



# Describing a Knowledge Base

Qingyun Wang<sup>1</sup>, Xiaoman Pan<sup>1</sup>, Lifu Huang<sup>1</sup>, Boliang Zhang<sup>1</sup>,  
Zhiying Jiang<sup>1</sup>, Heng Ji<sup>1</sup> and Kevin Knight<sup>2</sup>

<sup>1</sup>Rensselaer Polytechnic Institute

<sup>2</sup>DiDi Labs and University of Southern California

Codes and Data are public at:



# Knowledge Base Description

**Jean Jennings Bartik** (December 27, 1924 – March 23, 2011) was one of the original programmers for the **ENIAC** computer. She studied mathematics in school then began work at the **University of Pennsylvania**, first manually calculating ballistics trajectories, then using **ENIAC** to do so.



Jean Bartik



<b>Born</b>	Betty Jean Jennings December 27, 1924 <a href="#">Gentry County, Missouri</a>
<b>Died</b>	March 23, 2011 (aged 86)
<b>Nationality</b>	American
<b>Alma mater</b>	<a href="#">Northwest Missouri State Teachers College</a> (honorary D. Sc., 2002), <a href="#">University of Pennsylvania</a> (B.S., 1945; Master's 1967)
<b>Spouse(s)</b>	William Bartik
<b>Awards</b>	<a href="#">Computer Pioneer Award</a> (2008)
	<b>Engineering career</b>
<b>Employer(s)</b>	<a href="#">University of Pennsylvania</a> , <a href="#">Eckert–Mauchly Computer Corporation</a> , <a href="#">Auerbach Publishers</a> , <a href="#">Data Decisions</a>
<b>Projects</b>	<a href="#">ENIAC</a>
<b>Awards</b>	<a href="#">WITI Hall of Fame</a> <a href="#">Computer History Museum Fellow</a> (2008)

# Background

- Template based description generation
  - Pro: Tailored to the domain and high quality
  - Con: Hard to find appropriate template
- Language-model based generation
  - Pro: Flexible planning of each slot
  - Con: Might miss words and dependencies

Animal Antics

Read the words in the word box.  
Read each sentence.  
Fill in the blank with a word from the word box to complete each sentence.

eat runs can went jump

The frog can \_\_\_\_\_ 

 The duck \_\_\_\_\_ swim.

The rabbit likes to \_\_\_\_\_ carrots. 

 My dog \_\_\_\_\_ fast!

The turtle and the snail \_\_\_\_\_ for a walk. 

frederick parker-rhodes ( 21 november 1914 – 2 march 1987 ) was an australian rules footballer who played with carlton in the victorian football league ( vfl ) during the XXXXs and XXXXs .

# Dataset construction

- Based on Wikipedia dump (2018/04/01) and Wikidata (2018/04/12)
- For Wikidata table, filter out the slot type of which frequency is less than 3
- For Wikipedia article, use anchor links to locate entities
- Match values and entities between each table and article
- Remove unused entities and values in each table
- Remove the sentences which contain no values

Entity Type	# entity	# types before filtering	# types after filtering	# slots / sentence	# words / sentence	# slots / table	# words / entity	# sentence / entity
Person	100,000	109	76	1.9	16.8	8.0	70.9	4.2
Animal	6,216	30	12	1.3	17.1	3.2	42.2	2.5



Dataset: [https://drive.google.com/file/d/1TzcNdjZ0EsLh\\_rC1pBC7dU70jINcsVJd/view](https://drive.google.com/file/d/1TzcNdjZ0EsLh_rC1pBC7dU70jINcsVJd/view)

# KB Input

Silvi Jan  
סילבי ז'אן



## Personal information

**Date of birth** 27 October 1973 (age 44)  
**Place of birth** Israel  
**Playing position** Striker

## Senior career\*

Years	Team	Apps. (Gls)
1995–1999	Kolbotn I.L. <sup>[1]</sup>	
1999–2002	Hapoel Tel Aviv	
2002–2007	Maccabi Holon	
2009–2012	ASA Tel Aviv	76 (150)

## National team

1997–2007	Israel	22 (29)
-----------	--------	---------

\* Senior club appearances and goals counted for the domestic league only



Slot Type (s)	Row (r)	Slot Value (v)		
Name	1	Silvi Jan		
Member of Sports Team	2	ASA Tel Aviv University		
	3	Hapoel Tel F.C. (women)		
	4	Maccabi Holon F.C. (women)		
	5	Israel women's national football team	Match	22
			Goal	29
Date of birth	6	27 October 1973		
Country of Citizenship	7	Israel		
Position	8	Forward (association football)		

# Reference

**Silvi Jan** (**Hebrew**: סילבי ז'אן, born 27 October 1973) is a retired female **Israeli football** player.

## Club career [ edit ]

Jan played school football in **Netanya**, and was part of the boys' football team of Shorashim school.<sup>[2]</sup> As there was no women's league, Jan tried to find an active team in Europe, eventually signing with Kolbotn I.L.. In 1999, with the establishment of the **Israeli Women's League**, Jan returned to Israel and signed with **Hapoel Tel Aviv**.

Jan played at Hapoel Tel Aviv for four seasons, until the team folded, winning the league and cup in 2001–02. After Hapoel Tel Aviv folded, Jan signed with **Maccabi Holon**, where she played until her retirement in 2007.<sup>[1]</sup> With Maccabi Holon Jan won the league 4 times and the cup 5 times.

In January 2009, Jan returned to league action and joined **ASA Tel Aviv University**.<sup>[3]</sup> Jan played at ASA Tel Aviv until her retirement in May 2012.<sup>[4]</sup> With ASA Tel Aviv Jan Jan won three championships and two cups.

## International career [ edit ]

Silvi Jan has been a **striker** for the **Israel women's national football team** for many years,<sup>[5]</sup> appearing in 22 matches and scoring 29 goals. Jan made her debut against **Romania** on 2 November 1997, and scored her debut goal against **Bosnia and Herzegovina** on 8 April 2008, one of four she scored on the day. Her last international goal was also scored against **Bosnia and Herzegovina**, in a match played on 23 November 2006, in which Israel won 5–2. Jan made her last international appearance against **Poland** on 10 May 2007.



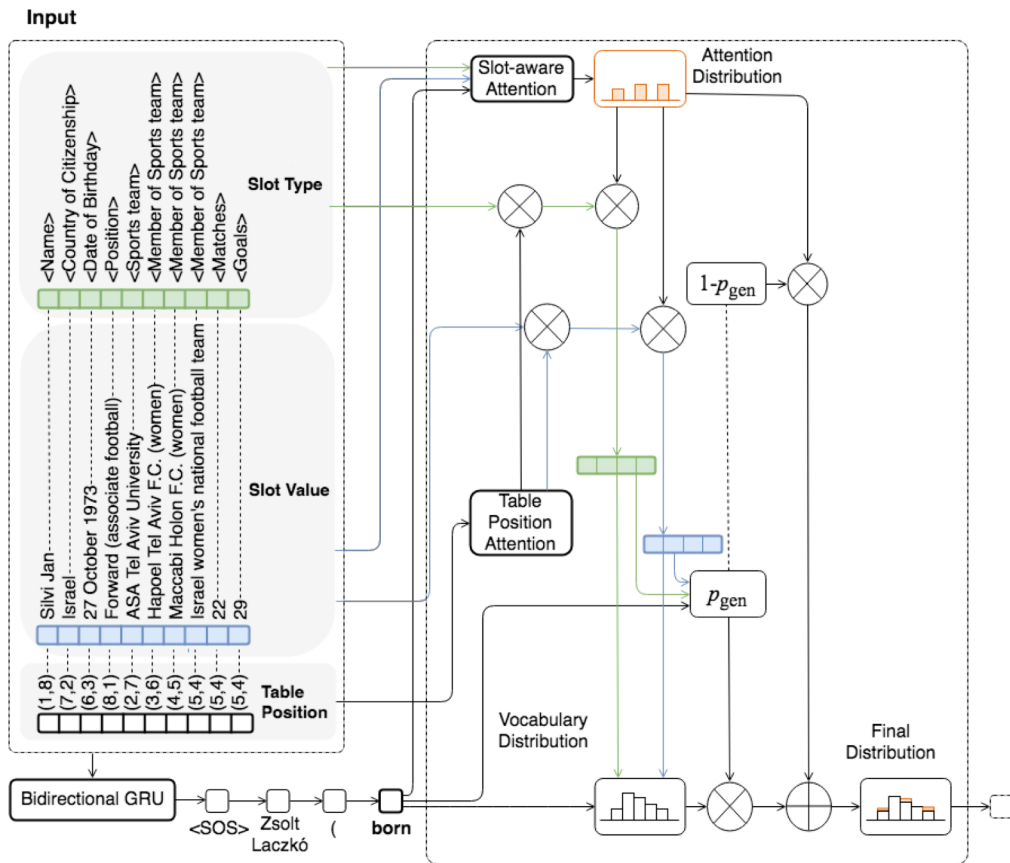
**Silvi Jan** (born 27 October 1973 ) is a retired female **Israeli** football player. **Silvi Jan** has been a **Forward** (association football) for the **Israel women's national football team** for many years appearing in **22** matches and scoring **29** goals. After **Hapoel Tel F.C. (women)** folded, Jan signed with **Maccabi Holon F.C. (women)** where she played until her retirement in 2007. In January 2009, Jan returned to league action and joined **ASA Tel Aviv University**. In 1999, with the establishment of the **Israeli Women's League**, Jan returned to **Israel** and signed with **Hapoel Tel F.C. (women)**.

# Motivation

- How do human describe a KB?
  - Generalize linguistic expressions for slot types
  - Cover facts contained in the input KB
  - Capture interdependence among facts
- Can machine repeat such a process?
  - Neural network based generator
  - Pointer network<sup>[1]</sup> to copy facts
  - Slot type and table position attention to relate slots



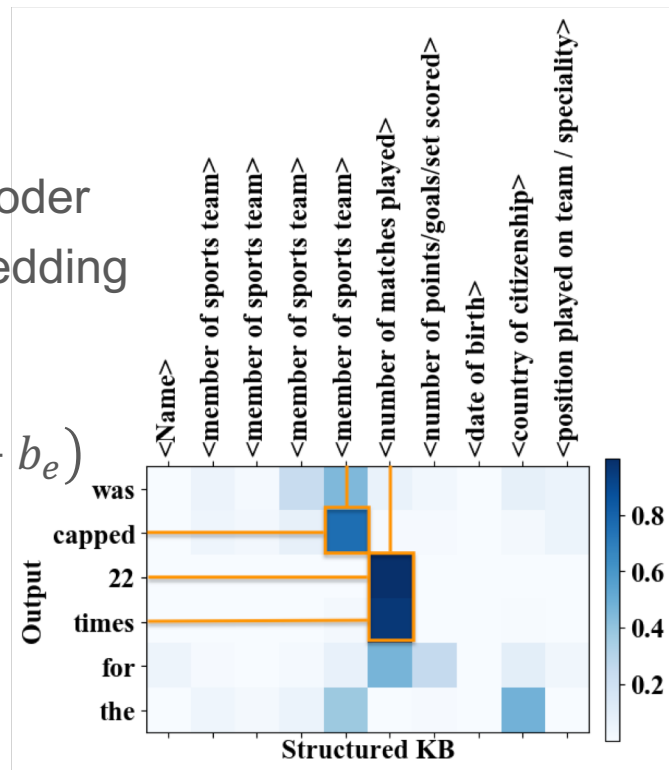
# Model





# Capture slot-aware information

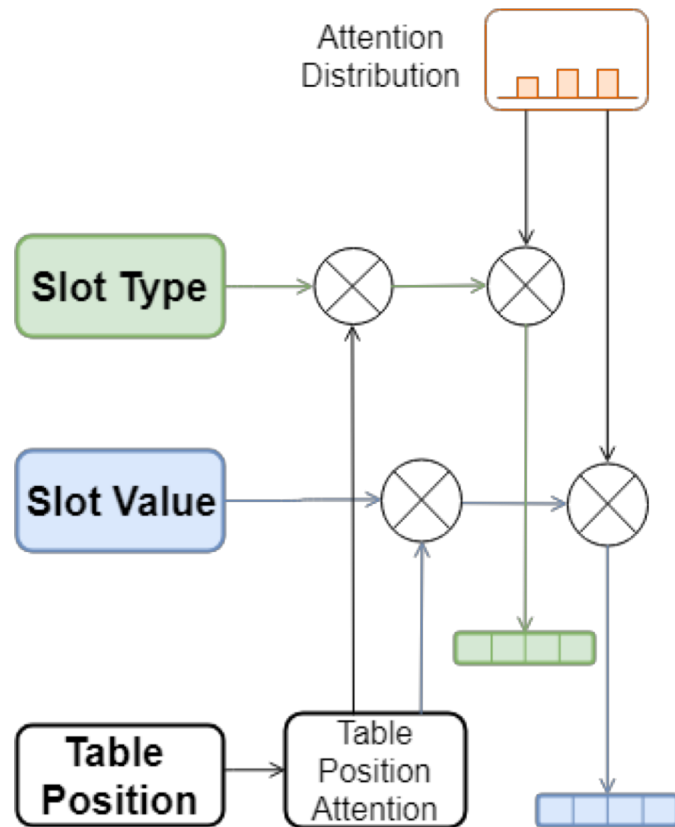
- At each step  $t$ , for each triple  $i$ :
  - Compute *Slot-aware attention*  $\alpha^t$  from decoder hidden  $\tilde{h}^t$ , embedding of slot type  $s_i$ , embedding of slot value  $v_i$ , the coverage vector  $c_i^t$
  - $e_i^t = v^T \tanh(W_h \tilde{h}^t + W_s s_i + W_v v_i + W_c c_i^t + b_e)$
  - $\alpha^t = \text{Softmax}(e^t)$
  - $c_i^t = \sum_{k=0}^{t-1} \alpha_i^k$





# Capture slot dependency

- Update each embedding of slot type  $s_i$  and slot value  $v_i$ 
  - $s_i^* = \sum_{k=1}^n f_{ik} s_k$
  - $v_i^* = \sum_{k=1}^n f_{ik} v_k$
- Update their context vectors  $L_t^*$  and  $L_v^*$  with attention score
  - $L_t^* = \sum_{k=1}^n \alpha^t s_i^*$
  - $L_v^* = \sum_{k=1}^n \alpha^t v_i^*$



# Structure Generator

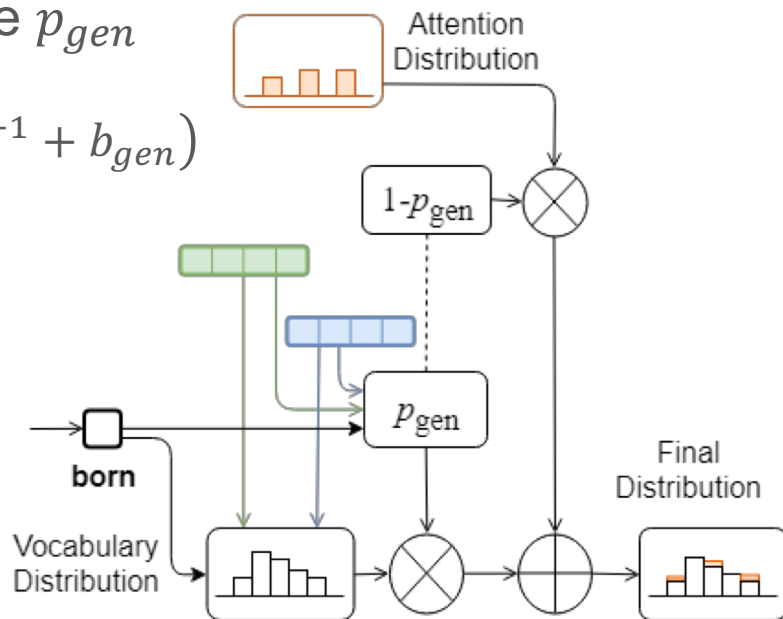
- Obtain vocabulary distribution from decoder and attention, and combine them with structure-aware gate  $p_{gen}$

- $$p_{gen} = \sigma(W_s^T L_s^* + W_v^T L_v^* + W_h^T \tilde{\mathbf{h}}^t + W_y^T y^{t-1} + b_{gen})$$

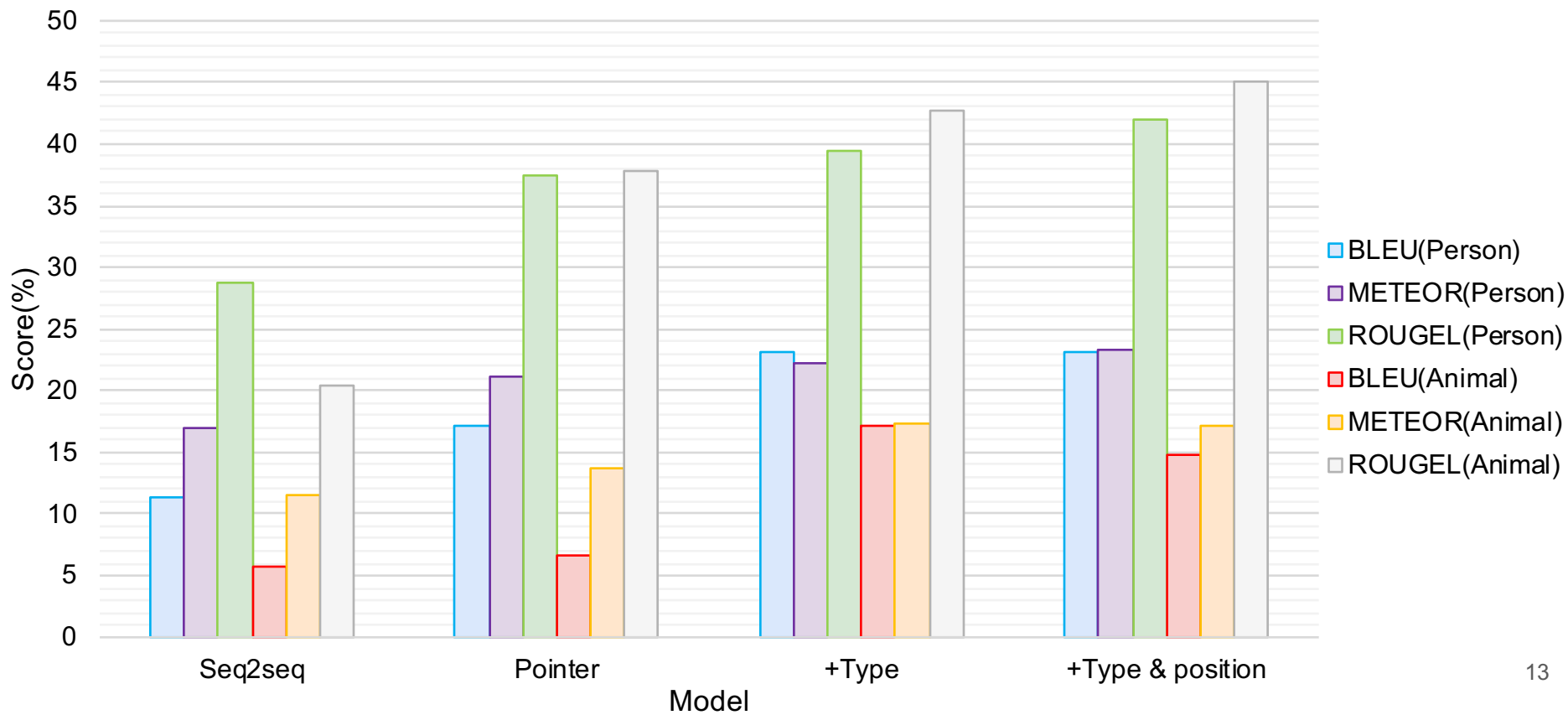
- $$P_{vocab} = \text{Softmax}(V[\tilde{\mathbf{h}}^t; L_s^*; L_v^*] + b_{vocab})$$

- $$P_{source}^j = \sum_{m|v_m=v_j} \alpha_m^t$$

- $$P(y_t) = p_{gen} P_{vocab} + (1 - p_{gen}) P_{source}$$

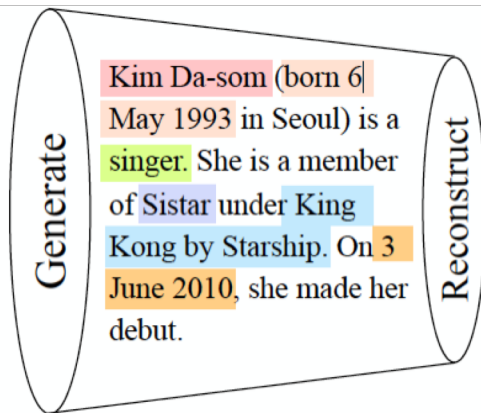


# Standard Metrics (%)



# KB reconstruction metric

Slot Type	Slot Value				
Name	Kim Da-som				
Date of Birth	6 May 1993				
Place of Birth	Gwangju				
Occupation	Singer				
Occupation	Actress				
Genres	K-pop				
Start Active Year	3 June 2010	Start Active Place	Seoul	Start Active Song	PUSH PUSH
Agent	King Kong by Starship				
Associated acts	Sistar				

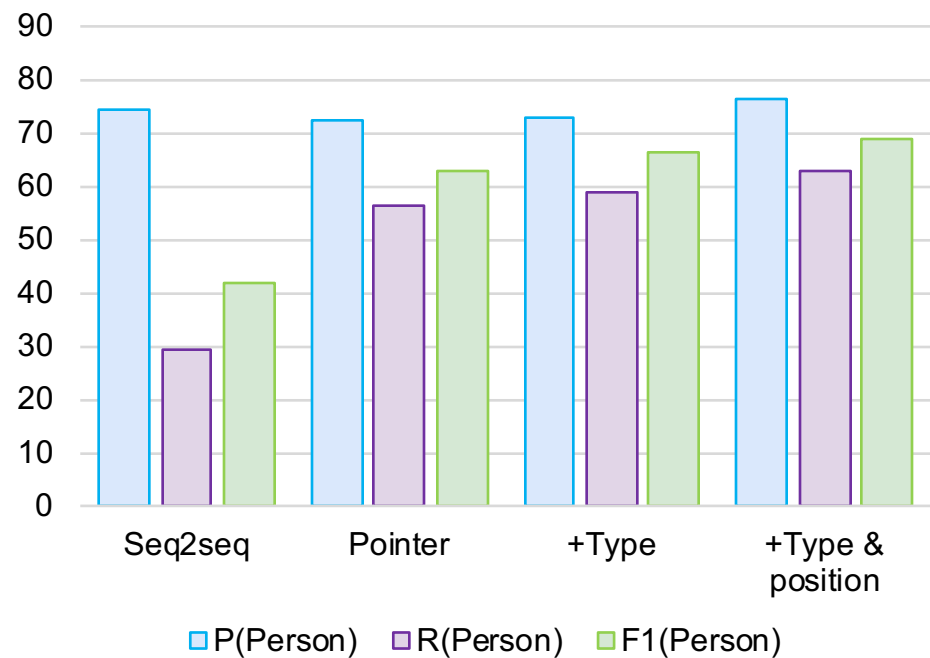


Slot Type	Slot Value
Name	Kim Da-som
Date of Birth	6 May 1993
Place of Birth	Seoul
Occupation	Singer
Associated acts	Sistar
Agent	King Kong by Starship
Start Active Year	3 June 2010

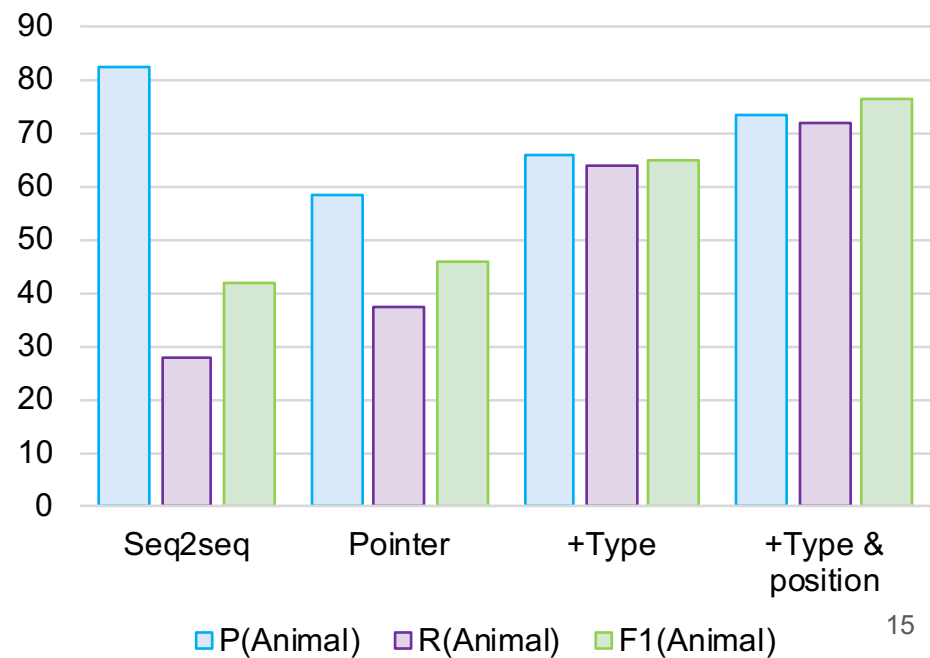
KB Reconstruction based Evaluation (Scores for the example: Overall Slot Filling  $P=6/7=85.7\%$ ,  $R=6/11=54.5\%$ ,  $F1=66.7\%$ ; Inter-dependent Slot Filling  $P=5/7=71.4\%$ ,  $R=5/9=55.6\%$ ,  $F1=62.5\%$ )

# Overall Slot Filling Precision(P), Recall(R), F1 %

## Person

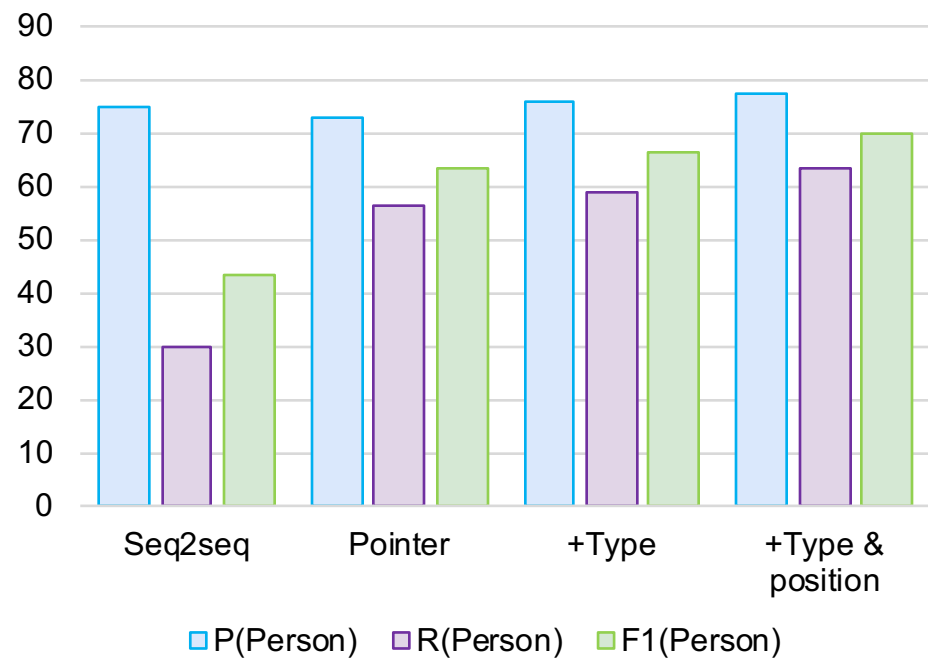


## Animal

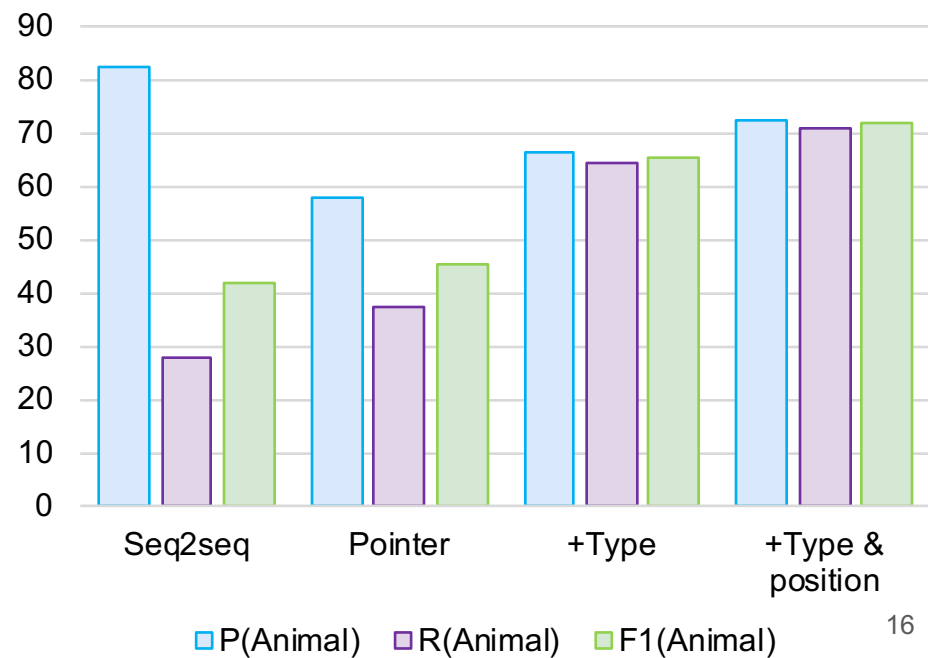


# Interdependent Slot Filling Precision(P), Recall(R), F1 %

## Person



## Animal





# Result of Seq2Seq

(born 23 April 1981) is a retired Israeli footballer. He played for the Thailand 's (scoring one goal) and was a member of the team that won the first ever player in the history of the National Basketball League. She played for the team from 1997 to 2001 scoring 29 goals. She played for the team from 1997 to 2001 scoring 29 goals. She played for the team from 1999 to 2001 and played for the team in the 1997 and 2003 seasons.

Problem: Lack of important entity (e.g. football team)

Solution: Pointer Network

# Result of Pointer Network

**Silvi Jan** the fourth past the Maccabi Holon F.C. (women). On 27 October 1973 in 29 2014) (born 22) is a former **Israel**. She was a **Forward (association football)** and currently plays for **Hapoel Tel Aviv F.C.(women)** in the Swedish league. She played for the **ASA Tel Aviv University** in the Swedish league. She was a member of the **Israel women's national football team** at the beginning of the 2008 season.

Problem: Incorrect slot type (e.g. **Israel**)

Solution: Slot type attention

# Result of Pointer + Type

Silvi Jan (born 27 October 1973) is a former Israeli footballer. He played for Hapoel Tel Aviv F.C.(women) and ASA Tel Aviv University.

Problem: Missing alignment slot (e.g. No. of goals )

Solution: Table position attention

# Result of Pointer + Type & Position

Silvi Jan (born 27 October 1973) is a former Israel. He played for Israel women's national football team, Hapoel Tel Aviv F.C.(women), ASA Tel Aviv University and Maccabi Holon F.C. (women). He was capped 22 times for the Israel women's national football team.

Result: Capture most of slots as well as their dependencies

# Impact of Slot-aware Attention

- Fill the value into correct slot types
  - Pointer result: “*24 September 1979 was born 3 October 1903 in 17 on 33 October 1906*” where 33 should be the number and 17 should be the number of goals
  - Our final result: “*he made 33 appearances and scored 17 goals*”.

# Impact of Table Position Attention

- Capture interdependent slots
  - “*Bill Sampy ... who played for **Sheffield United F.C. 41** times.*”
  - “*Giancarlo Antognoni ... he was also a member of the **Italy national football team** at the **1982 FIFA World Cup.**”*
  - “*Ernő Solymosi ... played **38** games and scored **7** goals for the **Hungary national football team.**”*

# Remaining challenges

- Noun to the adjective form when it appears right before a position slot
  - “**Italy** professional Association football player”
- Low KB Reconstruction recall
- Unmatched temporal expressions
  - “Aleksei Gasilin ( born **1 March 1996** ) is a ... He made his professional debut in the Russian Second Division in **1992** for ...”
- Person genders
  - “Silvi Jan (born 27 October 1973) is a former Israel. **He** played for Israel women’s national football team, ...”
- More details about the entity

# Conclusion

- Provide effective generator for the description of a structured knowledge base
- Propose KB reconstruction based metric to evaluate how many facts are correctly expressed in the generated output
- Achieve KB reconstruction F-scores up to 73%
- Create a large dataset of KBs paired with natural language descriptions for 106,216 entities



# Future Work

- Work on the remaining challenges
- Deal with setting where multiple facts are not in temporal order
- Generate foreign language for a KB
- Generate video/image for a KB



Thank you!

Code: [https://github.com/EagleW/Describing\\_a\\_Knowledge\\_Base](https://github.com/EagleW/Describing_a_Knowledge_Base)

Dataset: [https://drive.google.com/file/d/1TzcNdjZ0EsLh\\_rC1pBC7dU70jINcsVJd/view](https://drive.google.com/file/d/1TzcNdjZ0EsLh_rC1pBC7dU70jINcsVJd/view)