Highlights

- We propose the task of **dialogue state induction** (DSI), which is to **automatically induce** dialogue states from raw dialogue data.
- We introduce two **neural latent variable models** for DSI called DSI-base and DSI-GM.
- Review: This paper focuses on an important topic and has **great potential of motivating follow-up research**.

Dialogue State Induction (DSI)

Input: A set of customer service records without annotation.
Target: Automatically discover information that the user is looking for at each turn (dialogue states).
Difference from dialogue state tracking (DST): DSI task does not rely on manual labeling or even ontology and can generate slot-value pairs over raw dialogues automatically.

User: I want an expensive restaurant that serves Turkish food.
Manual labeling:
`inform(price=expensive, food=Turkish)`

System: Anatolia serves Turkish food.
User: What is the area?
Manual labeling:
`inform(price=expensive, food=Turkish); request(area)`

Ontology(optional):
price: [cheap, expensive, moderate, ...]
food: [Turkish, Italian, polish, ...]
area: [south, north, center, ...]

DST

Turn 1:
`inform(price=expensive, food=Turkish)`

Turn 2:
`inform(price=expensive, food=Turkish); request(area)`

User: I want an expensive restaurant that serves Turkish food.
~~Manual labeling:~~
~~`inform(price=expensive, food=Turkish)`~~

System: Anatolia serves Turkish food.
User: What is the area?
~~Manual labeling:~~
~~`inform(price=expensive, food=Turkish); request(area)`~~

~~**Ontology(optional):**~~
~~price: [cheap, expensive, moderate, ...]~~
~~food: [Turkish, Italian, polish, ...]~~
~~area: [south, north, center, ...]~~

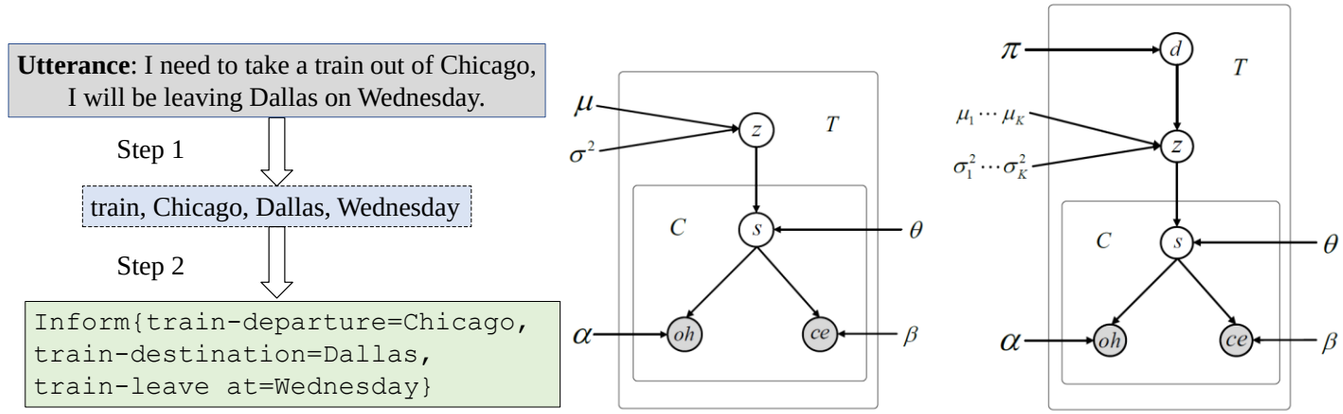
DSI

Turn 1:
`inform(price=expensive, food=Turkish)`

Turn 2:
`inform(price=expensive, food=Turkish); request(area)`

Method

Two steps:
- Candidate (values) extraction (POS tag, NER, coreference)
- Slot assignment: two neural latent variable models (DSI-base and DSI-GM)



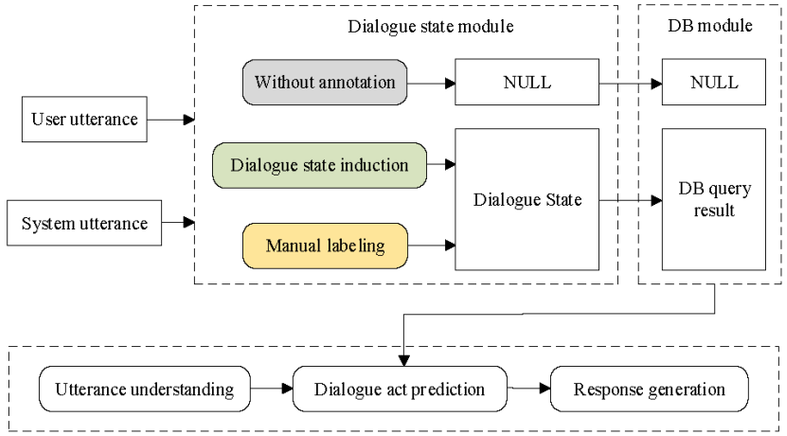
(a) *DSI-base* (b) *DSI-GM*

- The whole state and each slot are treated as latent variables, from which values observed in dialogue data are generated.
- Both models are generative probabilistic models, which generate a value by first generating a latent dialogue state vector, and then generating a slot.
- DSI-GM further considers the service domain explicitly by taking the dialogue state as a mixture of Gaussians.

Results

Models	MultiWOZ 2.1								SGD							
	Turn level				Joint level				Turn level				Joint level			
	Precision	Recall	F1	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1	Accuracy
Random	1.49	1.51	1.49	1.39	0.21	0.28	0.23	0.02	0.94	0.95	0.94	0.92	0.05	0.08	0.06	0.02
DSI-base	38.8	37.7	37.3	25.7	33.9	32.1	32.1	2.3	27.0	26.0	26.0	21.1	13.9	17.5	14.5	2.3
DSI-GM	52.5	39.3	49.6	36.1	49.2	43.2	44.8	5.0	34.7	33.4	33.5	27.5	19.0	22.9	19.5	3.1

Table 1: Overall results of DSI.



Dialogue State	Dialog Act Prediction			Delexicalized	
	Precision	Recall	F1	BLEU	Entity F1
None	71.0	67.4	69.1	18.7	54.6
DSI-GM	72.0	70.5	71.2	20.8	56.5
Manual labeling	75.6	73.0	74.2	21.6	61.3

paper GitHub