Knowledge-Retrieval Task-Oriented Dialog Systems with Semi-Supervision

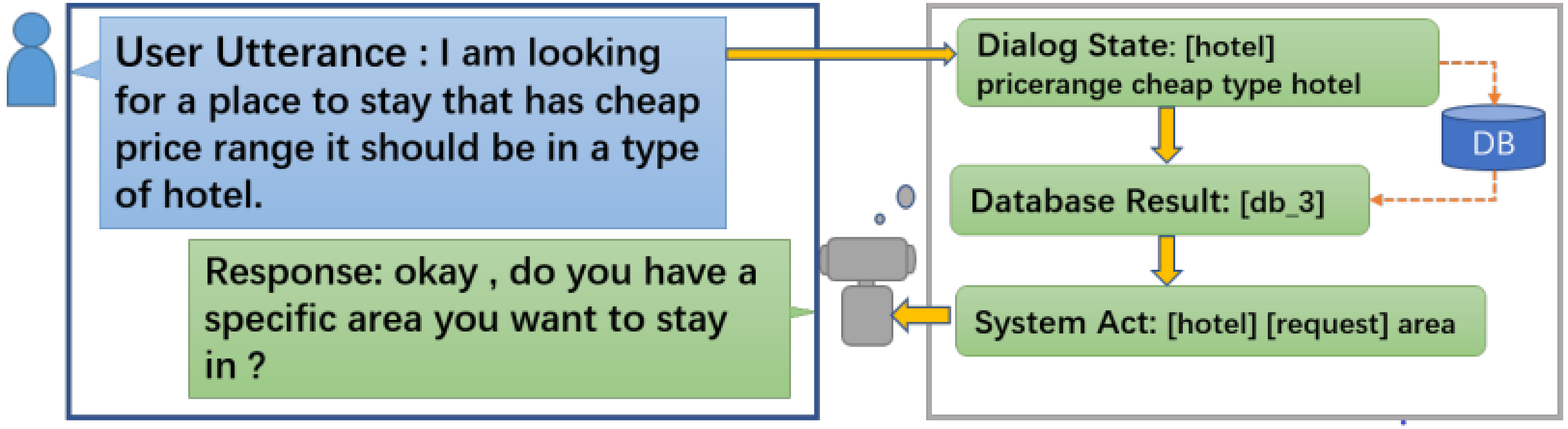
Yucheng Cai, Hong Liu, Zhijian Ou*
Speech Processing and Machine Intelligence Lab
Tsinghua University, Beijing, China
cyc22@mails.tsinghua.edu.cn
liuhong21@mails.tsinghua.edu.cn
ozj@tsinghua.edu.cn

Yi Huang, Junlan Feng
China Mobile Research Institute
Beijing, China
{huangyi,fengjunlan}
@chinamobile.com

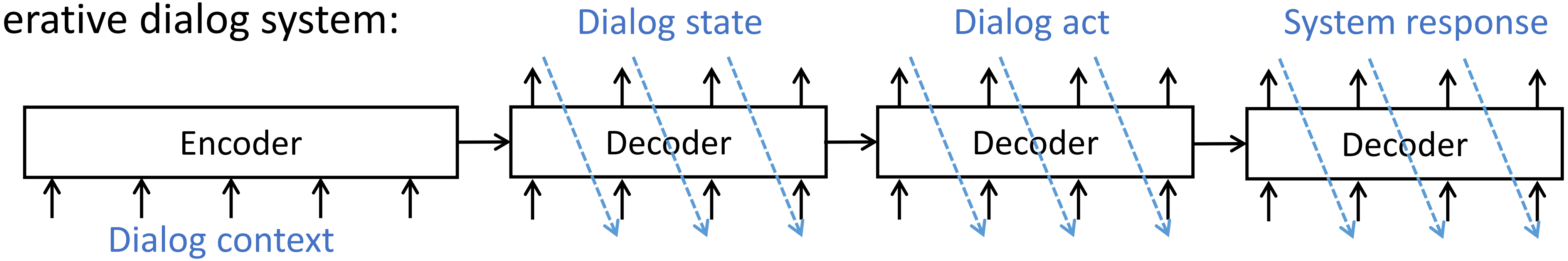
Outline

- Motivation
- Related Work
- Method
- Experiments
- Conclusion

The traditional information flow of a TOD system: dialog state tracking (DST), database querying (DB), policy (POL), and response generation (NLG).

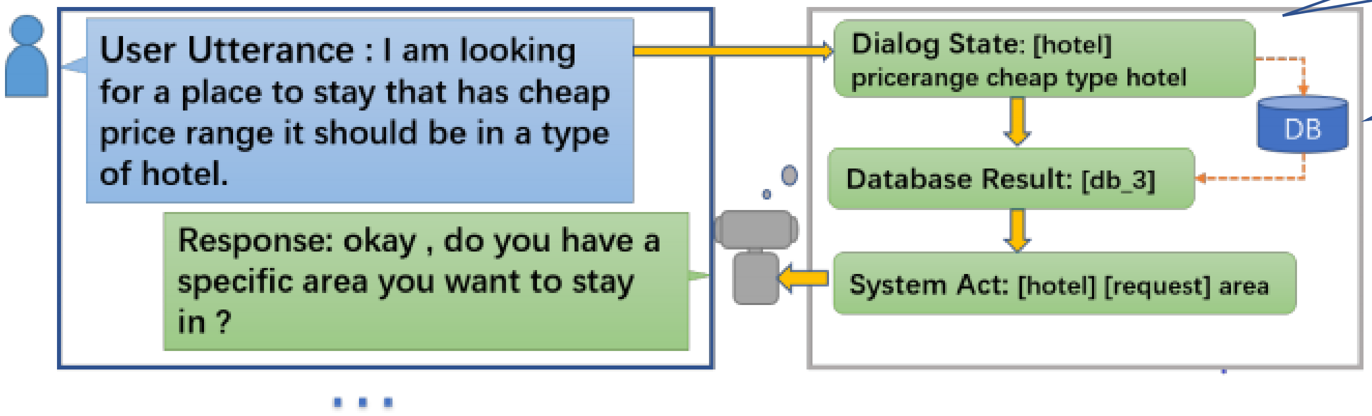


Generative dialog system:



Y. Zhang, Z. Ou, Z. Yu. "Task-Oriented Dialog Systems that Consider Multiple Appropriate Responses under the Same Context", AAAI 2020.

Introuduction (Our Improvements)

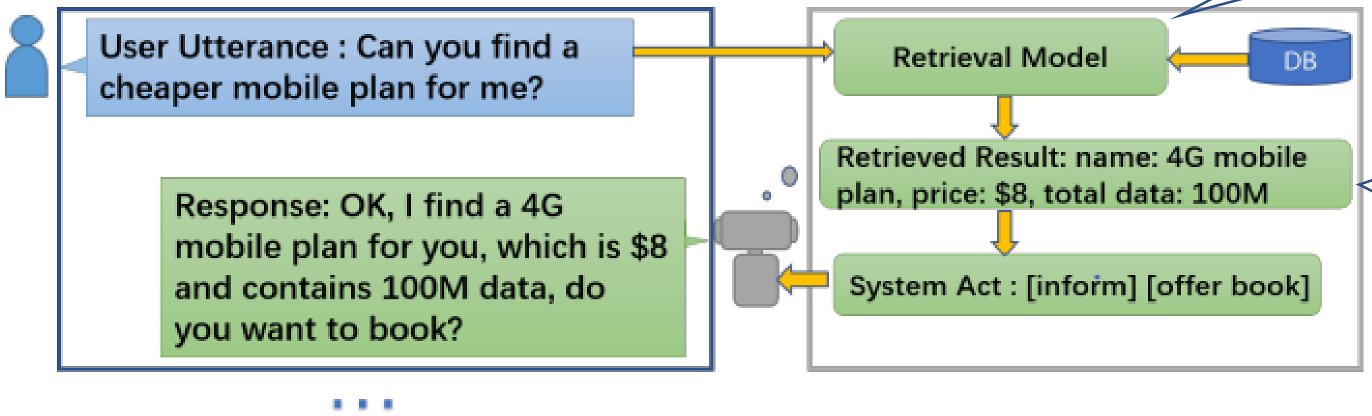


(a) Traditional TOD system

Difficulty in correctly tracking dialog state

Inflexibility of rule-based database query

- **Knowledge retriever**: retrieve appropriate external knowledge based on dialog context
- Avoid tracking of dialog states



(b) Our KRTOD system

- Obtaining labels of ground-truth knowledge is expensive
- We further develop **latent-variable model based semi-supervised learning**

Outline

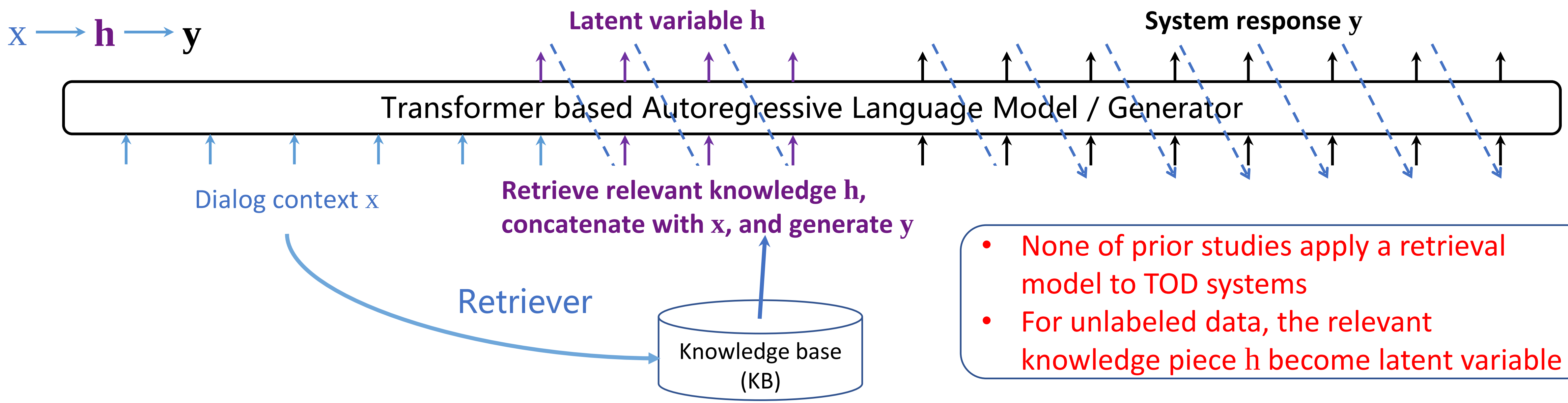
- Motivation
- Related Work
- Method
- Experiments
- Conclusion

Related Work

1. Knowledge Retriever for Conditional Generation in dialog systems
2. Semi-Supervised Learning (SSL) in dialog systems

Related Work (Knowledge Retrieval)

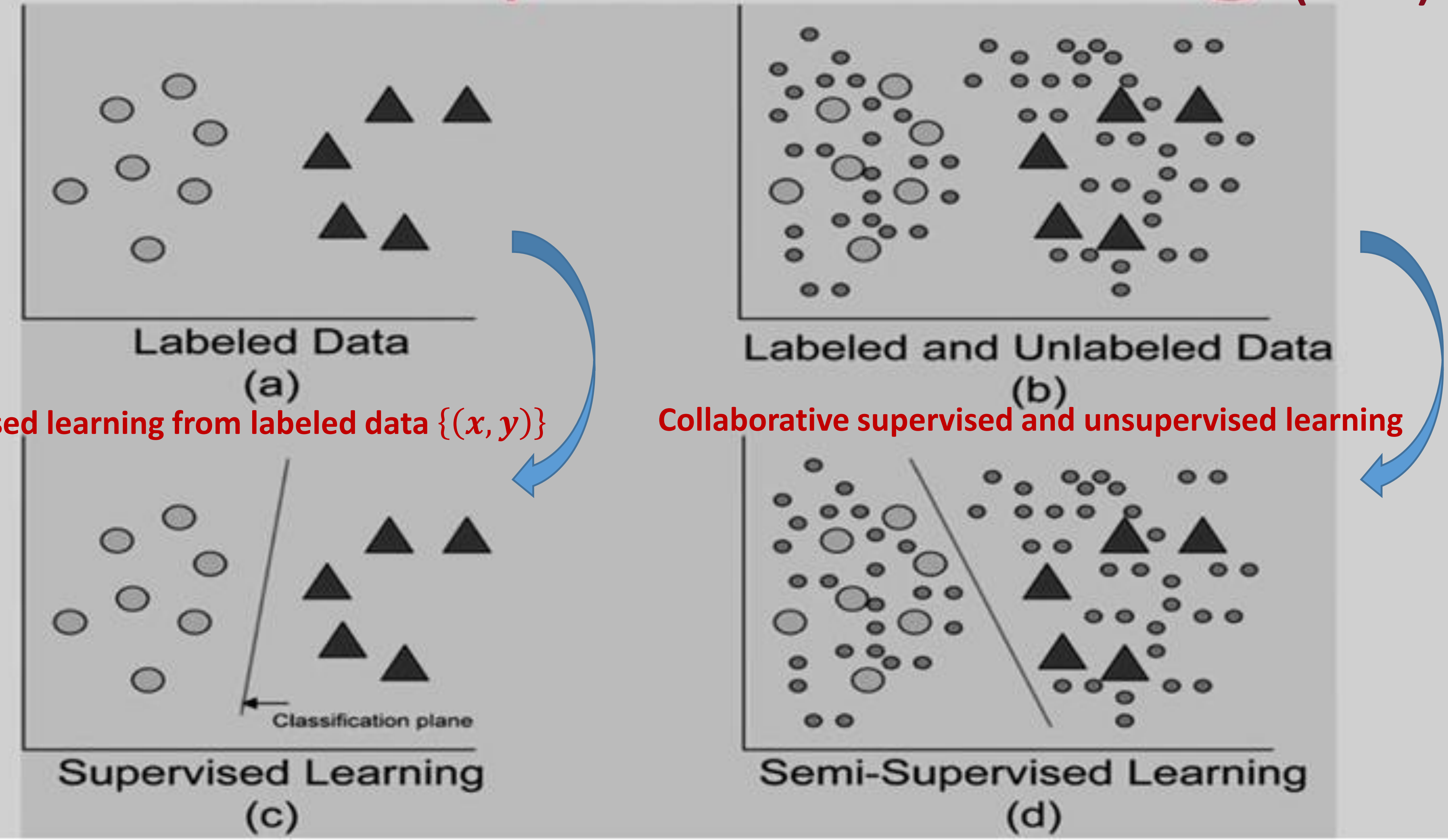
- Pure generative dialog systems
  - Without any access to an external knowledge base
  - Their ability to access and precisely manipulate knowledge is limited
- For open-domain question answering and knowledge-grounded dialog systems, recent studies such as RAG (Retrieval Augmented Generation):
  - Introduce a knowledge retriever model into conditional generation



- None of prior studies apply a retrieval model to TOD systems
- For unlabeled data, the relevant knowledge piece h become latent variable

P. Lewis, et al., "Retrieval-Augmented Generation (RAG) for Knowledge-Intensive NLP Tasks", NeurIPS 2020.

Semi-supervised learning (SSL)



Supervised learning from labeled data $\{(x, y)\}$

Collaborative supervised and unsupervised learning

Related work: Two approaches for semi-supervised dialog systems

Pre-training (serial collaboration)

- Unsupervised pre-training followed by supervised fine-tuning
- Large-scale language models , like GPT-x, pre-trained on open-domain texts are fine-tuned with in-domain labels

Joint-training (parallel collaboration)

- Formulate a latent variable model (**LVM**) of observations and labels
- Unsupervised learning with LVM usually maximizes marginal likelihood via variational learning over unlabeled data

- Remarkably, the two approaches, are not exclusive to take and can be jointly used, and, complement each other.
- Joint stochastic approximation (JSA) performs better than variational learning, particularly for discrete latent variable models.

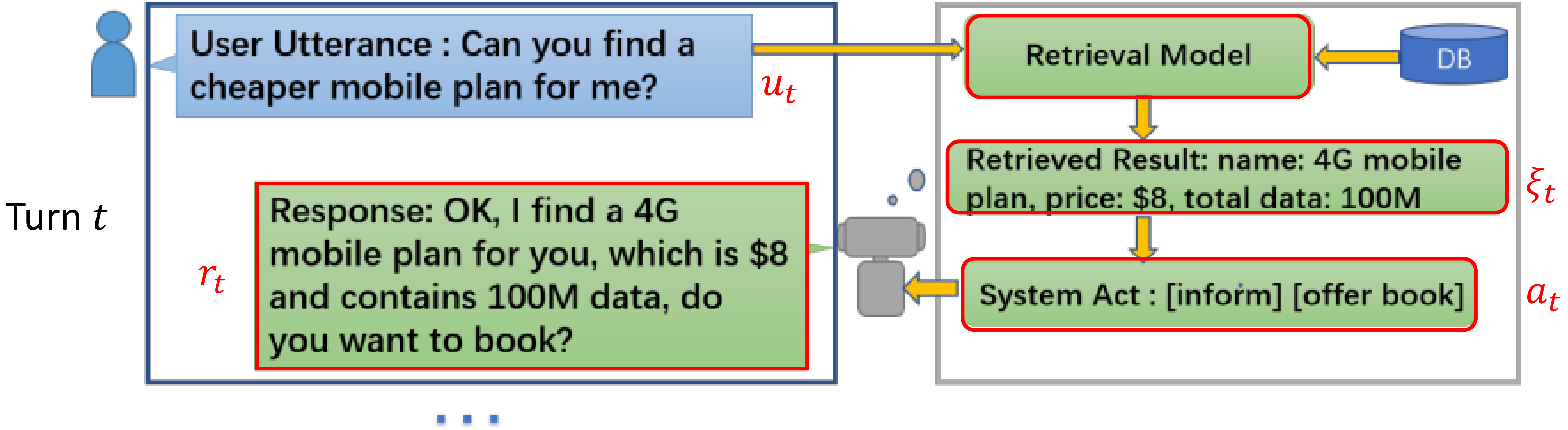
- Zhang, Y.; Ou, Z.; et al, "A Probabilistic End-To-End Task-Oriented Dialog Model with Latent Belief States towards Semi-Supervised Learning". EMNLP, 2020.
- Song, Y., et al, "An empirical comparison of joint-training and pre-training for domain-agnostic semi-supervised learning via energy-based models", IEEE Workshop on Machine Learning for Signal Processing (MLSP), 2021.
- Y. Cai, et al., "Advancing semi-supervised task oriented dialog systems by JSA learning of discrete latent variable models", SIGDIAL 2022.

Outline

- Motivation
- Related Work
- Method (JSA-KRTOD)
- Experiments
- Conclusion

Method (Notation)

- Consider a dialog with T turns of user utterances and system responses, denoted by $u_1, r_1, \cdots, u_T, r_T$ respectively
- The KB is composed of entities with attributes, or say, slot-value pairs, denoted by $\{\mathrm{sv}^1, \cdots, \mathrm{sv}^N\}$
- Relevant knowledge piece: the slot-value pairs that are relevant for the system to respond at turn t are denoted by ξ_t
- At turn t, $\{\xi_t, a_t\} \triangleq h_t$, collectively denoted as the latent variable. In labeled data, h_t is observed, while in unlabeled dialogs, it becomes a latent variable.



Method (A latent-variable dialog model)

Given context and current user utterance, retrieve knowledge, make action, generate response

$$p_\theta(h_{1:T}, r_{1:T} | u_{1:T}) = \prod_{t=1}^{T} p_\theta(h_t, r_t | c_t, u_t)$$

For dialog turn t:
u_t - user utterance
r_t - system response
$h_t = \{\xi_t, a_t\}$
ξ_t - relevant knowledge piece
a_t - system act

where $c_t = u_1, r_1, \cdots, u_{t-1}, r_{t-1}$ denotes the dialog context at turn t

- The joint model of h_t, r_t is decomposed into a knowledge retriever p_θ^{ret} and a response generator p_θ^{gen}

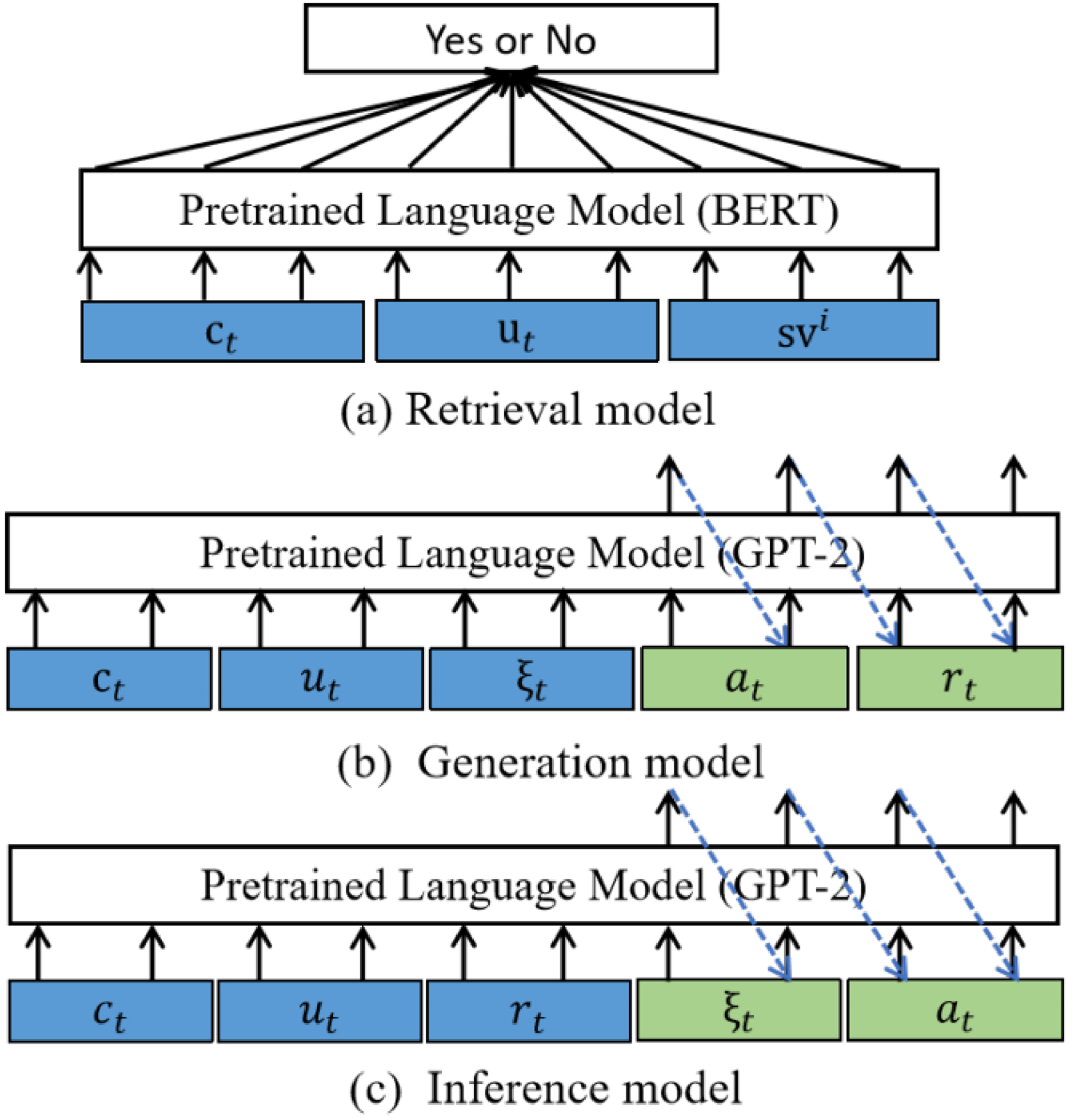
$$p_\theta(h_t, r_t | c_t, u_t) = p_\theta^{\text{ret}}(\xi_t | c_t, u_t) \times p_\theta^{\text{gen}}(a_t, r_t | c_t, u_t, \xi_t)$$

- Introduce an inference model $q_\phi(h_{1:T} | u_{1:T}, r_{1:T})$ as follows to approximate the true posterior

$$q_\phi(h_{1:T} | u_{1:T}, r_{1:T}) = \prod_{t=1}^{T} q_\phi(h_t \,| c_t, u_t, r_t) = \prod_{t=1}^{T} q_\phi(\xi_t, a_t | c_t, u_t, r_t)$$

Method (Model Implementation)

- Retrieval model $p_\theta^{\text{ret}}(\xi_t|c_t,u_t)$:
  - judge which knowledge to retrieve
- Generation model $p_\theta^{\text{gen}}(a_t,r_t|c_t,u_t,\xi_t)$:
  - use retrieved knowledge to generate response
- Inference model $q_\phi(\xi_t,a_t|c_t,u_t,r_t)$:
  - help infer knowledge in semi-supervised learning



Method (Implementation)

- **Supervised training:**
  - ➢ train retriever with cross-entropy loss
  - ➢ train generater use next-token prediction loss
  - ➢ Use the ground-truth knowledge label
- **Semi-Supervised training:**
  - ➢ generate knowledge with inference model
  - ➢ use JSA algorithm
- **Testing:**
  - ➢ retrieve knowledge first
  - ➢ use retrieved knowledge to generate

Algorithm 1 Semi-supervised training in JSA-KRTOD

Input: A mix of labeled and unlabeled dialogs.

```
Run supervised pre-training of θ and ϕ on labeled dialogs;
repeat
    Draw a dialog (u_{1:T}, r_{1:T});
    if (u_{1:T}, r_{1:T}) is not labeled then
        Generate h_{1:T} using the recursive turn-level MIS sampler
    end if
    J_θ = 0, J_ϕ = 0;
    for i = 1, ··· , T do
        J_θ += log p_θ^gen(a_t, r_t | c_t, u_t, ξ_t);
        J_ϕ += log q_ϕ(ξ_t, a_t | c_t, u_t, r_t);
    end for
    Update θ by ascending: ∇_θ J_θ;
    Update ϕ by ascending: ∇_ϕ J_ϕ;
until convergence
return θ and ϕ
```

Y. Cai, et al., “Advancing semi-supervised task oriented dialog systems by JSA learning of discretelatent variable models”, SIGDIAL 2022.

Related Work (Joint Stochastic Approximation)

JSA（Joint Stochastic Approximation）
= Expectation Maximization (EM) + Stochastic Approximation (SA) + Adaptive MCMC

SAEM (Delyon et al., 1999):

- Monte Carlo sampling: fill the missing values for latent variables through sampling $h' \sim p_\theta(h|x)$
- SA updating: perform gradient ascent over θ using $\nabla_\theta log p_\theta(x, h')$

sampling from $p_\theta(h|x)$ is intractable

JSA = coupling an SA version of EM (SAEM) with an adaptive MCMC procedure (UAI 2020)

- $q_\phi(h|x)$ acts like an adaptive proposal, using Metropolis Independent Sampling (MIS) to sample $q_\theta(h|x)$
- Jointly optimizing $q_\phi(h|x)$ with $p_\theta(x, h)$

Implementing JSA: Propose, Accept or Reject and Optimize (PARO)

Zhijian Ou, et.al, "Joint stochastic approximation and its application to learning discrete latent variable models", UAI, 2020.

Outline

- Motivation
- Related Work
- Method (JSA-KRTOD)
- Experiments
- Conclusion

Experiments (setup)

- Dataset: MobileCS (China Mobile Customer Service)
  - A real-life human-human dialog dataset, instead of collected by Wizard-of-Oz
  - Released from EMNLP 2022 SereTOD Workshop (Towards Semi-supervised and Reinforced Task-Oriented Dialog Systems) and Challenge http://seretod.org/
  - A total of 100K dialogs, 10% annotated

- Experiments in both labeled-only and semi-supervised settings (over both labeled and unlabeled data) can be conducted and fairly compared.
- Backbone: GPT2

Table 1: Comparison of our MobileCS corpus to MultiWOZ

Metric	MultiWOZ	MobileCS	
		labeled	unlabeled
Dialogs	8,438	8,975	87,933
Turns	113,556	100,139	972,573
Tokens	1,490,615	3,991,197	39,491,883
Avg. turns per dialog	13.46	11.16	11.06
Avg. tokens per turn	13.13	39.86	40.61
Slots	24	26	-
Values	4,510	14,623	-

- Z. Ou et.al, “A challenge on semi-supervised and reinforced task oriented dialog systems” in *arXiv preprint* , 2022.
- H. Liu et.al, “Information extraction and human-robot dialogue towards real-life tasks: A baseline study with the mobilecs dataset”, in *EMNLP2022 SereTOD Workshop* .

Experiments (main result)

- Evaluation: end2end
- Metrics
  - *Success rate*: how often the system is able to provide all the entities and values requested by the user
  - *BLEU*: measure the fluency of the generated responses
  - *Combined score* = Success + 2*BLEU

Baseline: uses predicted dialog state to query KB (***KB-query***)

Top three teams in the SereTOD Challenge

- PRIS: concatenates the whole local KB to the dialog history (***KB-grounded***);
- TJU-LMC: uses coarse-to-fine intent detection;
- Passion: improves prompting scheme

Table 1: *Main results on the MobileCS dataset. Success, BLEU-4, and combined score are reported. Our approach achieves SOTA results on both labeled-only and semi-supervised settings. Within the parentheses show the backbone models and their number of parameter.*

Setting	Method	Success	BLEU-4	Combined
Semi-supervised	Baseline [18]	31.5	4.170	39.84
	Passion [30]	43.2	6.790	56.78
	TJU-LMC [29]	68.9	7.54	83.98
	PRIS [19]	78.9	14.51	107.92
	JSA-KRTOD	91.8	9.677	**111.15**
Labeled-only	KB-query (GPT2 100M) [18]	31.5	4.170	39.84
	KB-grounded (GPT2 100M) [19]	64.2	8.845	81.89
	KB-grounded (T5 1B) [19]	74.1	11.32	96.74
	KRTOD (GPT2 100M)	86.8	8.639	**104.08**

JSA-KRTOD greatly outperforms **KB-query** and **KB-grounded**, especially in *Success*, breaking record in MobileCS!

Experiments (ablation)

- Comparison of JSA with pseudo labeling (PL)
  - JSA-KRTOD outperforms PL constantly in all ratios.
  - The relative improvement of JSA over PL in reducing errors in Success rate is 23% under ratio 9:1.

Table 2: *Comparison between pseudo labeling (PL) and JSA learning methods. Ratio means the ratio between the number of unlabeled dialogs and the number of labeled dialogs in training.*

Ratio	Method	Success	BLEU-4	Combined	p-value
1:1	PL	87.5	8.853	105.21	0.589
	JSA	88.0	8.713	**105.43**	
2:1	PL	87.8	9.196	106.19	0.853
	JSA	88.7	9.490	**107.68**	
4:1	PL	88.5	9.341	107.18	0.037
	JSA	90.9	9.398	**109.70**	
9:1	PL	89.4	9.532	108.46	0.055
	JSA	91.8	9.677	**111.15**	

The p-values from matched-pairs significance tests in Combined score show:

- As the size of unlabeled data increases, the improvements of JSA-KRTOD over PL become more significant,
- Confirm the superiority of JSA-KRTOD in leveraging unlabeled data.

Outline

- Motivation
- Related Work
- Method (JSA-KRTOD)
- Experiments
- Conclusion

Conclusion

- Introduce a knowledge retriever, instead of the traditional database query method, which improves the knowledge acquisition in TOD systems.
- Propose to use the JSA algorithm to perform semi-supervised learning for KRTOD systems.
- Extensive experiments conducted on MobileCS, a real-life dialog dataset, show that JSA-KRTOD achieves SOTA results on MobileCS in both labeled-only and semi-supervised settings.
- Future work: JSA-KRTOD potentially can exploit more different types of knowledge sources, such as passages, documents and knowledge graphs, in addition to slot-value pairs used in this work.

Thanks!

Code released at

https://github.com/thu-spmi/JSA-KRTOD