Some Research on Table-to-Text Generation

Heng Gong

Tencent AI Lab, NLP Center

Overview

- Data-to-Text Generation with Content Selection and Planning (AAAI 2019)
- Data-to-text Generation with Entity Modeling (ACL 2019)
- Learning to Select, Track, and Generate for Data-to-Text (ACL 2019)

Task

- Table-to-Text Generation
  - Input：Table (tuples e.g. <PLAYER, POINTS, 11>)
  - Output：Descriptive text

Weather Forecast Report

NBA Game Report

WIKIBIO

Team	POINTS	WINS	LOSSES	…
Wizards	88	31	18	…
Hornets	92	21	27	

Player	PTS	AST	REB	…
Wizards				
Paul Pierce	11	1	3	…
Nene	8	1	7	…
Bradley Beal	18	1	11	…
John Wall	16	10	1	…
…	…	…	…	…
Kris Humphries	13	1	5	…
Hornets				
Michael Kidd-Gilchrist	13	3	13	…
Al Jefferson	18	1	12	…
Gerald Henderson	17	5	2	…
Brian Roberts	18	3	1	…
…	…	…	…	…
Gary Neal	12	1	0	…

Tables

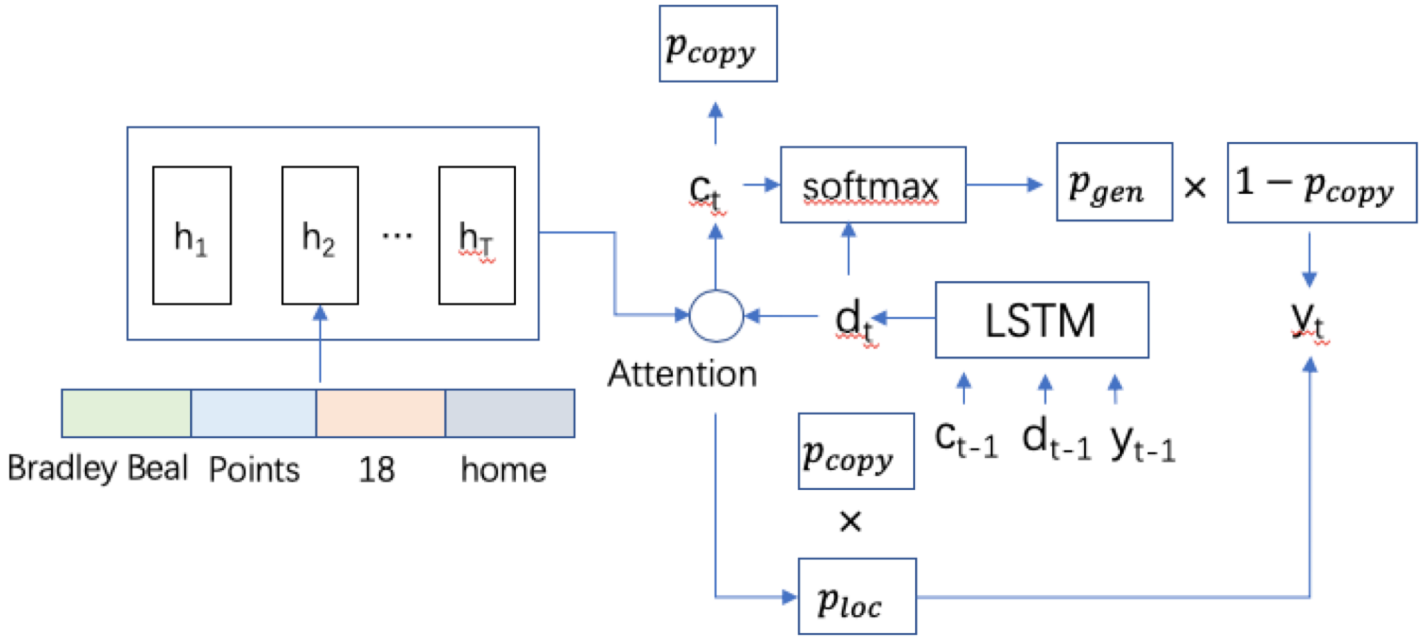
The Charlotte Hornets (21 - 27) defeated the Washington Wizards (31 - 18) 92 - 88 on Monday …The Hornets were led by **Al Jefferson** in this game , who went 9 - for - 19 from the floor to score 18 points ... It was the **second time** in the last three games he 's posted a double - double , while the two steals matched a season - high for the center … **Beal** has turned it on over his last two games , combining for 44 points and 14 rebounds ... This double - double marked the **second in a row** for **Wall** , who 's combined for 44 points and 22 asssists over his last two games …

Gold

	RC	WG	WB	RW	SBN	MLB
Vocab	409	394	400K	11.3K	68.6K	38.9K
Tokens	11K	0.9M	19M	1.6M	8.8M	14.3M
Examples	1.9K	22.1K	728K	4.9K	10.9K	26.3K
Avg Len	5.7	28.7	26.1	337.1	805.4	542.05
Rec. Types	4	10	1.7K	39	39	53
Avg Records	2.2	191	19.7	628	628	565

Base Model

- Standard neural table-to-text model consists of encoder-decoder structure.
  - Encoder: use MLP or RNN(LSTM) to transform records into continuous vectors
  - Decoder: LSTM decoder with attention mechanism.
  - Conditional copy (Gulcehre et al.) mechanism can improve the fidelity of text by copying information from table.



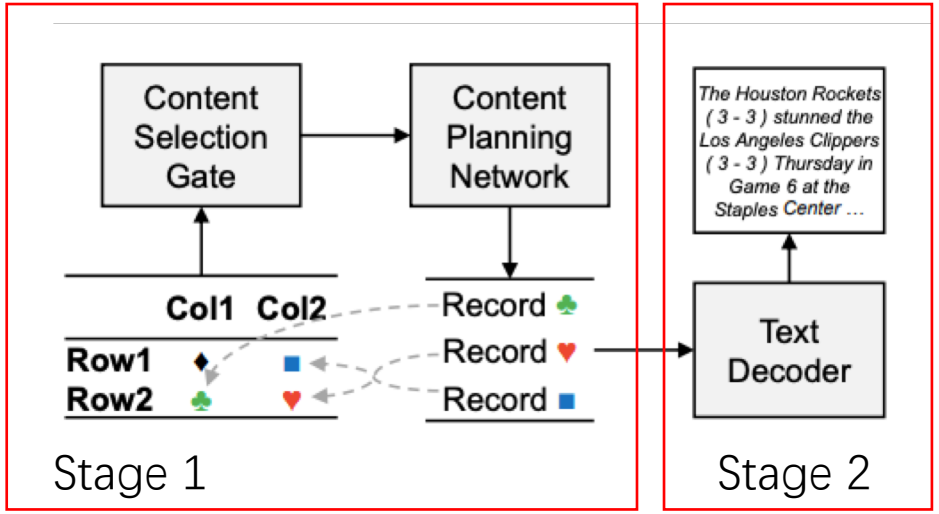
Data-to-Text Generation with Content Selection and Planning

Ratish Puduppully and **Li Dong** and **Mirella Lapata**
Institute for Language, Cognition and Computation
School of Informatics, University of Edinburgh
10 Crichton Street, Edinburgh EH8 9AB
r.puduppully@sms.ed.ac.uk li.dong@ed.ac.uk mlap@inf.ed.ac.uk

AAAI 2019

Motivation & Contribution

- Current **neural** table-to-text models neglect the **explicit** modeling of **what to say** and in **what order**
- It leads to (1) poor performance at **content selection**, (2) lack of a reasonable **ordering** of the selected facts and (3) not being very **faithful** to the input.



- They propose a model that explicitly modeling content selection and planning within a neural data-to-text architecture.
  - Provide high-level organization, allowing the decoder focus on easier tasks of sentence planning and surface realization
  - more interpretable
  - reduces redundancy

Method

1. Content Selection
   - Selection Gate
2. Content Planning
   - Pointer Network

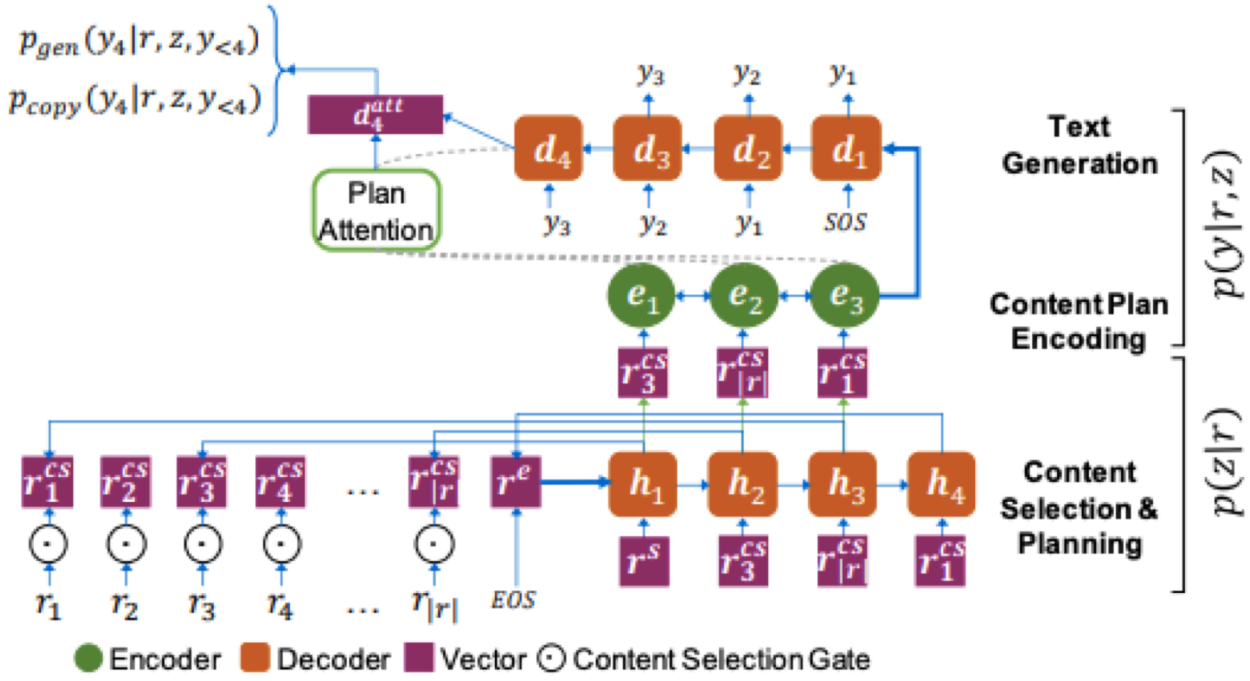
Selection Gate

$$\alpha_{j,k} \propto \exp(\mathbf{r}_j^\intercal \mathbf{W}_a \mathbf{r}_k)$$
$$\mathbf{c}_j = \sum_{k \neq j} \alpha_{j,k} \mathbf{r}_k$$
$$\mathbf{r}_j^{att} = \mathbf{W}_g [\mathbf{r}_j; \mathbf{c}_j]$$

Self attention

$$\mathbf{g}_j = \text{sigmoid}\left(\mathbf{r}_j^{att}\right)$$
$$\mathbf{r}_j^{cs} = \mathbf{g}_j \odot \mathbf{r}_j$$

Gate



Pointer Network (LSTM decoder):

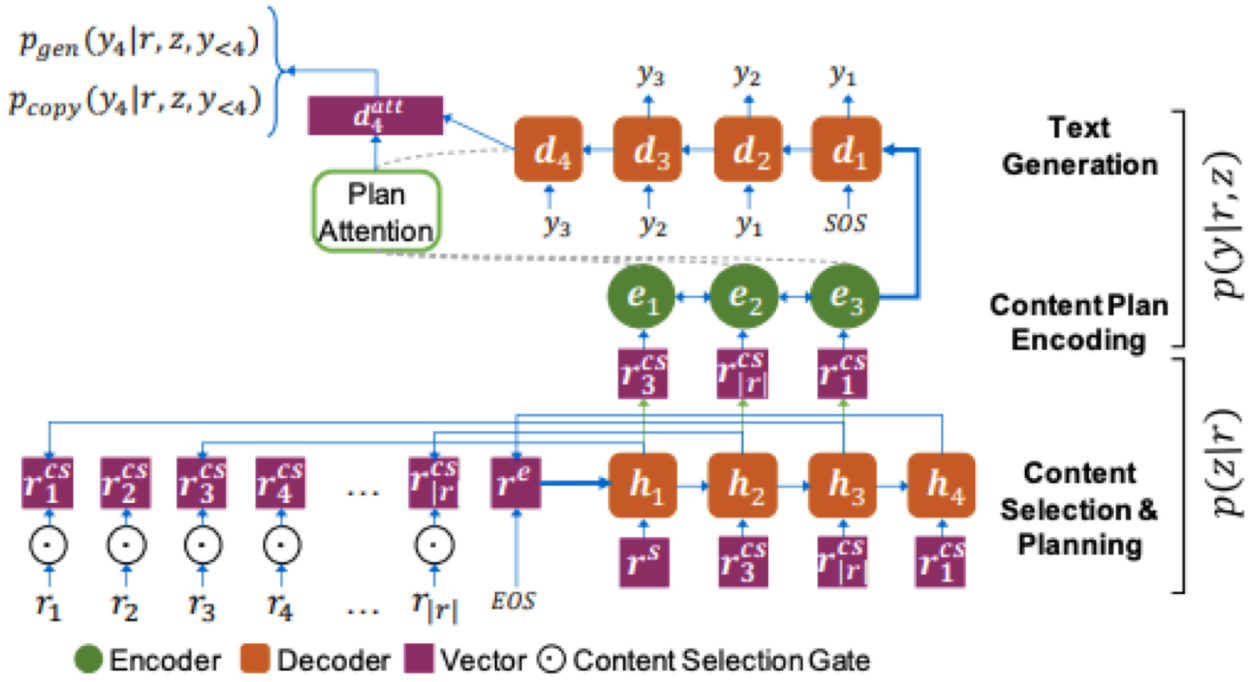
$$p(z_k = r_j | z_{<k}, r) \propto \exp(\mathbf{h}_k^\intercal \mathbf{W}_c \mathbf{r}_j^{cs})$$

Method

3. Text Generation
   - Attention with Conditional Copy

Training: $$\max \sum_{(r,z,y)\in\mathcal{D}} \log p\left(z|r\right) + \log p\left(y|r,z\right)$$

Inference: $$\hat{z} = \arg\max_{z'} p(z'|r)$$
$$\hat{y} = \arg\max_{y'} p(y'|r,\hat{z})$$



Experiments

- Metrics
  - BLEU
  - Extractive Metrics
    - Relation Generation (RG): Fidelity
    - Content Selection (CS)
    - Content Ordering (CO)
- Result
  - Outperform baselines across all
  - NCP+OR shows text generator’ s good performance on surface realization
  - NCP shows the bottleneck of CS&CP

	#Support	#Contra	Gram	Cohere	Concise
Gold	2.98	0.28	11.78	16.00	13.78
TEMPL	6.98	0.21	-0.89	-4.89	1.33
WS-2017	3.19	1.09	-4.22	-4.89	-6.44
NCP+CC	4.90	0.90	-2.44	-2.44	-3.55

Human Evaluation

Model	RG		CS		CO	BLEU
	#	P%	P%	R%	DLD%	
TEMPL	**54.23**	**99.94**	26.99	**58.16**	14.92	8.46
WS-2017	23.72	74.80	29.49	36.18	15.42	14.19
NCP+JC	34.09	87.19	32.02	47.29	17.15	14.89
NCP+CC	34.28	87.47	**34.18**	51.22	**18.58**	**16.50**

Table 5: Automatic evaluation on ROTOWIRE test set using

Model	RG		CS		CO	BLEU
	#	P%	P%	R%	DLD%	
TEMPL	54.29	99.92	26.61	59.16	14.42	8.51
WS-2017	23.95	75.10	28.11	35.86	15.33	14.57
ED+JC	22.98	76.07	27.70	33.29	14.36	13.22
ED+CC	21.94	75.08	27.96	32.71	15.03	13.31
NCP+JC	33.37	87.40	32.20	48.56	17.98	14.92
NCP+CC	33.88	87.51	33.52	51.21	18.57	16.19
NCP+OR	21.59	89.21	88.52	85.84	78.51	24.11

Table 3: Automatic evaluation on ROTOWIRE development

Model	RG		CS		CO	BLEU
	#	P%	P%	R%	DLD%	
ED+CC	21.94	75.08	27.96	32.71	15.03	13.31
CS+CC	24.93	80.55	28.63	35.23	15.12	13.52
CP+CC	33.73	84.85	29.57	44.72	15.84	14.45
NCP+CC	33.88	87.51	33.52	51.21	18.57	16.19
NCP	34.46	—	38.00	53.72	20.27	—

Table 4: Ablation results on ROTOWIRE development set us-

Example

System	Summary
Template	The Washington Wizards (8-13) defeated the Denver Nuggets (8-15) 92-85. Bradley Beal scored 26 points (8-15 FG, 4-7 3PT, 6-6 FT) to go with 3 rebounds. Nikola Jokic scored 17 points (6-10 FG, 0-0 3PT, 5-7 FT) to go with 11 rebounds. Markieff Morris scored 15 points (5-12 FG, 0-0 3PT, 5-5 FT) to go with 3 rebounds. John Wall scored 15 points (5-14 FG, 0-4 3PT, 5-6 FT) to go with 7 rebounds. Danilo Gallinari scored 14 points (3-11 FG, 1-8 3PT, 7-9 FT) to go with 4 rebounds. Jusuf Nurkic scored 13 points (6-6 FG, 0-0 3PT, 1-2 FT) to go with 7 rebounds. The Washington Wizards' next game will be at home against the Dallas Mavericks, while the Denver Nuggets will travel to play the Bulls.
WS-2017	The Washington Wizards (8-13) defeated the Denver Nuggets (8-15) 92-85 on Wednesday at the Verizon Center in Washington. The Wizards got off to a hot start in this one, out-scoring the Nuggets 29-18 in the first quarter alone. The Nuggets were able to out-score the Nuggets 28-12 in the third quarter, while the Wizards were able to coast to a victory in front of their home crowd. The Nuggets were the superior shooters in this game, going 45 percent from the field and 32 percent from the three-point line, while the Wizards went just 41 percent from the floor and 32 percent from beyond the arc. The Nuggets were also able to shoot 41 percent from the field and 32 percent from the three-point line , while the Nuggets went just 41 percent from the floor and 32 percent from beyond the arc. The Nuggets were led by the duo of John Wall and Bradley Beal. Wall went 5-for-14 from the field and 0-for-4 from the three-point line to score a game-high of 15 points, while also adding seven rebounds and five assists. It was his second double-double in a row, as he's combined for 47 points and 12 rebounds over his last two games. The only other Nugget to reach double figures in points was Wilson Chandler, who finished with 11 points and eight rebounds. It was his second double-double in a row, as he's combined for 47 points and 14 rebounds over that span. The only other Nugget to reach double figures in points was Wilson Chandler, who finished with 11 points and eight rebounds. The Nuggets' next game will be on the road against the Detroit Pistons on Friday, while the Nuggets will be at home against the New York Knicks on Friday.
NCP+CC	The Washington Wizards defeated the visiting Denver Nuggets 92-85 at Verizon Center on Monday. The Wizards (8-13) came into this game winners of five of their last eight games, but the Wizards (8-15) jumped out to a 10-point lead at the end of the first quarter. Bradley Beal led the way for the Wizards with a game-high 26 points on 8-of-15 shooting from the field. John Wall shot 5-of-14 from the field on his way to 15 points, to go along with seven rebounds, five assists, three steals and one block. Jusuf Nurkic chipped in 13 points, seven rebounds and one assist. Jameer Nelson filled out the stat sheet with 10 points, eight assists, four rebounds and two steals in 36 minutes. As a team, it was a forgettable shooting night for the Nuggets, as the team shot just 46 percent from the field. Next up, the Nuggets play the second game of a back-to-back when they host the Denver Nuggets on Wednesday, while the Wizards host the Portland Trail Blazers on Friday.

Summary

- They presented a data-to-text generation model which is enhanced with content selection and planning modules.
- Experimental results demonstrate that generation quality improves both in terms of the number of relevant facts contained in the output text (**content fidelity and content selection ability**), and the **order** according to which these are presented.
- Positive side-effects of content planning are additional improvements in the grammaticality, and conciseness of the generated text.

Summary

- Strength
  - Intuitive and effective compared with the strong baselines.
- Weakness
  - Suffers from error propagation.
  - Lack comparison with a COLING 2018 paper on this dataset at that time.

Data-to-text Generation with Entity Modeling

Ratish Puduppully and **Li Dong** and **Mirella Lapata**
Institute for Language, Cognition and Computation
School of Informatics, University of Edinburgh
10 Crichton Street, Edinburgh EH8 9AB
r.puduppully@sms.ed.ac.uk li.dong@ed.ac.uk mlap@inf.ed.ac.uk

ACL 2019

Motivation & Contribution

- Recent neural models treat entities as nothing more than vocabulary tokens.
- Descriptive texts are often characterized as "**entity coherent**" which means that their coherence is based on the way entities are introduced and discussed in the discourse (Karamanis et al., 2004).
- They propose a **entity-centric** neural architecture for data-to-text generation that **dynamically** update entity-specific representations while decoding (discourse in descriptive texts may shift from one entity to the next).
- Introduce a (five times larger) **new dataset**

on the baseball domain.

	ROTOWIRE	MLB
Vocab Size	11.3K	38.9K
# Tokens	1.5M	14.3M
# Instances	4.9K	26.3K
Avg Length	337.1	542.05
# Record Types	39	53
Avg Records	628	565

Method

- Introduce a **memory cell** and a **processor** for each entity. (A)
- Hierarchical attention that focus on entities at first, then on the records corresponding to entities. (B,C)

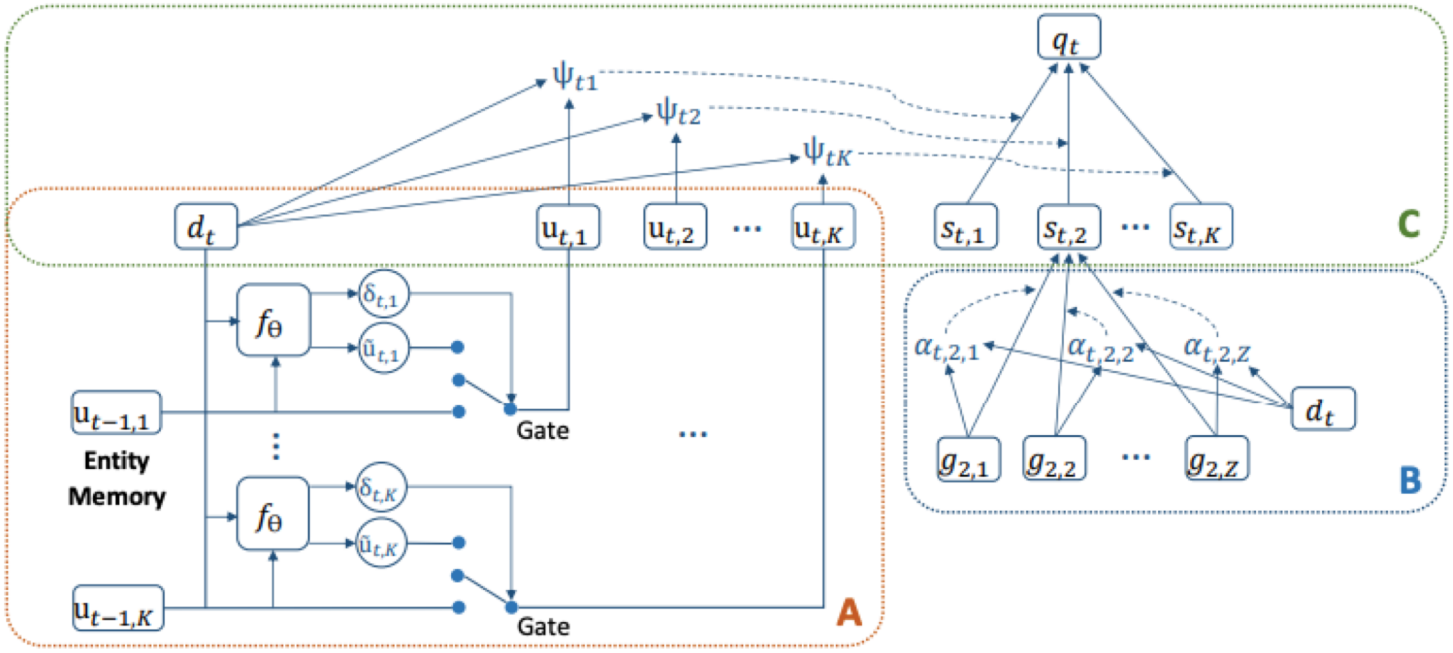


Figure 2: Diagram of entity memory network (block A) and hierarchical attention (blocks B and C). Module f_θ

Method

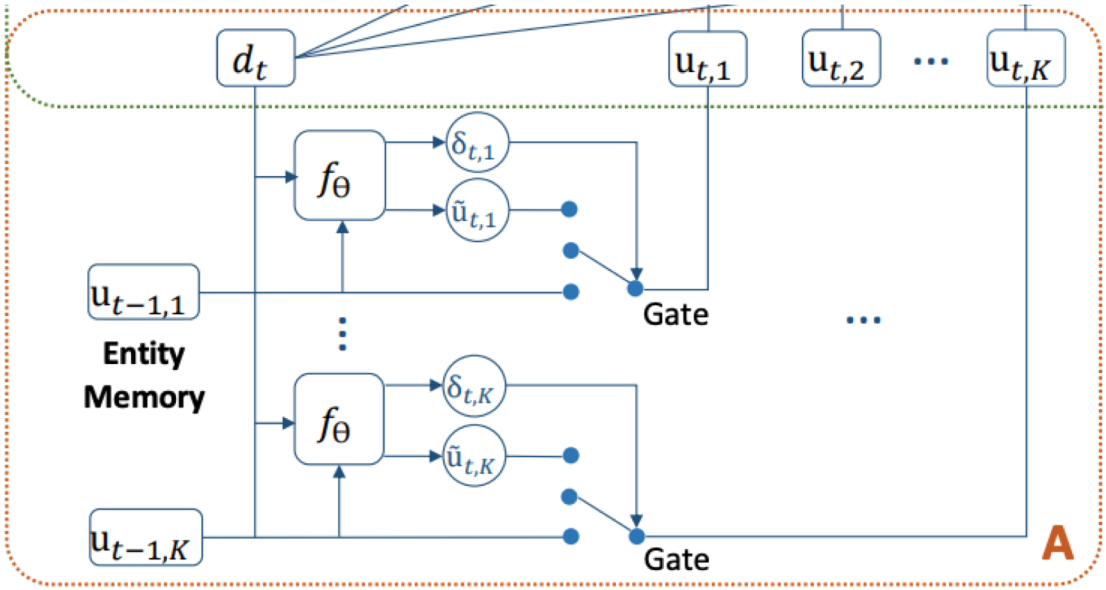
- Initialize entity memory with (linear transformation of) mean pooling over corresponding records.
- It has two update gate
  - Overall entity update gate

$$\boldsymbol{\gamma}_t = \sigma(\mathbf{W}_d \mathbf{d}_t + \mathbf{b}_d)$$

  - Individual entity update gate

$$\boldsymbol{\delta}_{t,k} = \boldsymbol{\gamma}_t \odot \sigma(\mathbf{W}_e \mathbf{d}_t + \mathbf{b}_e + \mathbf{W}_f \mathbf{u}_{t-1,k} + \mathbf{b}_f)$$

- New memory information
  - $\tilde{\mathbf{u}}_{t,k} = \mathbf{W}_g \mathbf{d}_t$
- Updated entity memory
  - $\mathbf{u}_{t,k} = (1 - \boldsymbol{\delta}_{t,k}) \odot \mathbf{u}_{t-1,k} + \boldsymbol{\delta}_{t,k} \odot \tilde{\mathbf{u}}_{t,k}$



Method

- Hierarchical attention
  - Focus on entity at first.

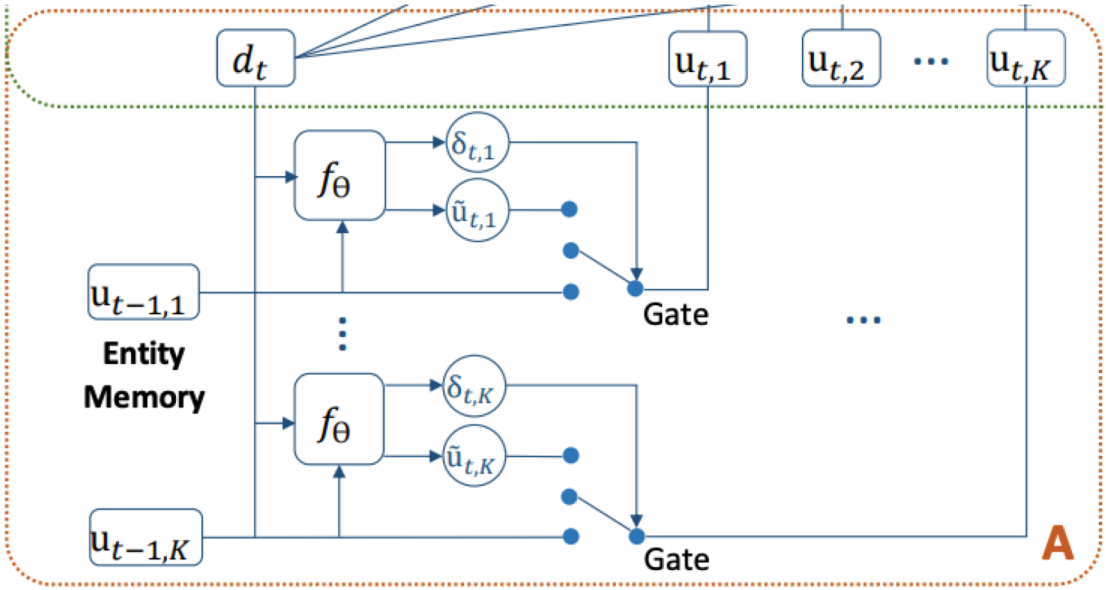
$$\Psi_{t,k} \propto \exp(\mathbf{d}_t^\top \mathbf{W}_h \mathbf{u}_{t,k})$$

  - Constraint on attention to records corresponding to each entity.

$$\alpha_{t,k,z} \propto \exp(\mathbf{d}_t^\top \mathbf{W}_a \mathbf{g}_{k,z})$$

$$\mathbf{q}_t = \sum_k \Psi_{t,k} \sum_z \alpha_{t,k,z} \mathbf{g}_{k,z}$$

Sum to 1



Experiments

- ENT performs better w.r.t content fidelity (RG on ROTOWIRE), content selection ability and content ordering (CO) without planning.
- On MLB, ED+CC tends to focus on one or two players getting most of the facts about them right, whereas ENT sometimes gets the coreference wrong, and thus lower RG precision

RW	RG #	RG P%	CS P%	CS R%	CO DLD%	BLEU
TEMPL	**54.23**	**99.94**	26.99	**58.16**	14.92	8.46
WS-2017	23.72	74.80	29.49	36.18	15.42	14.19
NCP+CC	34.28	87.47	34.18	51.22	18.58	**16.50**
ENT	30.11	92.69	**38.64**	48.51	**20.17**	16.12

MLB	RG #	RG P%	CS P%	CS R%	CO DLD%	BLEU
TEMPL	**59.93**	**97.96**	22.82	**68.46**	10.64	3.81
ED+CC	18.69	92.19	**62.01**	50.12	25.44	9.69
NCP+CC	17.93	88.11	60.48	55.13	**26.71**	9.68
ENT	21.35	88.29	58.35	61.14	24.51	**11.51**

Table 2: Evaluation on ROTOWIRE (RW) and MLB test sets using relation generation (RG) count and

RW	RG #	RG P%	CS P%	CS R%	CO DLD%	BLEU
ED+CC	22.68	79.40	29.96	34.11	16.00	14.00
+Hier	30.76	93.02	33.99	44.79	19.03	14.19
+Dyn	27.93	90.85	34.19	42.27	18.47	15.40
+Gate	31.84	91.97	36.65	48.18	19.68	15.97

MLB	RG #	RG P%	CS P%	CS R%	CO DLD%	BLEU
ED+CC	18.69	92.65	62.29	51.36	25.93	9.55
+Hier	19.02	93.71	62.84	52.12	25.72	10.38
+Dyn	20.28	89.19	58.19	58.94	24.49	10.85
+Gate	21.32	88.16	57.36	61.50	24.87	11.13

Table 3: Ablation results on ROTOWIRE (RW) and MLB development set using relation generation (RG)

Human Evaluation

ROTOWIRE	#Supp	#Contra	Gram	Coher	Concis
Gold	2.98*	0.28*	4.07*	3.33	-10.74*
TEMPL	6.98*	0.21*	-3.70*	-3.33*	17.78*
NCP+CC	4.90	0.90	-3.33*	-3.70*	-3.70
ENT	4.77	0.80	2.96	3.70	-3.33

MLB	#Supp	#Contra	Gram	Coher	Concis
Gold	2.81	0.15*	1.24*	3.48*	-9.33*
TEMPL	3.98*	0.04*	-10.67*	-7.30*	8.43*
ED+CC	3.24*	0.40	0.22*	-0.90*	-2.47*
NCP+CC	2.86	0.88*	0.90*	-1.35*	-1.80*
ENT	2.86	0.52	8.31	6.07	5.39

Example

The **Houston Rockets** (18–5) defeated the **Denver Nuggets** (10–13) 108–96 on Tuesday at the Toyota Center in Houston. The **Rockets** had a strong first half where they outscored . . . The **Rockets** were led by **Donatas Motiejunas**, who scored a game–high of 25 points . . . **James Harden** also played a factor in the win, as he went 7–for . . . Coming off the bench, **Donatas Motiejunas** had a big game and finished with 25 points . . . The only other player to reach double figures in points was **Arron Afflalo**, who came off the bench for 12 points . . . Coming off the bench, **Arron Afflalo** chipped in with 12 points . . . The **Nuggets**' next game will be on the road against the Boston Celtics on Friday, while the **Nuggets** will travel to Boston to play the Celtics on Wednesday.
The **Houston Rockets** (18–5) defeated the **Denver Nuggets** (10–13) 108–96 on Monday at the Toyota Center in Houston. The **Rockets** were the superior shooters in this game, going . . . The **Rockets** were led by the duo of **Dwight Howard** and **James Harden**. **Howard** shot 9–for–11 from the field and . . . **Harden** on the other hand recorded 24 points (7–20 FG, 2–5 3Pt, 8–9 FT), 10 rebounds and 10 assists, The only other Nugget to reach double figures in points was **Arron Afflalo**, who finished with 12 points (4–17 FG,. . . The **Rockets**' next game will be on the road against the New Orleans Pelicans on Wednesday, while the **Nuggets** will travel to Los Angeles to play the Clippers on Friday.

Table 4: Examples of model output for NCP+CC (top) and ENT (bottom) on ROTOWIRE. Recurring entities

Summary

- A novel entity-aware model for data-to-text generation which is linguistically motivated.
- A new dataset for data-to-text generation.
- A comprehensive evaluation and comparison study which highlights the merits and shortcomings.

Summary

- Strength
  - Provide a larger scale dataset which could alleviate reviewers' concern that any proposed model is dataset-specific.
- Weakness
  - Lack intuitive interpretation of those two gates' performance in addition to automatic evaluation result.

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**[♮‡]
[†]Nara Institute of Science and Technology [‡]Artificial Intelligence Research Center, AIST
[♮]Tokyo Institute of Technology [♭]Ochanomizu University [♯]The University of Tokyo
{iso.hayate.id3,aramaki}@is.naist.jp koba@is.ocha.ac.jp
{yui.uehara,ishigaki.t,hiroshi.noji,takamura.hiroya}@aist.go.jp
yusuke@is.s.u-tokyo.ac.jp okazaki@c.titech.ac.jp

ACL 2019

Motivation

- Simulate the human-like writing process that gradually selects the information by determining the intermediate variables while writing the summary.
  - input data is too large for a naive model to find its salient part
  - the salient part moves as the summary explains the data
- There are some writer-specific patterns and characteristics
  - how data records are selected to be mentioned
  - how data records are expressed as text
  - 32 writers wrote ~4K game summaries in ROTOWIRE

Method

- Proposed Model consists of **tracking** module and **text generation** module
- Two hidden states
  - $\boldsymbol{h}^{\mathrm{ENT}}$: to remember data records that have been referred to
  - $\boldsymbol{h}^{\mathrm{LM}}$: to generate summary
- Four variables
  - Z_t: binary number that decide whether to refer records at time step t.
  - E_t: indicates salient entities (players/teams)
  - A_t: indicates salient attributes (e.g. points, rebounds)
  - N_t: decides if a numeric value should be expressed in Arabic (e.g. 5) or English word (five)

Method

- Initialization
  - Record representation

$$\boldsymbol{r}_{e,a,v} = \tanh\left(\boldsymbol{W}^{\mathrm{R}}(\boldsymbol{e} \oplus \boldsymbol{a} \oplus \boldsymbol{v})\right)$$

  - Entity representation for performance in the game

$$\bar{\boldsymbol{e}} = \tanh\left(\sum_{a \in \mathcal{A}} \boldsymbol{W}_a^{\mathrm{A}} \boldsymbol{r}_{e,a,\boldsymbol{x}[e,a]}\right)$$

  - Initialize hidden state
    - $<\mathrm{SOD}>$ for $\boldsymbol{h}^{\mathrm{LM}}$
    - Average embeddings of $\bar{\boldsymbol{e}}$ for $\boldsymbol{h}^{\mathrm{ENT}}$

Method

- Saliency transition
  - whether to transition to another data record or not at time

$$p(Z_t = 1 \mid \boldsymbol{h}_{t-1}^{\text{LM}}, \boldsymbol{h}_{t-1}^{\text{ENT}}) = \sigma(\boldsymbol{W}_z(\boldsymbol{h}_{t-1}^{\text{LM}} \oplus \boldsymbol{h}_{t-1}^{\text{ENT}}))$$

  - If it decides not to transition, tracking states remain the same

$$\boldsymbol{h}_t^{\text{ENT}} = \boldsymbol{h}_{t-1}^{\text{ENT}}$$

- Selection and tracking
  - If it decides to transition, select entity and attribute

$$p(E_t = e \mid \boldsymbol{h}_{t-1}^{\text{LM}}, \boldsymbol{h}_{t-1}^{\text{ENT}}) \propto \begin{cases} \exp\left(\boldsymbol{h}_s^{\text{ENT}} \boldsymbol{W}^{\text{OLD}} \boldsymbol{h}_{t-1}^{\text{LM}}\right) & \text{if } e \in \mathcal{E}_{t-1} \\ \exp\left(\bar{\boldsymbol{e}} \boldsymbol{W}^{\text{NEW}} \boldsymbol{h}_{t-1}^{\text{LM}}\right) & \text{otherwise} \end{cases}$$

$$p(A_t = a \mid e_t, \boldsymbol{h}_{t-1}^{\text{LM}}, \boldsymbol{h}_t^{\text{ENT}'}) \propto \exp\left(\boldsymbol{r}_{e_t,a,\boldsymbol{x}[e_t,a]} \boldsymbol{W}^{\text{ATTR}} (\boldsymbol{h}_{t-1}^{\text{LM}} \oplus \boldsymbol{h}_t^{\text{ENT}'})\right)$$

$$\boldsymbol{h}_t^{\text{ENT}'} = \begin{cases} \boldsymbol{h}_{t-1}^{\text{ENT}} & \text{if } e_t = e_{t-1} \\ \text{GRU}^{\text{E}}(\bar{\boldsymbol{e}}, \boldsymbol{h}_{t-1}^{\text{ENT}}) & \text{else if } e_t \notin \mathcal{E}_{t-1} \\ \text{GRU}^{\text{E}}(\boldsymbol{W}_s^{\text{S}} \boldsymbol{h}_s^{\text{ENT}}, \boldsymbol{h}_{t-1}^{\text{ENT}}) & \text{otherwise.} \end{cases}$$

$$\boldsymbol{h}_t^{\text{ENT}} = \text{GRU}^{\text{A}}(\boldsymbol{r}_{e_t,a_t,\boldsymbol{x}[e_t,a_t]}, \boldsymbol{h}_t^{\text{ENT}'}).$$

Method

- Summary Generation
  - If the model refers to a new data record ($Z_t = 1$)
    - Copy and decide whether in Arabic numerals or English word

$$p(N_t = 1 \mid \boldsymbol{h}_{t-1}^{\mathrm{LM}}, \boldsymbol{h}_t^{\mathrm{ENT}}) = \sigma(\boldsymbol{W}^{\mathrm{N}}(\boldsymbol{h}_{t-1}^{\mathrm{LM}} \oplus \boldsymbol{h}_t^{\mathrm{ENT}}))$$

  - New hidden states of the language model

$$\boldsymbol{h}'_t = \tanh\left(\boldsymbol{W}^{\mathrm{H}}(\boldsymbol{h}_{t-1}^{\mathrm{LM}} \oplus \boldsymbol{h}_t^{\mathrm{ENT}})\right)$$

  - If the salient data record is the same as the previous one($Z_t = 0$)

$$p(Y_t \mid \boldsymbol{h}'_t) = \mathbf{softmax}(\boldsymbol{W}^{\mathrm{Y}}\boldsymbol{h}'_t)$$

  - Update hidden states of the LM

$$\boldsymbol{h}_t^{\mathrm{LM}} = \mathbf{LSTM}(\boldsymbol{y}_t \oplus \boldsymbol{h}'_t, \boldsymbol{h}_{t-1}^{\mathrm{LM}})$$

  - Incorporate Writer Information

$$\boldsymbol{h}'_t = \tanh\left(\boldsymbol{W}'^{\mathrm{H}}(\boldsymbol{h}_{t-1}^{\mathrm{LM}} \oplus \boldsymbol{h}_t^{\mathrm{ENT}} \oplus \boldsymbol{w})\right)$$

Method

- Learning Objective
  - Fully supervised

$$
\begin{aligned}
&\log p(Y_{1:T}, Z_{1:T}, E_{1:T}, A_{1:T}, N_{1:T} \mid \boldsymbol{x}) \\
&= \sum_{t=1}^{T} \log p(Z_t = z_t \mid \boldsymbol{h}_{t-1}^{\mathrm{LM}}, \boldsymbol{h}_{t-1}^{\mathrm{ENT}}) \\
&+ \sum_{t:Z_t=1} \log p(E_t = e_t \mid \boldsymbol{h}_{t-1}^{\mathrm{LM}}, \boldsymbol{h}_{t-1}^{\mathrm{ENT}}) \\
&+ \sum_{t:Z_t=1} \log p(A_t = a_t \mid e_t, \boldsymbol{h}_{t-1}^{\mathrm{LM}}, \boldsymbol{h}_{t}^{\mathrm{ENT}'}) \\
&+ \sum_{t:Z_t=1, a_t \text{is num\_attr}} \log p(N_t = n_t \mid \boldsymbol{h}_{t-1}^{\mathrm{LM}}, \boldsymbol{h}_{t}^{\mathrm{ENT}}) \\
&+ \sum_{t:Z_t=0} \log p(Y_t = y_t \mid \boldsymbol{h}'_t)
\end{aligned}
$$

Experiments

- Construct ROTOWIRE-MODIFIED
  - In the original dataset ROTOWIRE, some NBA games have two documents, one of which is sometimes in the training data and the other is in the test or validation data
  - Original dataset lacks writer information
  - Collected approximately 78% of the documents in the original dataset

Method	RG		CS			CO	BLEU
	#	P%	P%	R%	F1%	DLD%	
GOLD	27.36	93.42	100.	100.	100.	100.	100.
TEMPLATES	54.63	100.	31.01	58.85	40.61	17.50	8.43
Wiseman et al. (2017)	22.93	60.14	24.24	31.20	27.29	14.70	14.73
Puduppully et al. (2019)	33.06	83.17	33.06	43.59	37.60	16.97	13.96
PROPOSED	39.05	**94.43**	**35.77**	**52.05**	**42.40**	**19.38**	**16.15**

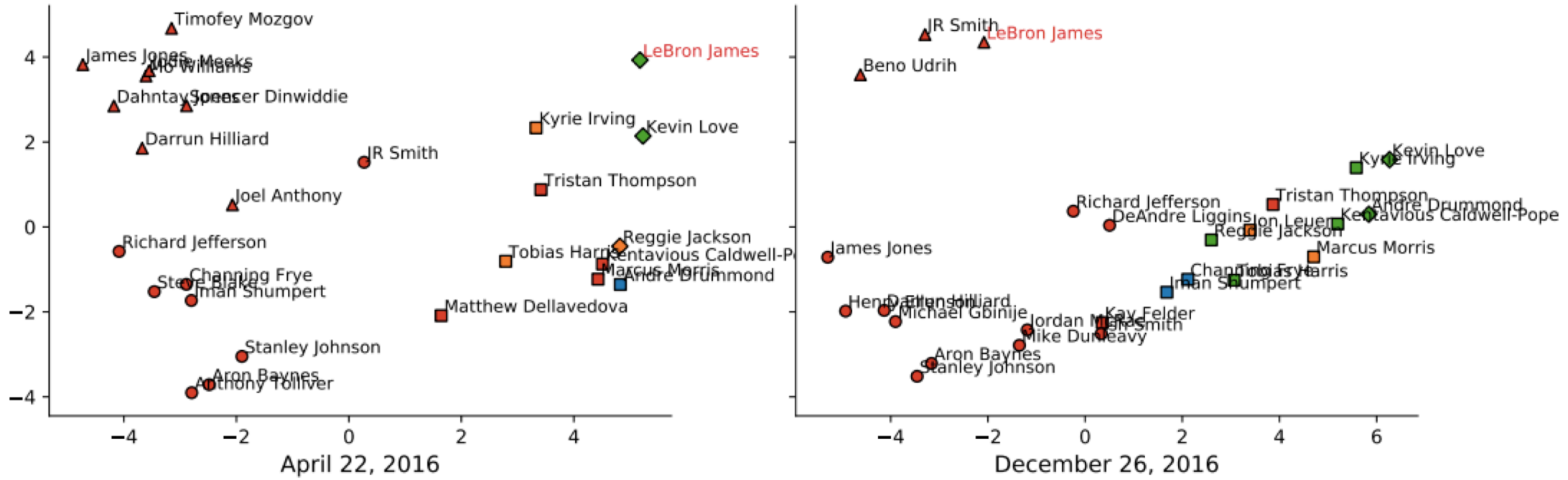
Method	RG		CS			CO	BLEU
	#	P%	P%	R%	F1%	DLD%	
Puduppully et al. (2019)	33.06	83.17	33.06	43.59	37.60	16.97	13.96
+ w in stage 1	28.43	**84.75**	**45.00**	**49.73**	**47.25**	22.16	18.18
+ w in stage 2	35.06	80.51	31.10	45.28	36.87	16.38	17.81
+ w in stage 1 & 2	28.00	82.27	44.37	48.71	46.44	**22.41**	**18.90**
PROPOSED	39.05	**94.38**	35.77	52.05	42.40	19.38	16.15
+ w	30.25	92.00	**50.75**	**59.03**	**54.58**	**25.75**	**20.84**

Experiments

- Visualization
  - Static entity embedding (blue indicates top 100 players for 2016-17 season)

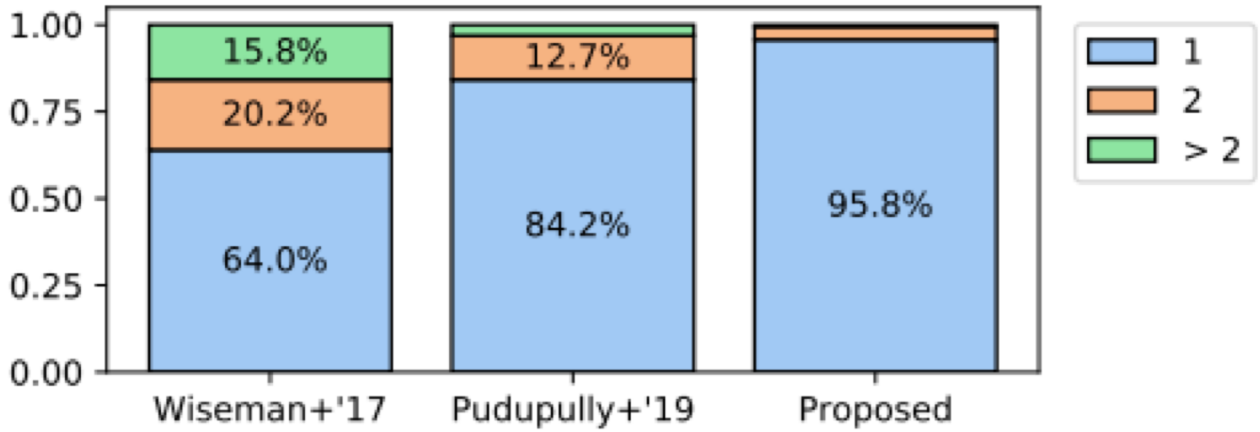
Experiments

- Visualization
  - Dynamic entity embedding



Experiments

- Visualization
  - Duplicate relation mentioning



Experiments

- Example

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

Summary

- They propose a new data-to-text model that produces a summary text while tracking the salient information that imitates a human-writing process.
- Their model outperformed the existing models in all evaluation measures.
- They also explored the effects of incorporating writer information to data-to-text models.

Summary

- Strength
  - 32 writer wrote ~4k game summaries in ROTOWIRE indicates potential research direction on modeling individual’ s writing style.

- Weakness
  - NCP+CC’ s result in the paper for comparison seems to be less satisfactory especially w.r.t BLEU.