Joint Reasoning on Hybrid-knowledge sources for Task-Oriented Dialog

Mayank Mishra, Danish Contractor and Dinesh Raghu
IBM Research

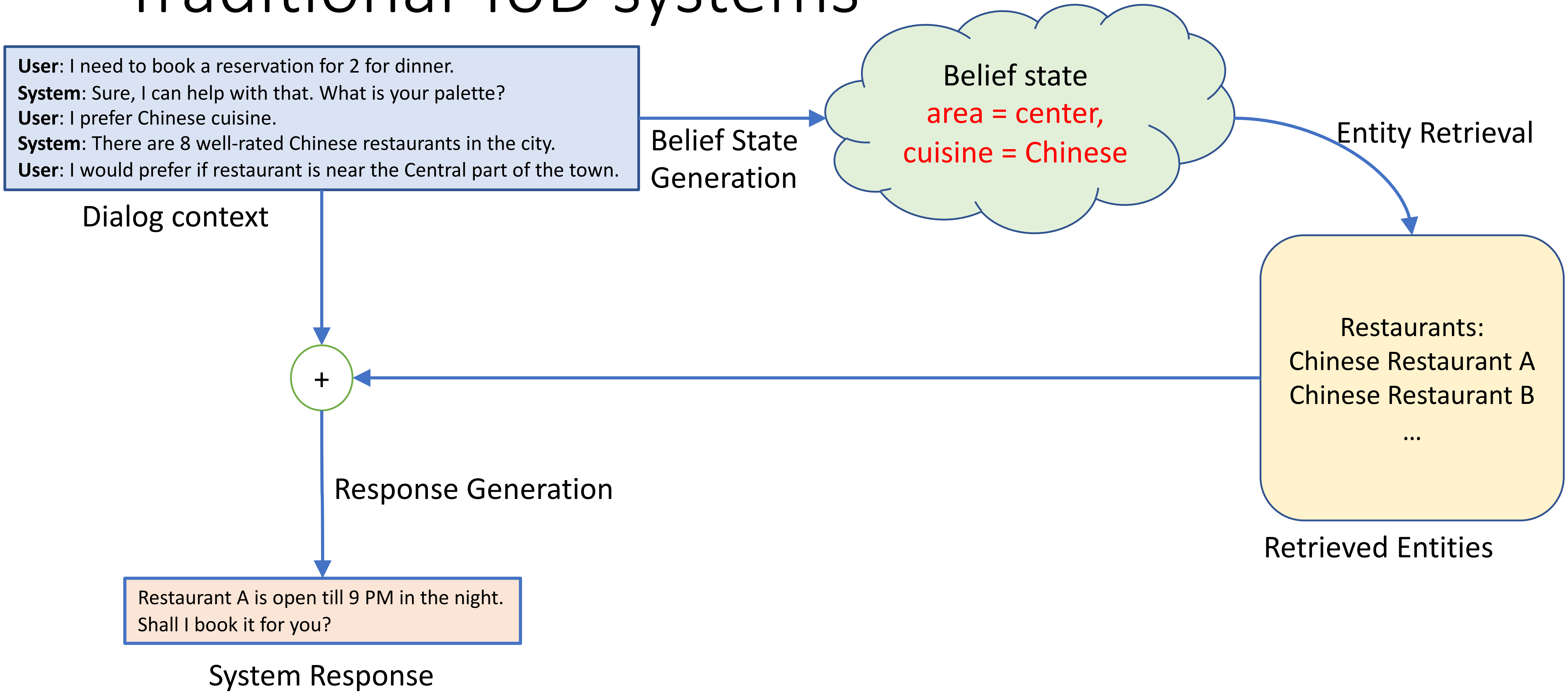
Introduction

- Traditional dialog systems use structured knowledge sources for response generation
- Relevant knowledge might also reside in unstructured knowledge

Terminology

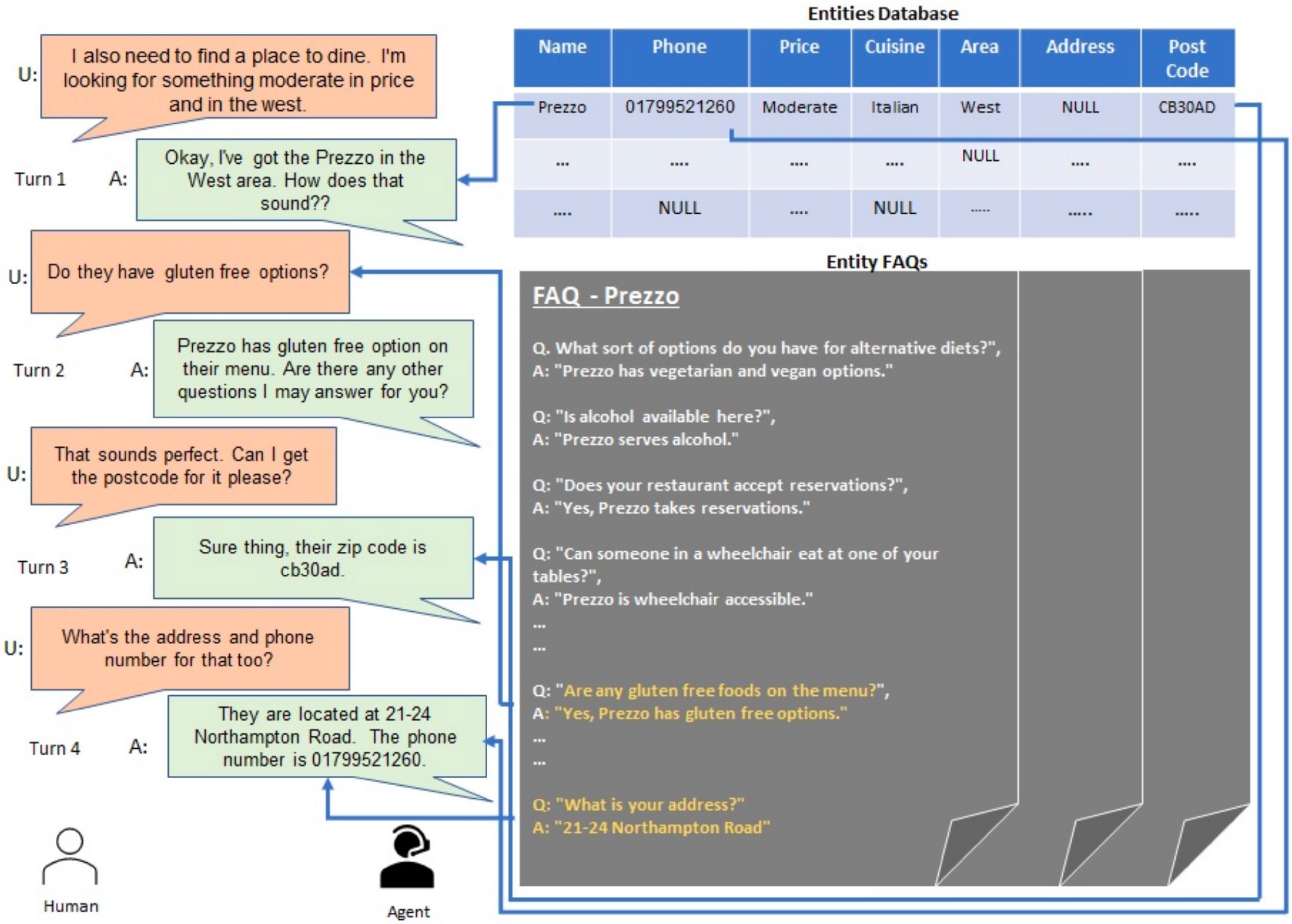
- Structured knowledge: Information about an entity in a tabular format which can be easily queried
- Unstructured knowledge: Information about an entity in documents. For example, menu for a restaurant, customer reviews etc
- Entity: All the information about a restaurant/hotel etc

Traditional ToD systems



Limitations of current methods

- Recent SOTA models make limiting assumptions for the said problem
- Traditional dialog systems use structured knowledge sources for response generation
- Relevant knowledge might also reside in unstructured knowledge


Entities Database
Name
Phone
Price
Cuisine
Area
Address
Post Code
Prezzo
01799521260
Moderate
Italian
West
NULL
CB30AD
NULL
NULL
NULL
Entity FAQs
FAQ - Prezzo
Q. What sort of options do you have for alternative diets?",
A: "Prezzo has vegetarian and vegan options."
Q: "Is alcohol available here?",
A: "Prezzo serves alcohol."
Q: "Does your restaurant accept reservations?",
A: "Yes, Prezzo takes reservations."
Q: "Can someone in a wheelchair eat at one of your tables?",
A: "Prezzo is wheelchair accessible."
Q: "Are any gluten free foods on the menu?",
A: "Yes, Prezzo has gluten free options."
Q: "What is your address?"
A: "21-24 Northampton Road"
U: I also need to find a place to dine. I'm looking for something moderate in price and in the west.
Turn 1
A: Okay, I've got the Prezzo in the West area. How does that sound??
U: Do they have gluten free options?
Turn 2
A: Prezzo has gluten free option on their menu. Are there any other questions I may answer for you?
U: That sounds perfect. Can I get the postcode for it please?
Turn 3
A: Sure thing, their zip code is cb30ad.
U: What's the address and phone number for that too?
Turn 4
A: They are located at 21-24 Northampton Road. The phone number is 01799521260.
Human
Agent

Our contributions

- Release a new version of DSTC-9 MultiWoZ dataset (called HybridToD) which is devoid of the mentioned limitations
- Propose a model that fuses information from both structured and unstructured sources to generate the response

HybridToD dataset

- We use the SeKnow-MultiWoZ dataset and modify it to remove the aforementioned limitations
- We construct a graph G(V, E):
  - Each vertex in V is a unique slot-value pair
  - An edge exists between 2 vertices if the slot-values represented by these vertices occur in the same utterance in the training dataset
- We find a maxcut of this graph which ensures that we modify the maximum number of utterances in the dataset
- We move entities from one of the resulting halves of maxcut from structured source to unstructured knowledge

HybridToD dataset

	Context-Response pairs			Number of entities
Domain	train	validation	test	train/validation/test
hotel	19370	2316	2295	33
restaurant	19716	2162	2188	110
attraction	8192	1226	1246	79
total	47278	5704	5729	222

Table 2: Number of context-response pairs in the dataset

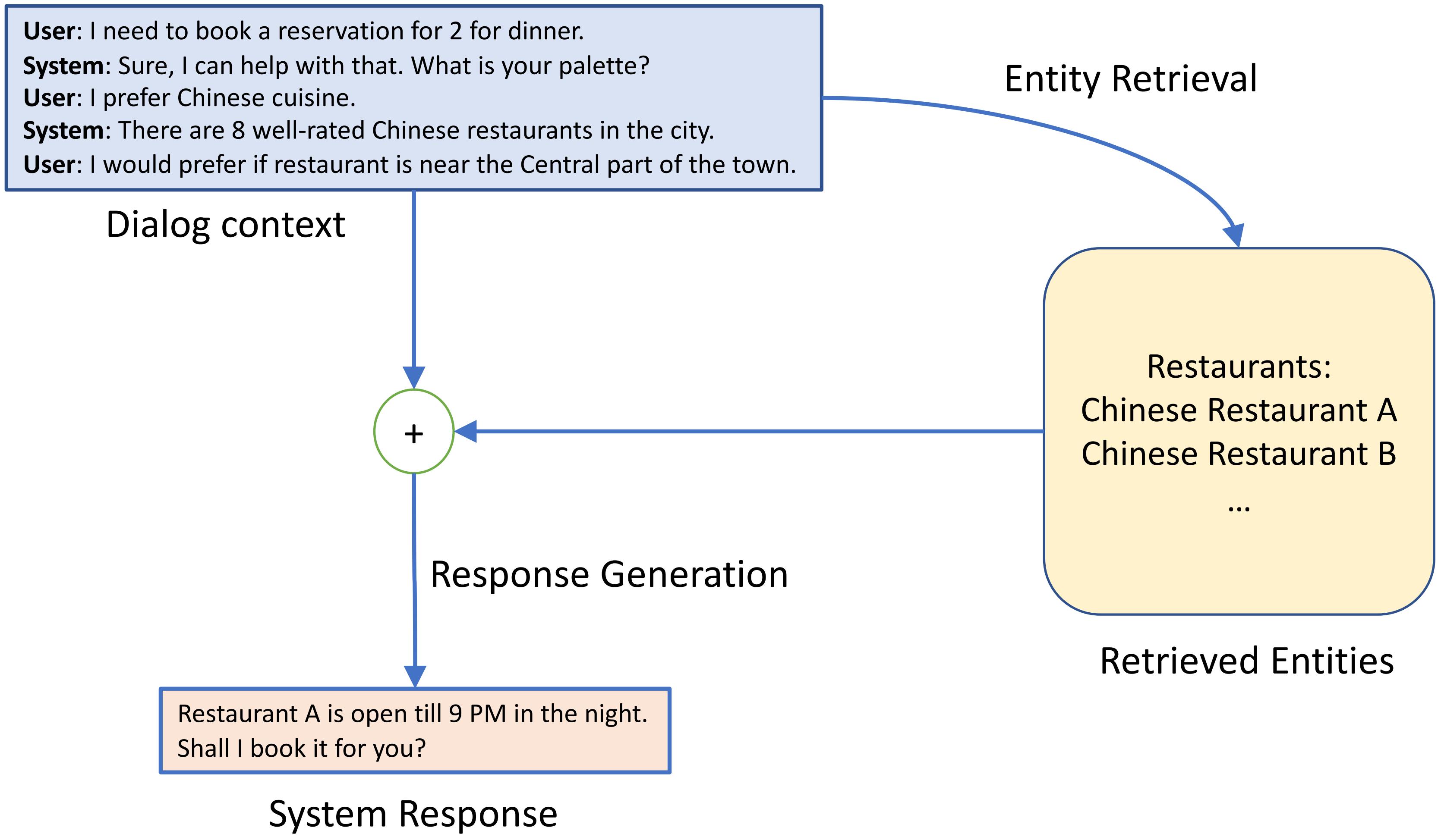
Domain	SEKNOW-MULTIWOZ	HYBRIDTOD
hotel	10.97	6.79
restaurant	8.12	5.25
attraction	9	6.38

Table 3: Average number of slot values by domain in the structured knowledge source for each dataset.

Domain	SEKNOW-MULTIWOZ	HYBRIDTOD	UNSTRUCTUREDTOD
hotel	36.52	40.58	46.48
restaurant	14.96	17.83	22.7
attraction	0	2.62	8

Table 4: Average number of FAQs for each domain in the unstructured knowledge source.

Our model


User: I need to book a reservation for 2 for dinner.
System: Sure, I can help with that. What is your palette?
User: I prefer Chinese cuisine.
System: There are 8 well-rated Chinese restaurants in the city.
User: I would prefer if restaurant is near the Central part of the town.
Dialog context
Entity Retrieval
Restaurants:
Chinese Restaurant A
Chinese Restaurant B
…
Retrieved Entities
+
Response Generation
Restaurant A is open till 9 PM in the night.
Shall I book it for you?
System Response

Our model

- We train a BART model for both the entity retrieval and response generation tasks
- We score each of the entities using the BART model and take the best one
- The best entity is used along with the dialog context to generate the response

$$[e^s] = \langle struct \rangle \, \langle slot \rangle \, slot_1 \, \langle val \rangle \, value_1$$
$$\langle slot \rangle \, slot_2 \, \langle val \rangle \, value_2 ...$$
$$[e^{us}] = \langle unstruct \rangle \, \langle doc \rangle \, document_1$$
$$\langle doc \rangle \, document_2 ...$$
$$[e] = [e^s] \, [e^{us}]$$

Entity Representation

$$\langle entity\_retrieval\_task \rangle \, \langle u \rangle \, u_1 \, \langle r \rangle \, r_1 ...$$
$$\langle u \rangle \, u_n \, \langle entity \rangle \, [e_j]$$

Input for Entity Retrieval

$$\langle response\_task \rangle \, \langle u \rangle \, u_1 \, \langle r \rangle \, r_1 ...$$
$$\langle u \rangle \, u_n \, \langle entity \rangle \, [e]$$

Input for Response Generation

Experiments

- We train all the models on HybridToD and evaluate on SeKnow-MultiWoZ, HybridToD and a completely unstructured knowledge source
- We also show that training on SeKnow-MultiWoZ leads to a model memorizing the source of information leading to sub-optimal performance when the knowledge modality is changed at test time
- We compare our models on response generation performance to the SOTA baseline SeKnow

Results

					slot-values		
Train Dataset	**Test Dataset**	**Model**	**Bleu-1**	**Bleu-4**	**prec.**	**recall**	**F1**
HYBRIDTOD	SEKNOW-MULTIWOZ	JOINTLM	**30.63**	**8.60**	**50.48**	**45.37**	**47.79**
		SEKNOW	29.20	7.83	43.16	28.65	33.14
HYBRIDTOD	HYBRIDTOD	JOINTLM	**30.59**	**8.67**	**50.56**	**45.83**	**48.08**
		SEKNOW	29.05	7.70	44.29	29.12	35.14
HYBRIDTOD	UNSTRUCTUREDTOD	JOINTLM	**30.30**	**8.44**	**51.05**	**45.37**	**48.04**
		SEKNOW	27.43	6.68	42.96	19.62	27.11

Table 5: All models trained on HYBRIDTOD and evaluated on the rest of the datasets

					slot-values		
Train Dataset	**Test Dataset**	**Model**	**Bleu-1**	**Bleu-4**	**prec.**	**recall**	**F1**
SEKNOW-MULTIWOZ	SEKNOW-MULTIWOZ	JOINTLM	29.07	8.06	49.74	41.31	45.13
		SEKNOW	**31.00**	**9.14**	**52.17**	**44.98**	**48.31**
SEKNOW-MULTIWOZ	HYBRIDTOD	JOINTLM	**27.77**	**7.54**	**44.48**	**36.39**	**40.03**
		SEKNOW	26.61	7.32	42.19	26.70	33.31
SEKNOW-MULTIWOZ	UNSTRUCTUREDTOD	JOINTLM	**27.03**	**7.17**	**46.29**	**34.93**	**39.82**
		SEKNOW	26.19	6.42	41.96	19.48	26.53

Table 6: All models trained on HYBRIDTOD and evaluated on the rest of the datasets

Conclusion

- In this paper, we have created and released a new dataset for advancing the research in ToD systems which require both structured and unstructured information
- Through extensive experiments, we have also shown that existing approaches don’t work well with changing slot-value distribution at inference time
- We also propose a model which shows superior performance to SeKnow and is also more robust to changing slot-value distribution at inference time

Thank you