

---

# Table-to-text Generation by Structure-aware Seq2seq Learning

---

Authors: **Tianyu Liu**, Kexiang Wang, Lei Sha, Baobao Chang and Zhifang Sui

Affiliation: Key Laboratory of Computational Linguistics(ICL), Peking University, Beijing, China

Speaker: **Tianyu Liu**

E-mail: [tianyu0421@pku.edu.cn](mailto:tianyu0421@pku.edu.cn)

# Knowledge Modeling



**Machine Reading**  
Unstructured text  Structured data



**Machine Writing**  
Structured data  Unstructured text



# Knowledge Modeling



Machine Reading

Unstructured text  $\longrightarrow$  Structured data



Machine Writing

Structured data  $\longrightarrow$  Unstructured text



北京大学  
PEKING UNIVERSITY

# Data-to-text



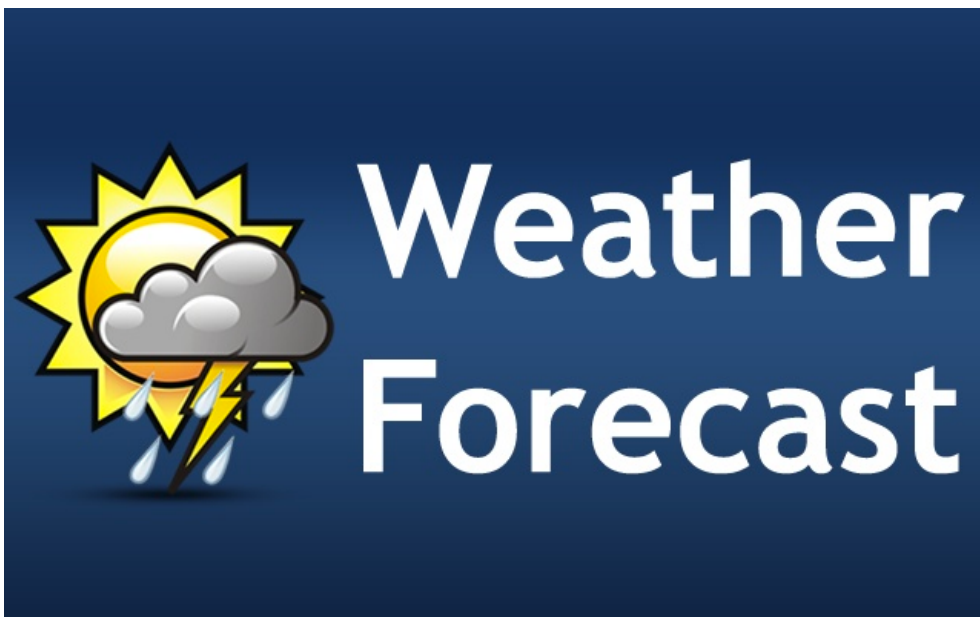
	WIN	LOSS	PTS	FG_PCT	RB	AS ...
TEAM						
Heat	11	12	103	49	47	27
Hawks	7	15	95	43	33	20

	AS	RB	PT	FG	FGA	CITY ...
PLAYER						
Tyler Johnson	5	2	27	8	16	Miami
Dwight Howard	4	17	23	9	11	Atlanta
Paul Millsap	2	9	21	8	12	Atlanta
Goran Dragic	4	2	21	8	17	Miami
Wayne Ellington	2	3	19	7	15	Miami
Dennis Schroder	7	4	17	8	15	Atlanta
Rodney McGruder	5	5	11	3	8	Miami
Thabo Sefolosha	5	5	10	5	11	Atlanta
Kyle Korver	5	3	9	3	9	Atlanta
...						

The Atlanta Hawks defeated the Miami Heat , 103 - 95 , at Philips Arena on Wednesday . Atlanta was in desperate need of a win and they were able to take care of a shorthanded Miami team here . Defense was key for the Hawks , as they held the Heat to 42 percent shooting and forced them to commit 16 turnovers . Atlanta also dominated in the paint , winning the rebounding battle , 47 - 34 , and outscoring them in the paint 58 - 26. The Hawks shot 49 percent from the field and assisted on 27 of their 43 made baskets . This was a near wire - to - wire win for the Hawks , as Miami held just one lead in the first five minutes . Miami ( 7 - 15 ) are as beat - up as anyone right now and it 's taking a toll on the heavily used starters . Hassan Whiteside really struggled in this game , as he amassed eight points , 12 rebounds and one blocks on 4 - of - 12 shooting ...

Wiseman S, Shieber S M, Rush A M. Challenges in data-to-document generation EMNLP 2017

# Data-to-text



```
temperature(time=17-06, min=48, mean=53, max=61)
windSpeed(time=17-06, min=3, mean=6, max=11)
windDir(time=17-06, mode=SSW)
gust(time=17-06, min=0, mean=0, max=0)
skyCover(time=17-21, mode=0-25)
skyCover(time=02-06, mode=75-100)
precipChance(time=17-06, min=2, mean=14, max=20)
rainChance(time=17-06, mode=someChance)
```

“a 20 percent chance of showers after midnight. increasing clouds, with a low around 48 southwest wind between 5 and 10 mph”

Mei H, Bansal M, Walter M R. What to talk about and how? selective generation using lstms with coarse-to-fine alignment NAACL2016.

# Data-to-text



inform(name=X-name, type=placetoeat, eatype=restaurant, area=riverside, food=French)  
X is a French restaurant on the riverside.  
X is a restaurant providing french and **continental** and by the river.  
X is a restaurant that serves french **takeaway**. [**riverside**]  
X is a french restaurant in the riverside area.  
X is a restaurant in the riverside that serves **italian** food. [**French**]  
X is a restaurant in the riverside that serves **italian** food. [**French**]  
X is a restaurant in the riverside area that serves french food.

Dušek O, Jurčiček F. Sequence-to-sequence generation for spoken dialogue via deep syntax trees and strings. ACL 2016.



# Data-to-text

George Mikell



Mikell in March 2017

Born	Jurgis Mikelaitis 4 April 1929 (age 88) Bildeniai, <a href="#">Lithuania</a>
Nationality	Lithuanian, Australian
Occupation	Actor, writer
Years active	1957—present
Known for	<a href="#">The Guns of Navarone</a> <a href="#">The Great Escape</a>
Height	6' 0" (1.83m)



WIKIPEDIA  
The Free Encyclopedia



**George Mikell** (born **Jurgis Mikelaitis**; 4 April 1929) is a Lithuanian-Australian actor and writer best known for his performances in [The Guns of Navarone](#) (1961) and [The Great Escape](#) (1963).

## Our work !



北京大学  
PEKING UNIVERSITY

# How to represent a table ?

## 1. Key-value map

(Name, George Mikell); (Nationality, Lithuanian Australian); (Years Active, 1957--present) ....



# How to represent a table ?

## 1. Key-value map

(Name, George Mikell); (Nationality, Lithuanian Australian); (Years Active, 1957--present) ....

## 2. Triple

(George Mikell, FullName, George Mikell);

(George Mikell, Nationality, Lithuanian Australian);

(George Mikell, Known for, The Guns of Navarone); (The Guns of Navarone, Directed by, J.Lee Thompson) ....

# How to represent a table ?

## 1. Key-value map

(Name, George Mikell); (Nationality, Lithuanian Australian); (Years Active, 1957--present) ....

## 2. Triple

(George Mikell, FullName, George Mikell);

(George Mikell, Nationality, Lithuanian Australian);

(George Mikell, Known for, The Guns of Navarone); (The Guns of Navarone, Directed by, J.Lee Thompson) ....

## 3. Multi-element Tuple

	WIN	LOSS	PTS	FG_PCT	RB	AS ...
TEAM						
Heat	11	12	103	49	47	27
Hawks	7	15	95	43	33	20



# How to represent a table ?

## 1. Key-value map ✓

(Name, George Mikell); (Nationality, Lithuanian Australian); (Years Active, 1957--present) ....

## 2. Triple

(George Mikell, FullName, George Mikell);

(George Mikell, Nationality, Lithuanian Australian);

(George Mikell, Known for, The Guns of Navarone); (The Guns of Navarone, Directed by, J.Lee Thompson) ....

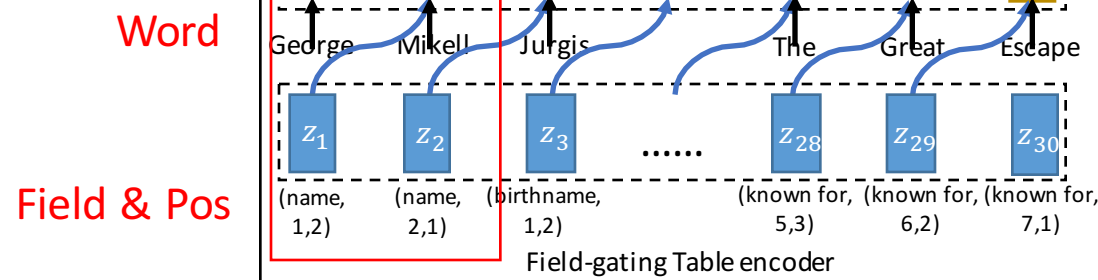
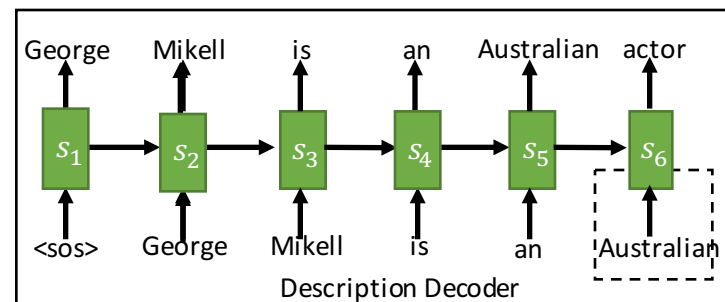
## 3. Multi-element Tuple

	WIN	LOSS	PTS	FG_PCT	RB	AS ...
TEAM						
Heat	11	12	103	49	47	27
Hawks	7	15	95	43	33	20



# Table Encoder

name	George Mikell
birthname	Jurgis Mikelaityis
birthdate	4 April 1929 (age 88)
birthplace	Bildeniai, Lithuania
nationality	Lithuanian, Australian
occupation	Actor, writer
years active	1957–present
known for	The Guns of Navarone The Great Escape



# Table Encoder

Word embedding + Structure embedding

(name, **George** Mikell)

		word	Field embedding
name	George Mikell	George	(name, 1, 2)
birthname	Jurgis Mikelaityis	Mikell	(name, 2, 1)
birthdate	4 April 1929 (age 88)	Jurgis	(birthdate, 1, 2)
birthplace	Bilideniai, Lithuania	Mikelaityis	(birthdate, 2, 1)
nationality	Lithuanian, Australian	4	(birthdate, 1, 7)
occupation	Actor, writer	April	(birthdate, 2, 6)
years active	1957–present	...	...
known for	The Guns of Navarone	The	(known for, 5, 3)
	The Great Escape	Great	(known for, 6, 2)
		Escape	(known for, 7, 1)



# Table Encoder

Word embedding

(name, **George** Mikell)

		<i>word</i>	<i>Field embedding</i>
name	George Mikell	George	(name, 1, 2)
birthname	Jurgis Mikelaityis	Mikell	(name, 2, 1)
birthdate	4 April 1929 (age 88)	Jurgis	(birthname, 1, 2)
birthplace	Bilteniai, Lithuania	Mikelaityis	(birthname, 2, 1)
nationality	Lithuanian, Australian	4	(birthdate, 1, 7)
occupation	Actor, writer	April	(birthdate, 2, 6)
years active	1957–present	...	...
known for	The Guns of Navarone	The	(known for, 5, 3)
	The Great Escape	Great	(known for, 6, 2)
		Escape	(known for, 7, 1)



# Table Encoder

Structure embedding (Field type embedding + Position embedding)

(name, **George** Mikell)

		<i>word</i>	<i>Field embedding</i>
name	George Mikell	George	(name, 1, 2)
birthname	Jurgis Mikelaity	Mikell	(name, 2, 1)
birthdate	4 April 1929 (age 88)	Jurgis	(birthdate, 1, 2)
birthplace	Bildeniai, Lithuania	Mikelaity	(birthdate, 2, 1)
nationality	Lithuanian, Australian	4	(birthdate, 1, 7)
occupation	Actor, writer	April	(birthdate, 2, 6)
years active	1957–present	...	...
known for	The Guns of Navarone	The	(known for, 5, 3)
	The Great Escape	Great	(known for, 6, 2)
		Escape	(known for, 7, 1)



# Table Encoder

Structure embedding (Field type embedding + Position embedding)

(name, **George** Mikell)

		<i>word</i>	<i>Field embedding</i>
name	George Mikell	George	(name, 1, 2)
birthname	Jurgis Mikelaityis	Mikell	(name, 2, 1)
birthdate	4 April 1929 (age 88)	Jurgis	(birthname, 1, 2)
birthplace	Bildeniai, Lithuania	Mikelaityis	(birthname, 2, 1)
nationality	Lithuanian, Australian	4	(birthdate, 1, 7)
occupation	Actor, writer	April	(birthdate, 2, 6)
years active	1957–present	...	...
known for	The Guns of Navarone The Great Escape	The	(known for, 5, 3)
		Great	(known for, 6, 2)
		Escape	(known for, 7, 1)



# Field-gating Encoder

## Vanilla LSTM

$$\begin{pmatrix} i_t \\ f_t \\ o_t \\ \hat{c}_t \end{pmatrix} = \begin{pmatrix} \text{sigmoid} \\ \text{sigmoid} \\ \text{sigmoid} \\ \tanh \end{pmatrix} W_{4n,2n}^c \begin{pmatrix} d_t \\ h_{t-1} \end{pmatrix}$$

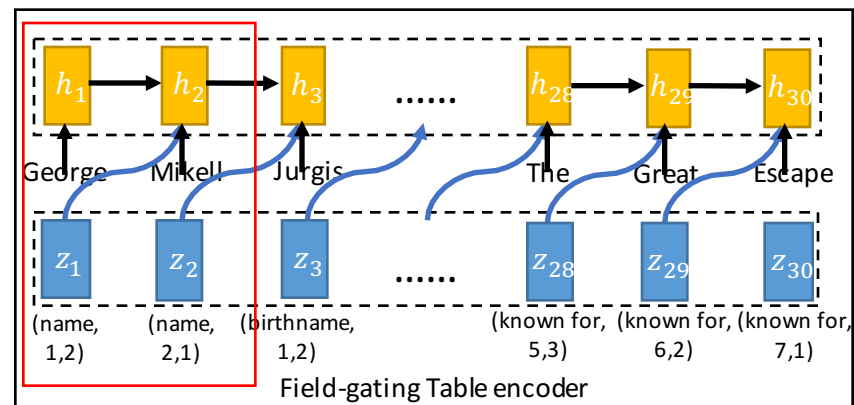
$$c_t = f_t \odot c_{t-1} + i_t \odot \hat{c}_t$$

$$h_t = o_t \odot \tanh(c_t)$$

## Undate cell state

$$\begin{pmatrix} l_t \\ \hat{z}_t \end{pmatrix} = \begin{pmatrix} \text{sigmoid} \\ \tanh \end{pmatrix} W_{2n,2n}^d (z_t)$$

$$c'_t = f_t \odot c_{t-1} + i_t \odot \hat{c}_t + l_t \odot \hat{z}_t$$



# Human-like Generation

name	George Mikell
birthname	Jurgis Mikelaityis
birthdate	4 April 1929 (age 88)
birthplace	Bildeniai, Lithuania
nationality	Lithuanian, Australian
occupation	Actor, writer
years active	1957–present
known for	The Guns of Navarone The Great Escape

George Mikell



# Human-like Generation

name	George Mikell
birthname	Jurgis Mikelaïtis
birthdate	4 April 1929 (age 88)
birthplace	Bildeniai, Lithuania
nationality	Lithuanian, Australian
occupation	Actor, writer
years active	1957–present
known for	The Guns of Navarone The Great Escape

**George Mikell** (born **Jurgis Mikelaïtis**; 4 April 1929)



# Human-like Generation

name	George Mikell
birthname	Jurgis Mikelaityis
birthdate	4 April 1929 (age 88)
birthplace	Bildeniai, Lithuania
nationality	Lithuanian, Australian
occupation	Actor, writer
years active	1957–present
known for	The Guns of Navarone The Great Escape

**George Mikell** (born **Jurgis Mikelaityis**; 4 April 1929) is a Lithuanian-Australian actor and writer



# Human-like Generation

name	George Mikell
birthname	Jurgis Mikelaityis
birthdate	4 April 1929 (age 88)
birthplace	Bildeniai, Lithuania
nationality	Lithuanian, Australian
occupation	Actor, writer
years active	1957–present

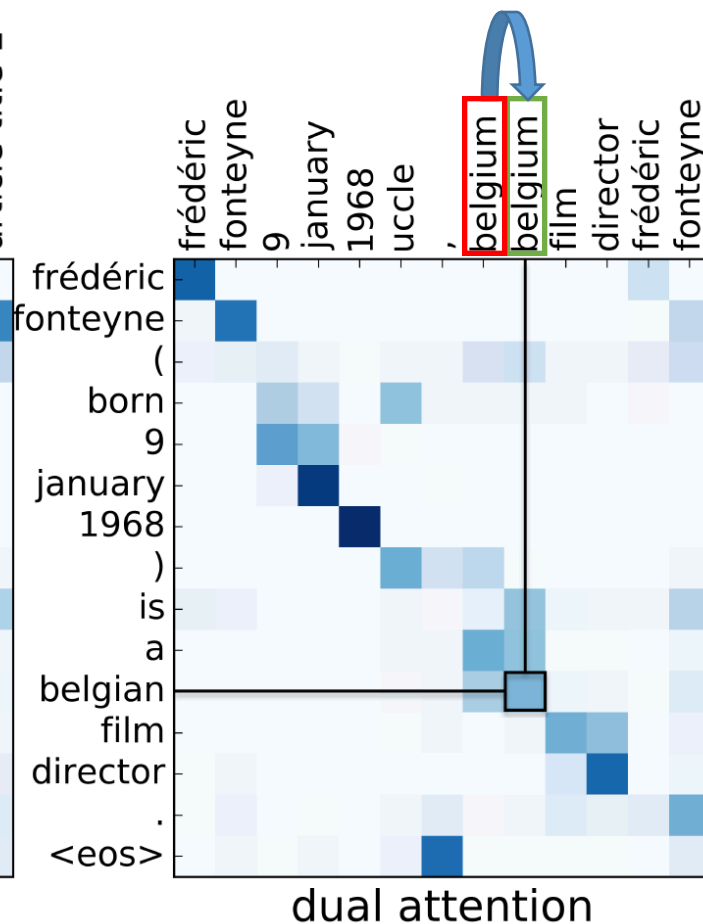
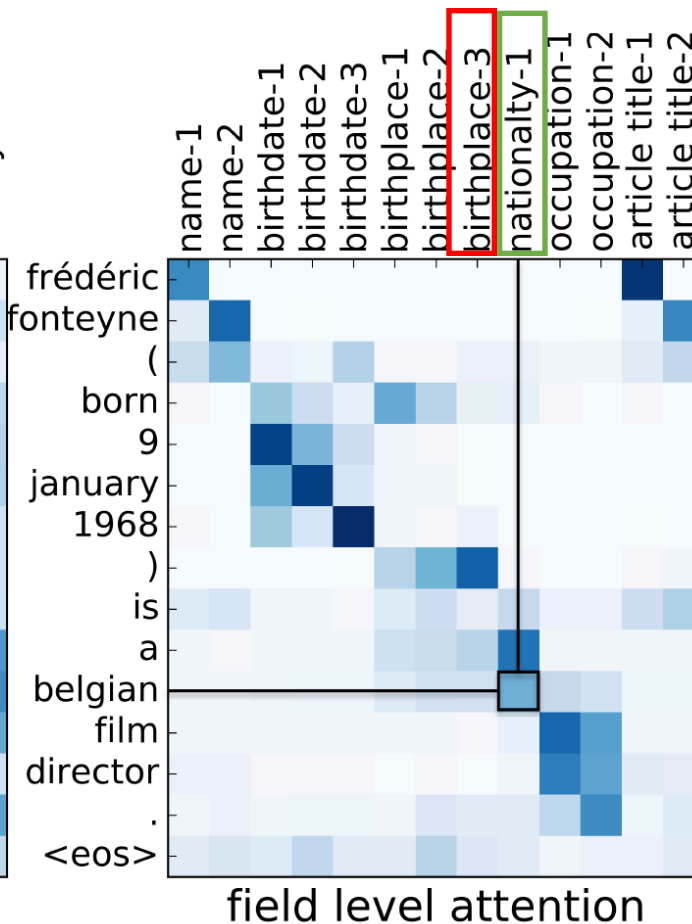
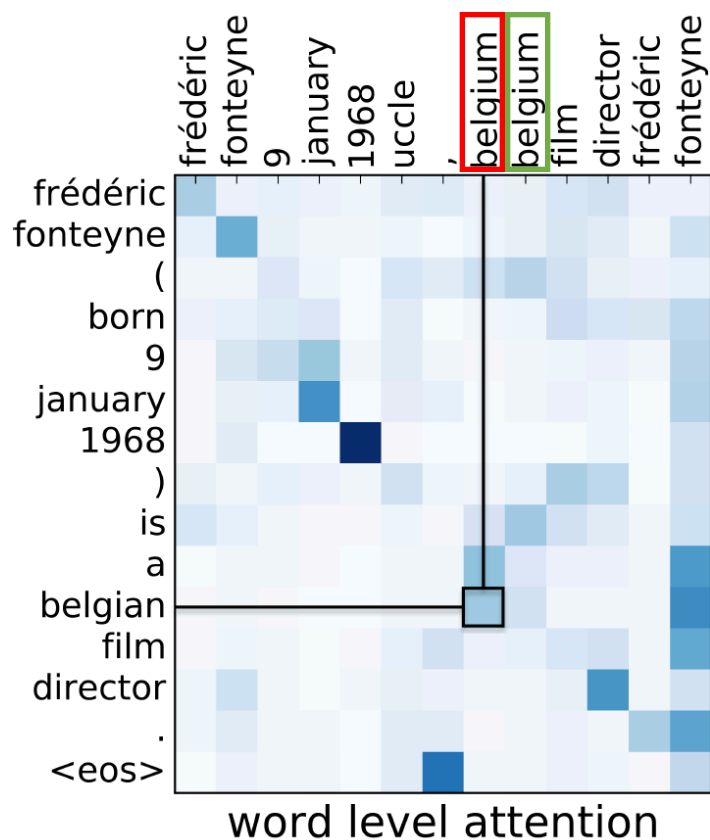
known for	The Guns of Navarone The Great Escape
-----------	--

**George Mikell** (born **Jurgis Mikelaityis**; 4 April 1929) is a Lithuanian-Australian actor and writer best known for his performances in *The Guns of Navarone* (1961) and *The Great Escape* (1963).





# Case study



# Dataset: WIKIBIO

## Infobox

name	George Mikell
birthname	Jurgis Mikelaitis
birthdate	4 April 1929 (age 88)
birthplace	Bildeniai, Lithuania
nationality	Lithuanian, Australian
occupation	Actor, writer
years active	1957–present
known for	The Guns of Navarone The Great Escape

## Description

**George Mikell** (born **Jurgis Mikelaitis**; 4 April 1929) is a Lithuanian-Australian actor and writer best known for his performances in *The Guns of Navarone* (1961) and *The Great Escape*(1963).

	# tokens per sentence	# table token per sent.	# tokens per table	# fields per table
Mean	26.1	9.5	53.1	19.7

**WIKIBIO** (Lebret, Grangier, and Auli (2016)) contains **728,321** articles from English Wikipedia (**Sep 2015**). The dataset uses the **first sentence** of each article as the description of the related infobox.



# Experiments

## Statistical Baselines:

**KN & Template KN:** The Kneser-Ney (KN) model is a widely used language model. Template KN is a KN model over templates which also serves as a baseline in [Lebret et al., 2016]

**NLM & Table NLM:** NLM is a naive statistical language model proposed by Lebret et al. 2016. Table NLM includes local and global conditioning over the tables

## Neural Network Baseline:


**Seq2seq:** vanilla seq2seq neural architecture is also provided as a strong baseline which uses the concatenation of word embedding, field embedding and position embedding as the model input.

Model	BLEU	ROUGE
KN	2.21	0.38
Template KN	19.80	10.70
NLM	$4.17 \pm 0.54$	$1.48 \pm 0.23$
Table NLM	$34.70 \pm 0.36$	$25.80 \pm 0.36$
Seq2seq	$42.06 \pm 0.32$	$38.06 \pm 0.36$
+ field (concat)	$43.34 \pm 0.37$	$39.84 \pm 0.32$
+ pos (concat)	$43.65 \pm 0.44$	$40.32 \pm 0.23$
Field-gating Seq2seq	$43.74 \pm 0.23$	$40.53 \pm 0.31$
+ dual attention	<b><math>44.89 \pm 0.33</math></b>	$41.21 \pm 0.25$
+ beam search (k=5)	44.71	<b>41.65</b>

Table 3: BLEU-4 and ROUGE-4 for structure-aware seq2seq model (last three rows), statistical language model (first four rows) and vanilla seq2seq model with field and position input (three rows in the middle).



# Does the order matter?

name	George Mikell	Shuffle 	years active	1957–present
birthname	Jurgis Mikelaityis		occupation	Actor, writer
birthdate	4 April 1929 (age 88)		known for	The Guns of Navarone The Great Escape
birthplace	Bildeniai, Lithuania		birthplace	Bildeniai, Lithuania
nationality	Lithuanian, Australian		nationality	Lithuanian, Australian
occupation	Actor, writer		birthname	Jurgis Mikelaityis
years active	1957–present		name	George Mikell
known for	The Guns of Navarone The Great Escape		birthdate	4 April 1929 (age 88)

Model	BLEU	ROUGE
Seq2seq	40.04 (-2.02)	36.85 (-1.21)
+ field & pos	42.10 (-1.55)	38.97 (-1.35)
Structure-aware	<b>44.28 (-0.61)</b>	<b>40.79 (-0.42)</b>

Table 4: Experiments on the disordered tables to show the efficiency of global addressing.







# Thanks.



北京大学  
PEKING UNIVERSITY