rebounds, three assists and a steal. Malcolm Brogdon went for 12 points, eight assists and six rebounds. Mirza Teletovic was productive in a reserve role as well, generating 13 points and a rebound. ... Courtney Lee checked in with 11 points, three assists, two rebounds, a steal and a block. ... The Bucks and Knicks face off once again in the second game of the home-and-home series, with the meeting taking place Friday night in Milwaukee.

(a) Box-score

(b) Gold summary

Generated summary

The Milwaukee Bucks defeated the New York Knicks, 105-104, at Madison Square Garden on Wednesday evening. The Bucks (18-16) have been one of the hottest teams in the league, having won five of their last six games, and they have now won six of their last eight games. The Knicks (16-19) have now won six of their last six games, as they continue to battle for the eighth and final playoff spot in the Eastern Conference. Giannis Antetokounmpo led the way for Milwaukee, as he tallied 27 points, 13 rebounds, four assists, three blocked shots and one steal, in 39 minutes. Jabari Parker added 15 points, four rebounds, three assists, one steal and one block, and 6-of-8 from long range. John Henson added two points, two rebounds, one assist, three steals and one block. John Henson was the only other player to score in double digits for the Knicks, with 15 points, four assists, three rebounds and one steal, in 33 minutes. The Bucks were led by Derrick Rose, who tallied 15 points, four assists, three rebounds and one steal in 33 minutes. Willy Hernangomez started in place of Porzingis and finished with 15 points, four assists, three rebounds and one steal in 33 minutes. Willy Hernangomez started in place of Jose Calderon (knee) and responded with one rebound and one block. The Knicks were led by their starting backcourt of Carmelo Anthony and Carmelo Anthony, but combined for just 13 points on 5-of-16 shooting. The Bucks next head to Philadelphia to take on the Sixers on Friday night, while the Knicks remain home to face the Los Angeles Clippers on Wednesday.

(a) Puduppully et al. (2019)

The Milwaukee Bucks defeated the New York Knicks, 105-104, at Madison Square Garden on Saturday. The Bucks (18-16) checked in to Saturday's contest with a well, outscoring the Knicks (16-19) by a margin of 39-19 in the first quarter. However, New York by just a 25-foot lead at the end of the first quarter, the Bucks were able to pull away, as they outscored the Knicks by a 59-46 margin into the second. 45 points in the third quarter to seal the win for New York with the rest of the starters to seal the win. The Knicks were led by Giannis Antetokounmpo, who tallied a game-high 27 points, to go along with 13 rebounds, four assists, three blocks and a steal. The game was a crucial night for the Bucks' starting five, as the duo was the most effective shooters, as they posted Milwaukee to go on a pair of low low-wise (Carmelo Anthony) and Malcolm Brogdon. Anthony added 11 rebounds, seven assists and two steals to his team-high scoring total. Jabari Parker was right behind him with 15 points, four rebounds, three assists and a block. Greg Monroe was next with a bench-leading 18 points, along with nine rebounds, four assists and three steals. Brogdon posted 12 points, eight assists, six rebounds and a steal. Derrick Rose and Courtney Lee were next with a pair of {11 / 11} -point efforts. Rose also supplied four assists and three rebounds, while Lee complemented his scoring with three assists, a rebound and a steal. John Henson and Mirza Teletovic were next with a pair of {two / two} -point efforts. Teletovic also registered 13 points, and he added a rebound and an assist. Jason Terry supplied eight points, three rebounds and a pair of steals. The Cavs remain in last place in the Eastern Conference's Atlantic Division. They now head home to face the Toronto Raptors on Saturday night.

(b) Our model

- ✓ Fluent text
- ✗ Many incorrect realizations (orange-colored)
- ✗ Many improper content selections (blue-colored)
- ✓ Fluent text
- ✓ Correct realization & appropriate content selection
- ✗ Difficult to capture complex relation (underlined)

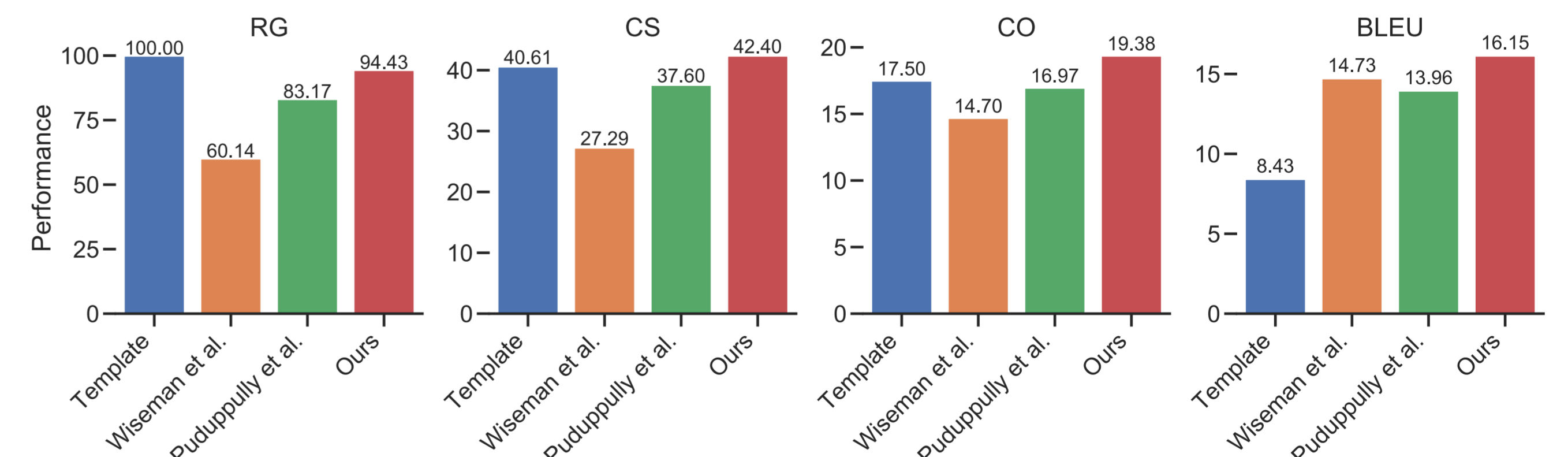
Experimental results

Baseline:

- ▶ Template: game results, 6 prominent players, next game information
 - ▶ Wiseman et al. [2]: Encoder-Decoder w/ Attention
 - ▶ Puduppully et al. [3]: Two stage model
1. Predicting the sequence of data records
 2. Generating a summary conditioned on the predicted sequence

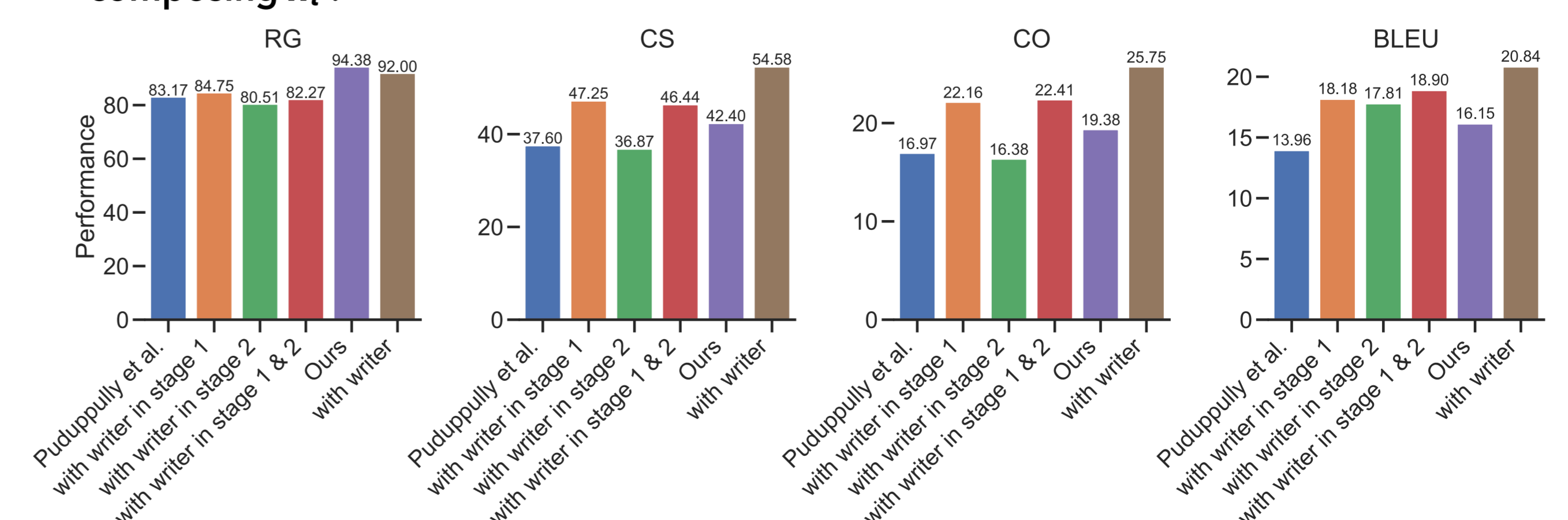
Evaluation Metrics:

- ▶ RG: the ratio of the correct relations out of all the extracted relations.
- ▶ CS: the F1 score of the relations extracted from the generated summary against those from the reference summary
- ▶ CO: the normalized Damerau-Levenshtein Distance (DLD) between the sequences of relations extracted from the generated and reference summary.
- ▶ BLEU



Writer Bias Evaluation:

- ▶ For Puduppully et al., concatenate the writer embedding w for decoder's input.
- ▶ For our model, concatenate w for hidden state of LM and Tracking module states for composing h_t' .



References

1. Ji et al. - Dynamic Entity Representations in Neural Language Models, In EMNLP 2017
2. Wiseman et al. - Challenges in Data-to-Document Generation, In EMNLP 2017
3. Puduppully et al. - Data-to-Text Generation with Content Selection and Planning, In AAAI 2019