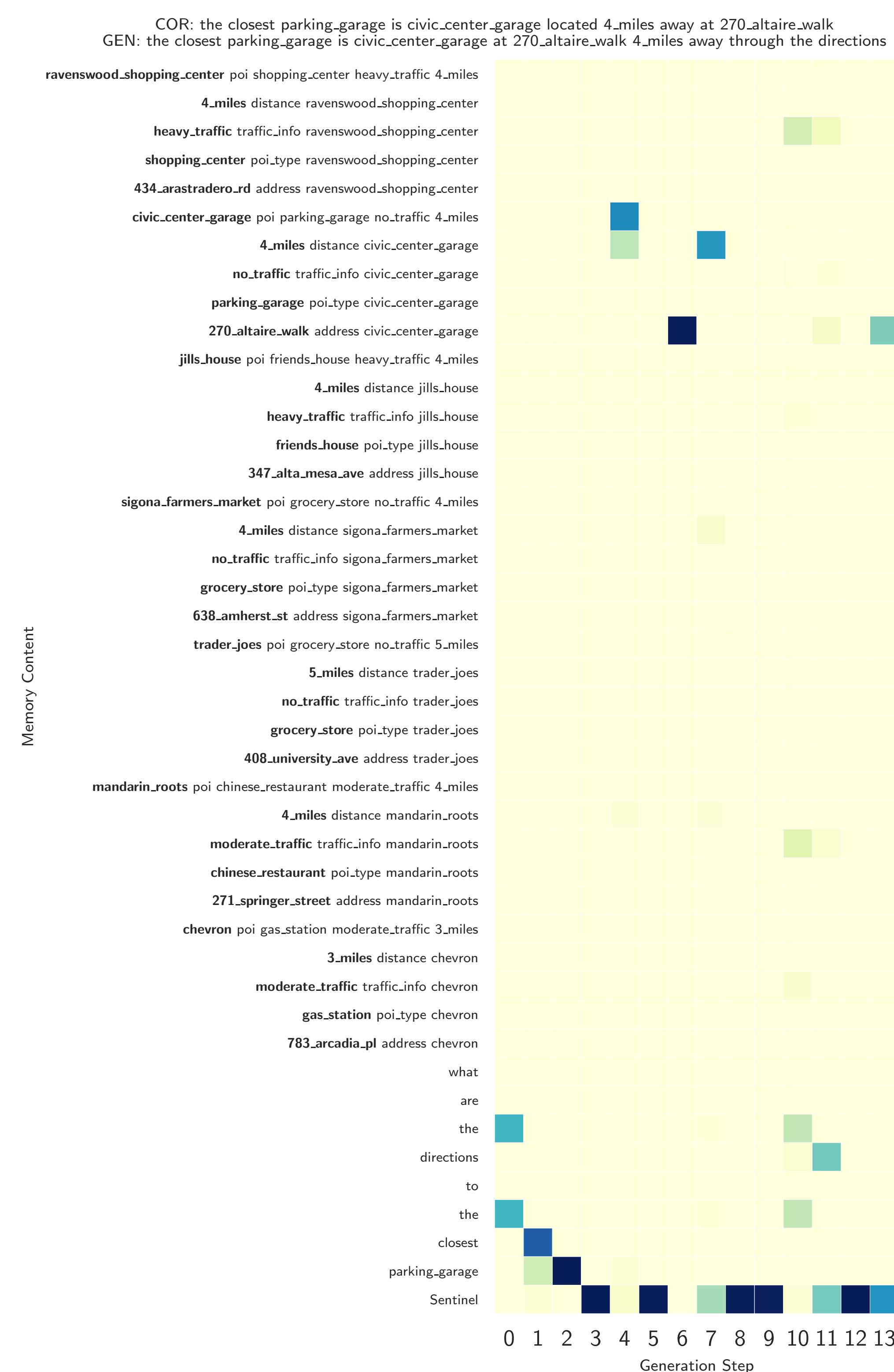


# Mem2Seq: Effectively Incorporating Knowledge Bases into End-to-End Task-Oriented Dialog Systems

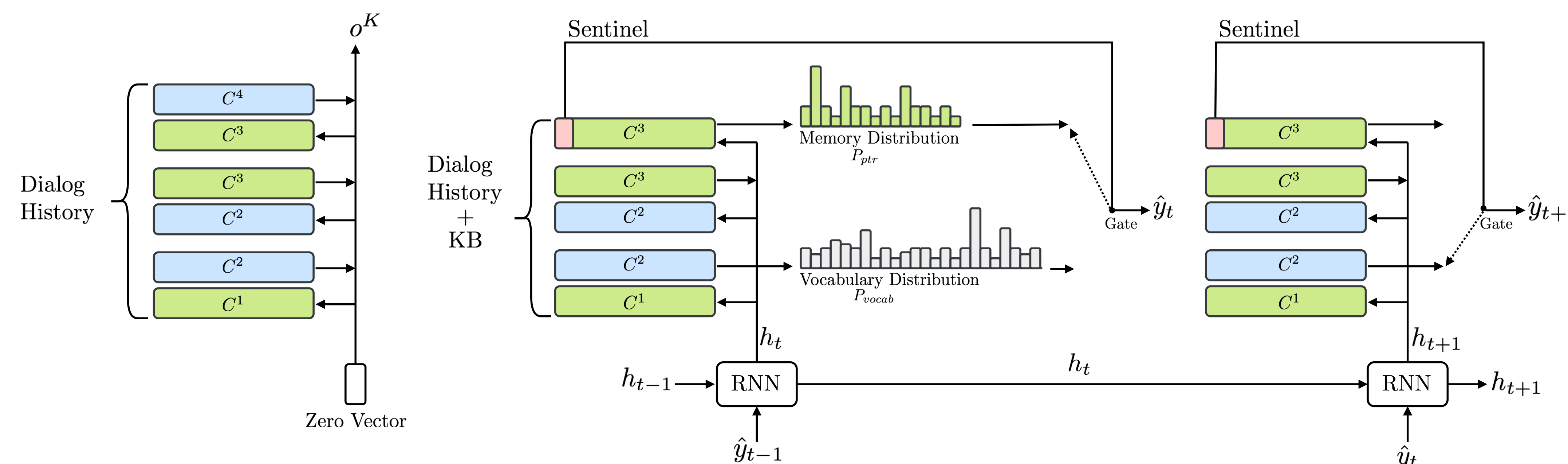
## Introduction

- End-to-end task-oriented dialog systems usually suffer from the challenge of incorporating knowledge bases (KBs).
- Mem2Seq** is the first neural generative model that combines the **multi-hop attention** over memories with the idea of **pointer network**.
- Mem2Seq** can be **trained faster** and attain the **state-of-the-art** performance on three different task-oriented dialog datasets.
- We empirically prove that multi-hop attention mechanism helps in learning **correlations** between memories.
- The model is **general** without complicated task-specific designs.

## Attention Read Out



## Mem2Seq



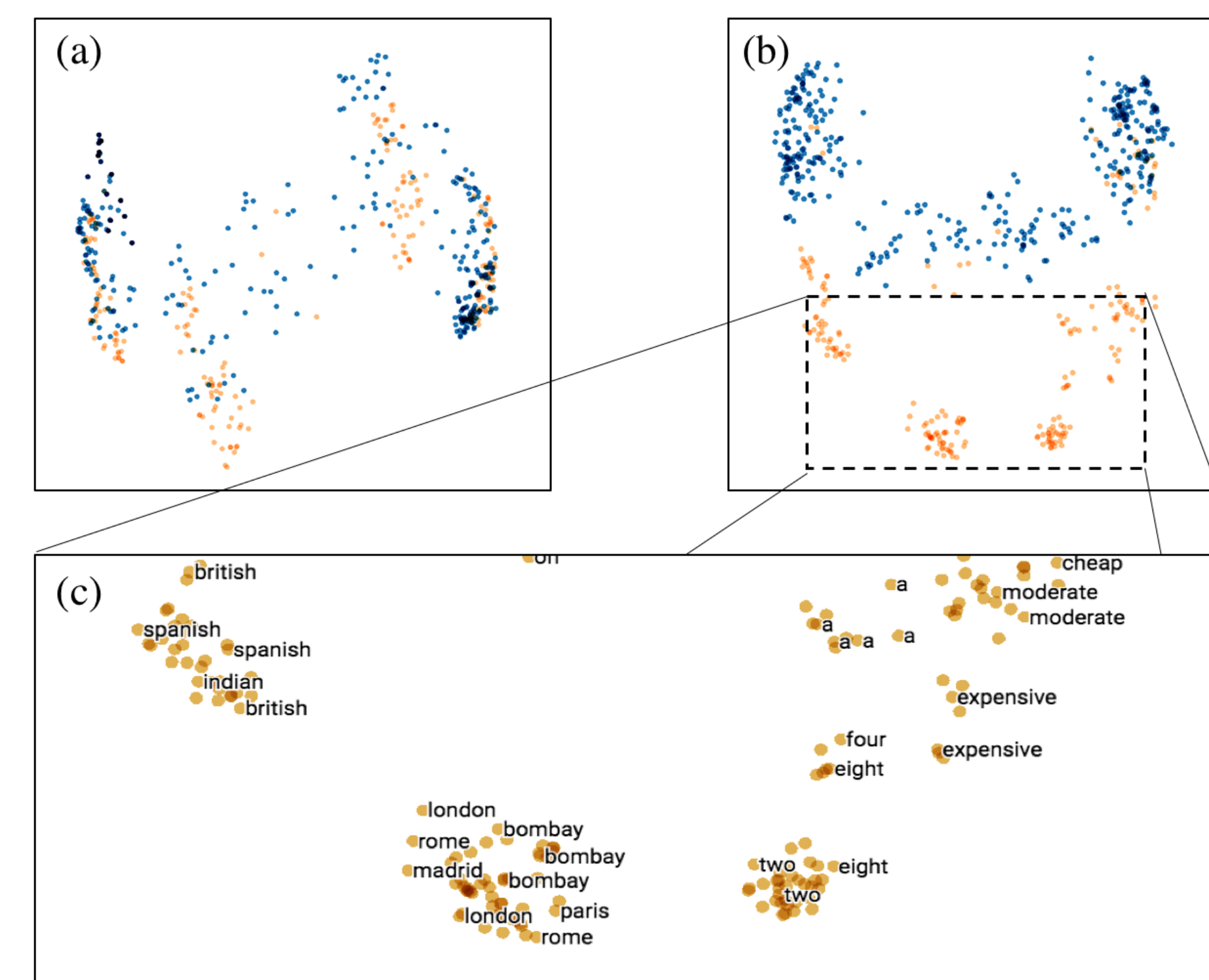
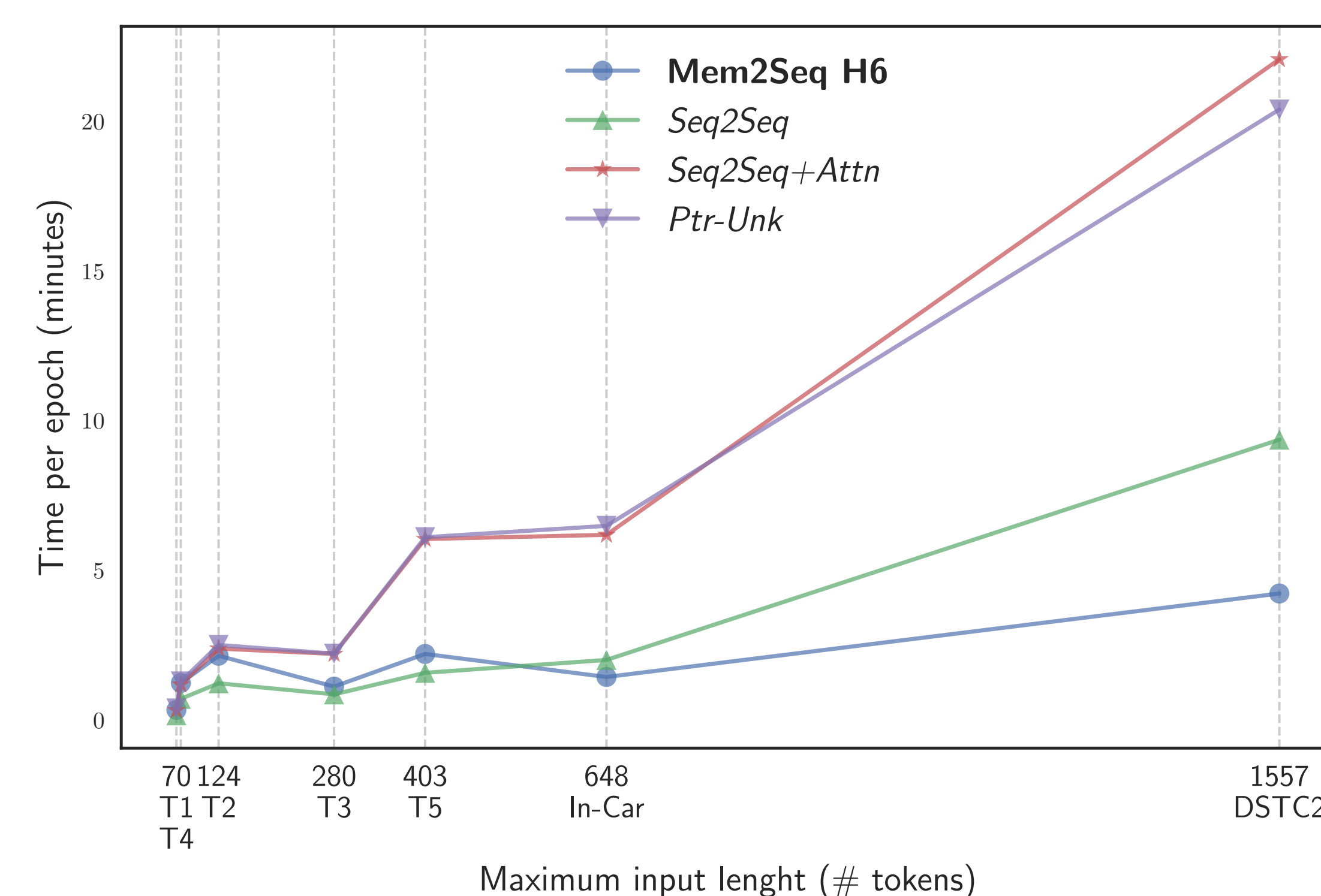
## Results

bAbI	QRN	MemNN	Seq2Seq	Ptr-Unk	Mem2Seq	DSTC-2	Ent. F1	BLEU	Per-Resp.
T1	99.4	99.9	100	100	100	Rule-Based	-	-	33.3
T2	99.5	100	100	100	100	QRN	-	-	43.8
T3	74.8	74.9	74.8	85.1	<b>94.7</b>	MemNN	-	-	41.1
T4	57.2	59.5	57.2	<b>100</b>	<b>100</b>	Seq2Seq	69.7	55.0	46.4
T5	<b>99.6</b>	96.1	98.4	99.4	97.9	+Attn	67.1	<b>56.6</b>	46.0
T1-OOV	83.1	72.3	81.7	92.5	<b>94.0</b>	+Copy	71.6	55.4	<b>47.3</b>
T2-OOV	78.9	78.9	78.9	83.2	<b>86.5</b>	Mem2Seq	<b>75.3</b>	55.3	45.0
T3-OOV	75.2	74.4	75.3	82.9	<b>93.2</b>				
T4-OOV	56.9	57.6	57	<b>100</b>	<b>100</b>				
T5-OOV	67.8	65.5	65.7	73.6	<b>84.5</b>				
In-Car	BLEU	Ent. F1	Sch. F1	Wea. F1	Nav. F1				
Human	13.5	60.7	64.3	61.6	55.2				
Seq2Seq	8.4	10.3	9.7	14.1	7.0				
+Attn	9.3	19.9	23.4	25.6	10.8				
Ptr-Unk	8.3	22.7	26.9	26.7	14.9				
Mem2Seq	<b>12.6</b>	<b>33.4</b>	<b>49.3</b>	<b>33.6</b>	<b>24.6</b>				

- **bAbI** dialogs: we report the per-response and per-dialog accuracy.

- **DSTC2**: Seq2Seq (+attn and +copy) is reported from Eric et. al. (2017).

- **In-Car** Assistant: both BLEU and Entity F1 are improved without using canonical form.



## Methodology

**Mem2Seq** is composed of two components: the MemNN encoder, and the memory decoder.

- The encoder uses a **MemNN** with adjacent weighted tying.
- The decoder uses a **RNN** to generating dynamic query for a **MemNN**.
- At each time step, two distributions are generated using the RNN hidden state  $h_t$ , one over all the words in the **vocabulary**,

$$P_{vocab}(\hat{y}_t) = \text{Softmax}(W_1[h_t; o^1])$$

and one over the **memory contents**

$$P_{ptr}(\hat{y}_t) = p_t^K$$

which  $p_t^K$  is the attention at the last MemNN hop. The next word is selected using an hard gating mechanism (i.e. **sentinel**).

- All the parameters are jointly learned by minimizing the sum of two standard cross-entropy losses.

## Example

Point of interest (poi)	Distance	Traffic info	Poi type	Address
The Westin	5 miles	moderate traffic	rest stop	329 El Camino Real
Round Table	4 miles	no traffic	pizza restaurant	113 Anton Ct
Mandarin Roots	5 miles	no traffic	chinese restaurant	271 Springer Street
Palo Alto Cafe	4 miles	moderate traffic	coffee or tea place	436 Alger Dr
Dominos	6 miles	heavy traffic	pizza restaurant	776 Arastradero Rd
Stanford Express Care	6 miles	no traffic	hospital	214 El Camino Real
Hotel Keen	2 miles	heavy traffic	rest stop	578 Arbol Dr

1<sup>th</sup> Turn

**DRIVER** Where can I get tea?  
**Seq2Seq** I have a away from away would you like the address  
**+Attn** I have a listing for a place that serves tea that is 5 miles away  
**Ptr-Unk** There is a away you like would you like more info

**Mem2Seq** The nearest tea shop is Palo Alto Cafe located 4 miles away, would you like directions there?

**GOLD** Palo Alto Cafe is 4 miles away and serves coffee and tea. Do you want the address?

2<sup>th</sup> Turn

**DRIVER** Yes.  
**Seq2Seq** The is at at  
**+Attn** The address is 329 El Camino Real and it's 3 miles away and there is no traffic  
**Ptr-Unk** Palo Alto Cafe is 4 miles away and PAD is no traffic  
**Mem2Seq** Palo Alto Cafe is 4 miles away at 436 Alger Drive  
**GOLD** Palo Alto is located at 436 Alger Dr.



Scan the QR to find our git-hub repository