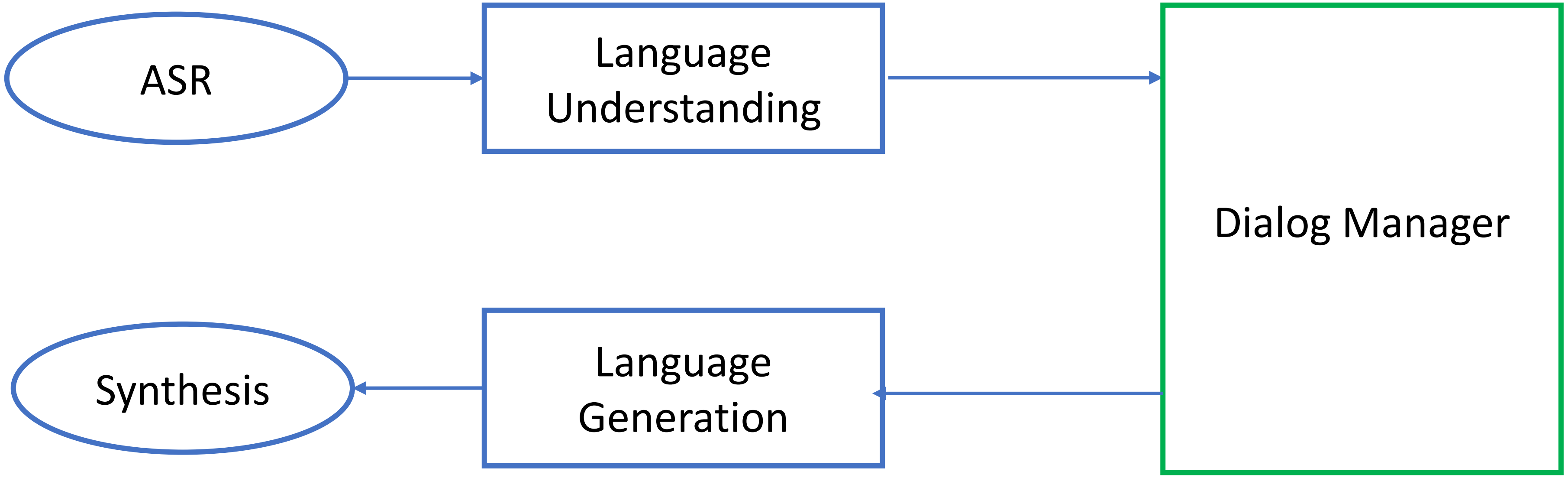
Building Dialog Systems with Less Supervision

Zhou Yu

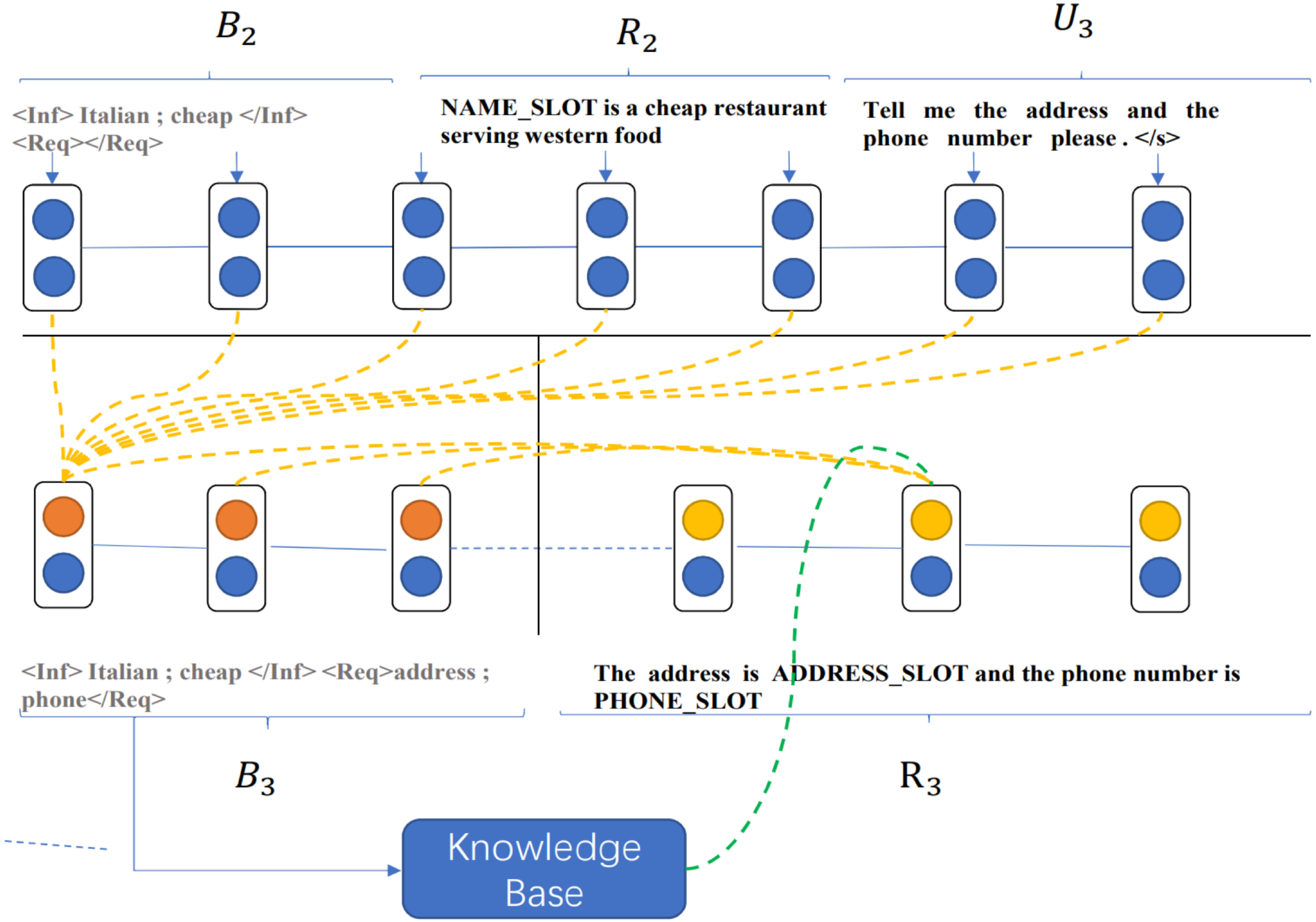
UC Davis

Dialog System: Modular Framework



Dialog System: End-to-End Framework

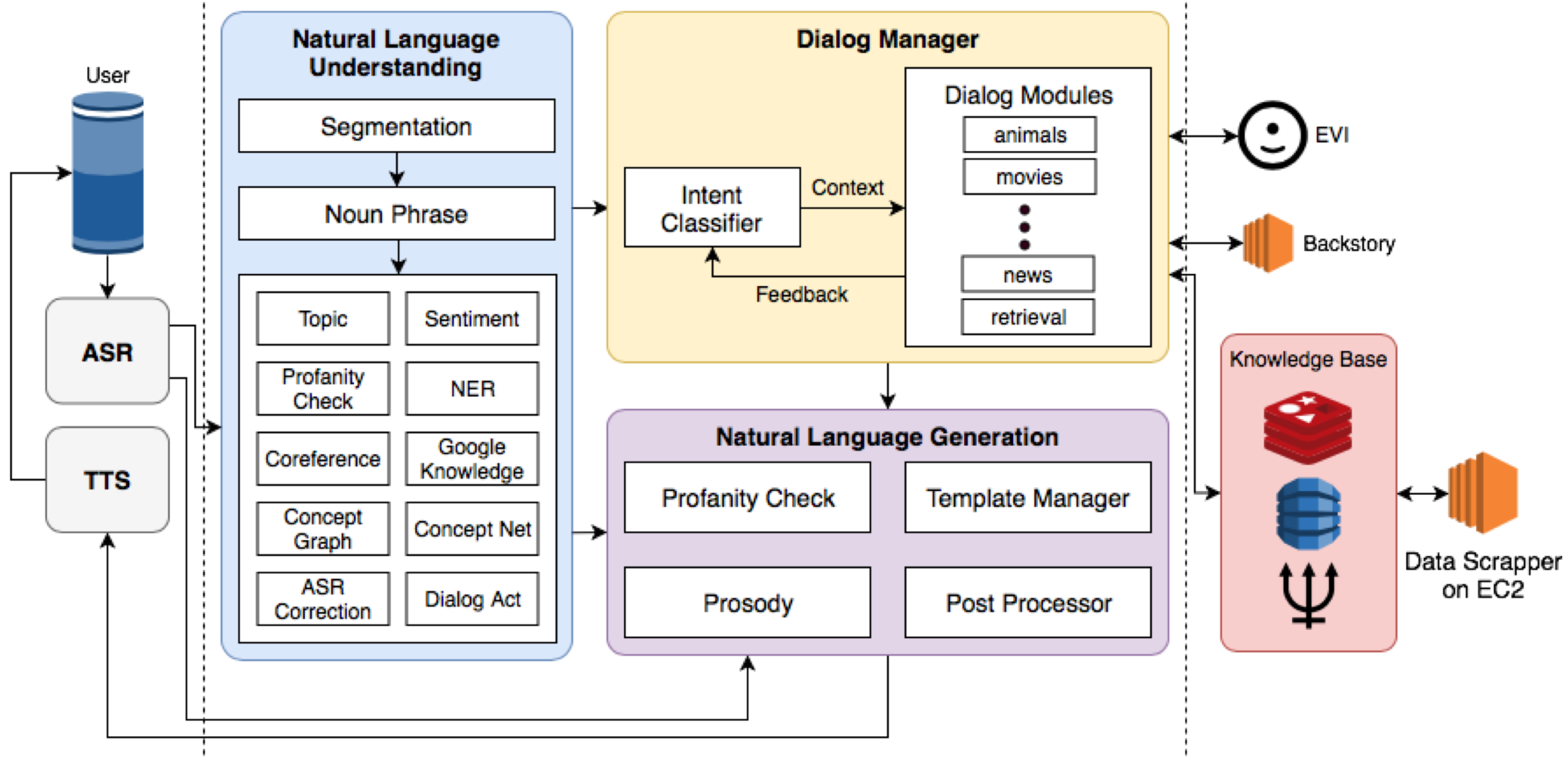
Turn	Dialogue
$User_1$	Can I have some Italian food please?
$Machine_1$	<Inf> Italian </Inf><Req> </Req> What price range are you looking for?
$User_2$	I want cheap ones.
$Machine_2$	<Inf> Italian ; cheap </Inf> <Req></Req> NAME_SLOT is a cheap restaurant serving western food
$User_3$	Tell me the address and the phone number please .
$Machine_3$	<Inf> Italian ; cheap </Inf> <Req>address ; phone</Req> The address is ADDRESS_SLOT and the phone number is PHONE_SLOT



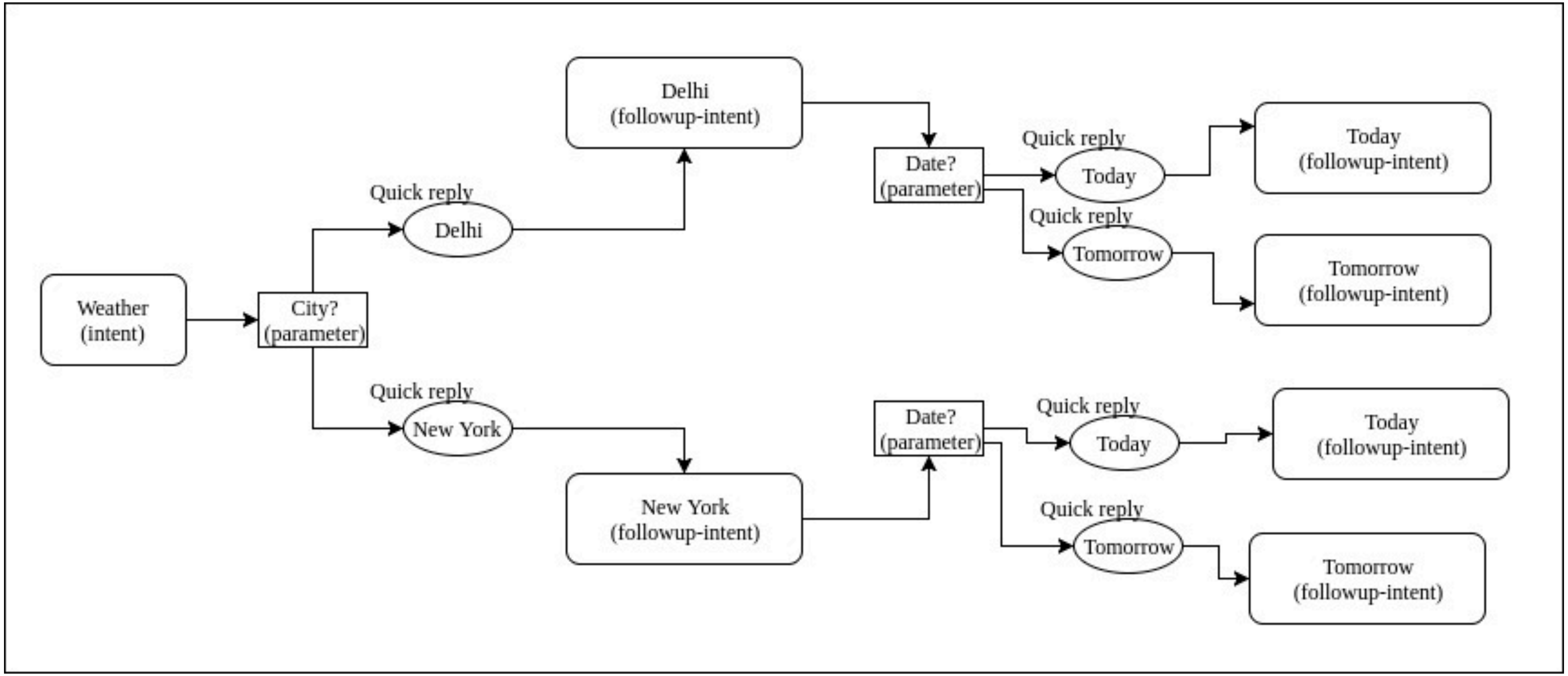
– $B_t = \text{Seq2seq}(B_{t-1} R_{t-1} U_t)$

– $R_t = \text{Seq2seq}(B_{t-1} R_{t-1} U_t | B_t, \text{KB search results})$

Gunrock: Alexa Prize Social Bot Winner 2018



Industry Dialog Flow



Unsupervised Dialogue Structure Learning

NAACL 2019

Goal: Extract dialogue structure represented by latent variable z_t.

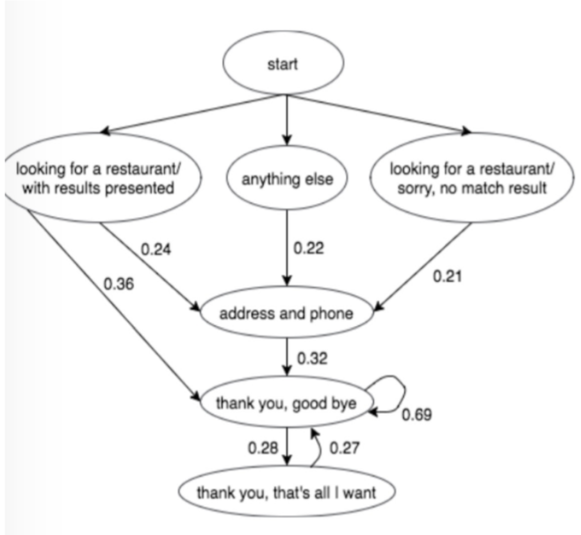
Method: Variational Inference with Discrete-Variational-RNN.

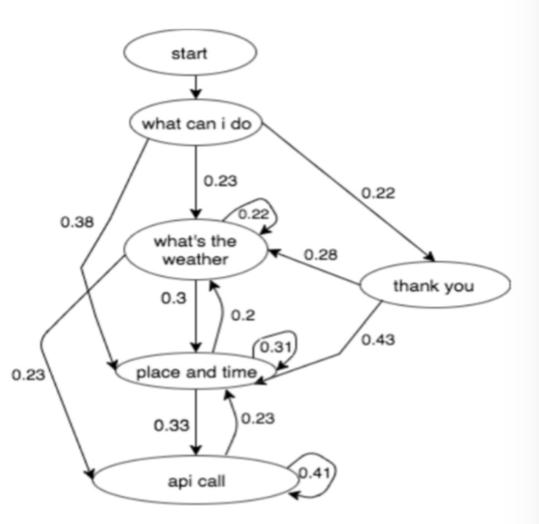
Qualitative Evaluation: visualize the learned dialogue structure

Quantitative Evaluation: the reconstruction probability of the test set.

Comparison with K-means, HMM and Model Variants

Application in RL: Transition probability between states as reward





Dialogue Structure by HMM

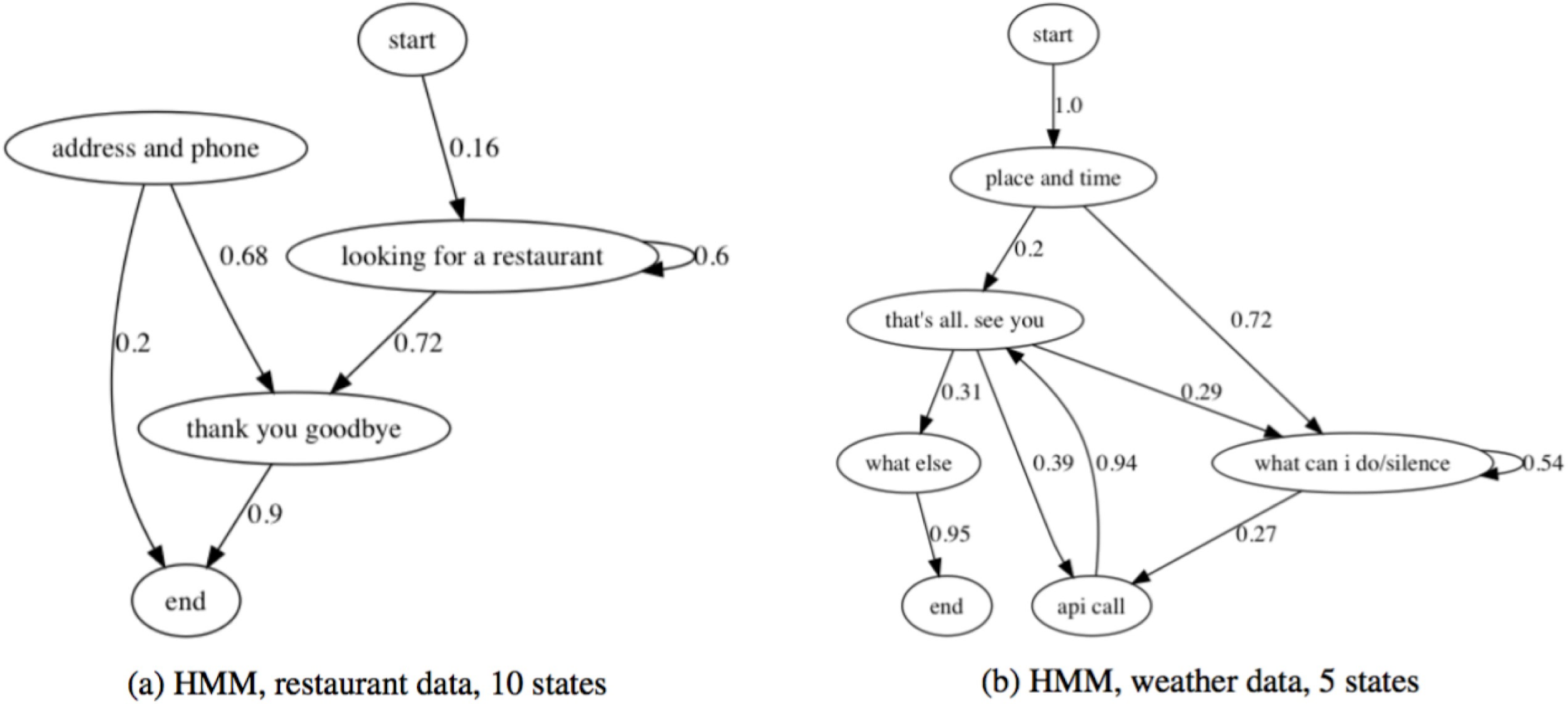
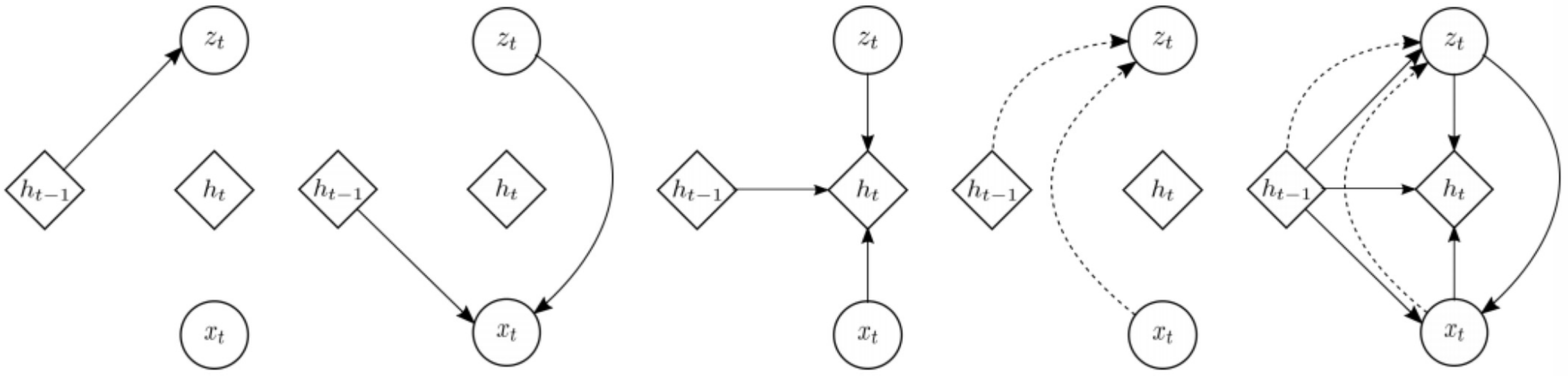


Figure 7: Dialog structures generated by HMM on different datasets.

Variational-RNN

Chung et al. NIPS 2015

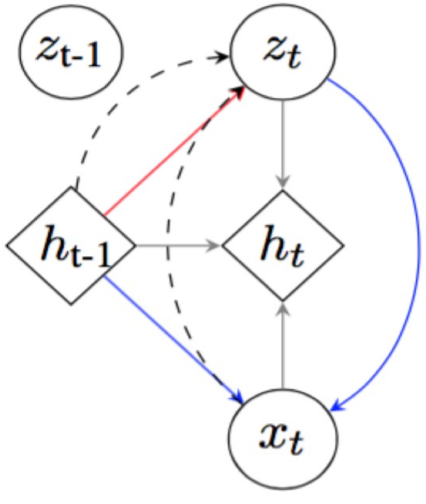


(a) Prior (b) Generation (c) Recurrence (d) Inference (e) Overall

Variable	**Corresponding Interpretation in dialogue structure**
Xt: sequantial data	usr: “I want a <Chinese> restaurant” sys: “what <area> do you prefer?”
Zt: latent vector	latent state interpreted as <looking for a restaurant> **(<Z0, Z1, …, Zn>, transition prob <Zi, Zi+1>)** represents the dialogue structure
ht: hiddent vector in RNN	hidden representations

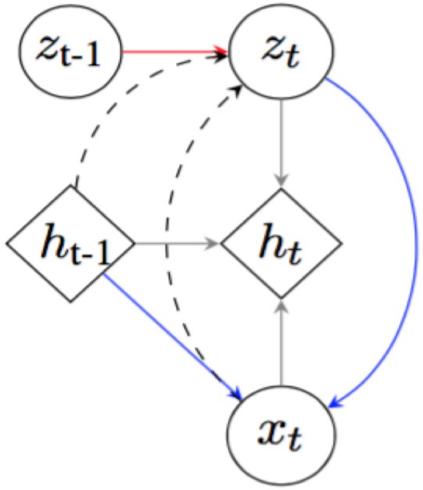
Discrete-Variational-RNN

Zt in VRNN: continuous → hard to interpret → discrete with Gumbel-Softmax



(a) Discrete-VRNN

Different in the prior!



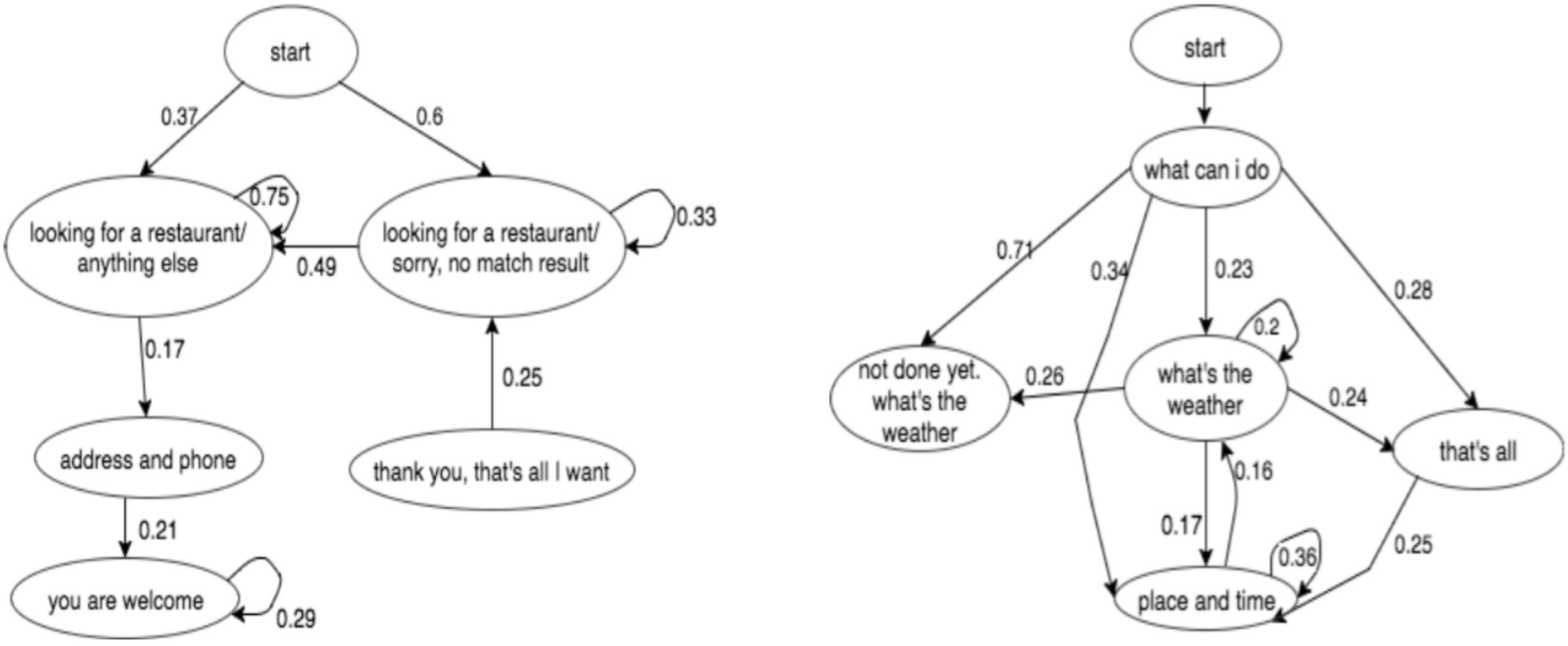
(b) Direct-Discrete-VRNN

+ penality on entity
→ NE-DD-VRNN

Transition prob <Zi, Zi+1>:
Frequency Count of predicted <Zi, Zi+1> tuple

Transition prob <Zi, Zi+1>:
These transition probabilities are model parameters in the prior, trained by minimizing the loss function.

Dialogue Structure by DD-VRNN

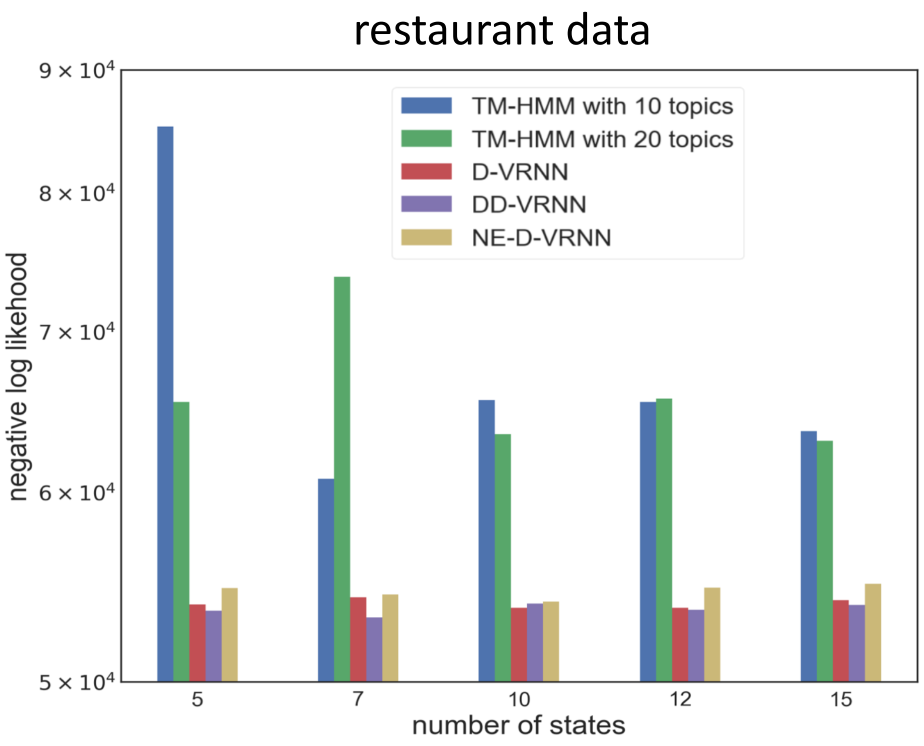


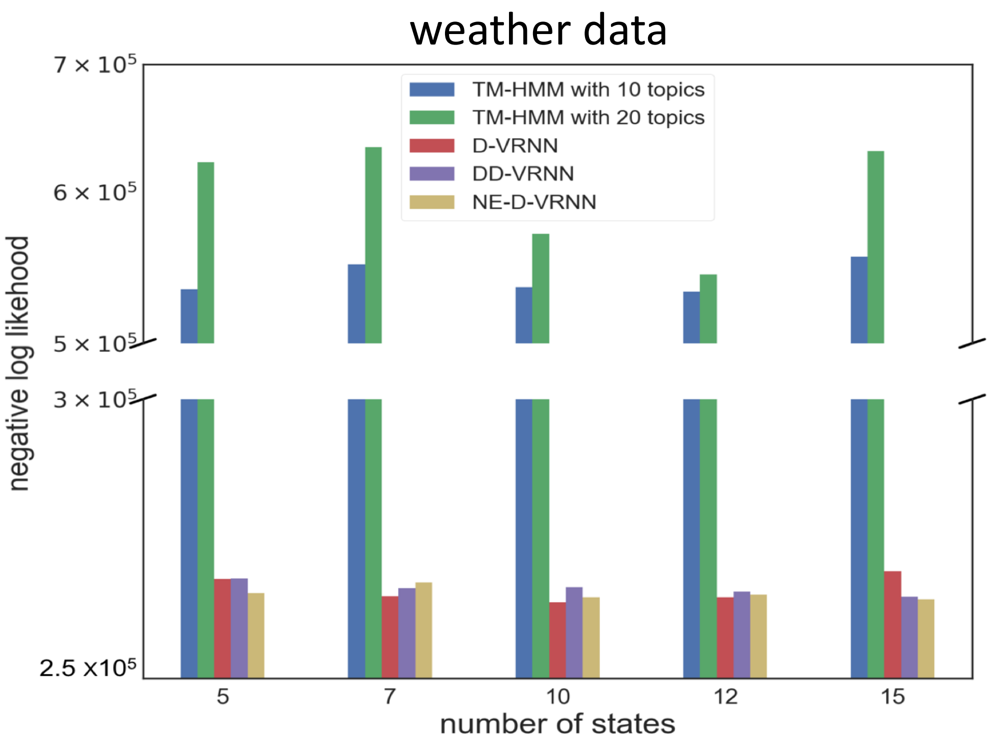
(a) DD-VRNN, restaurant data, 10 states

(b) DD-VRNN, weather data, 5 states

Figure 8: Dialog structures generated by DD-VRNN on different datasets.

Quantitative Evaluation





Comparison with HMM:
better than HMM on the negative log likelihood

Comparison with Model Variants:

1. When N is small, DD-VRNN performs the best
2. D-VRNN is the most stable one
3. The three models compensate each other

Comparison with K-means

From	Utterance
SYS:	Okay, you don't care place, do you?
USR:	That's correct.
SYS:	What date are you interested?
USR:	**Weather this morning.**
SYS:	**I believe you said this morning.**

(a) One example in State 2, "provide place and time"

From	Utterance
USR:	Weather tomorrow.
SYS:	[api_call]. your weather report is here. what else can I do?
USR:	**Weather this morning.**
SYS:	**I believe you said this morning.**

(b) One example in State 3, "provide time"

Because they have different context, they should be in differenet state!
But a simple K-means with just put them into the same cluster given their similar surface form

Application in RL

Ppred from state1 to state2 from the RL model

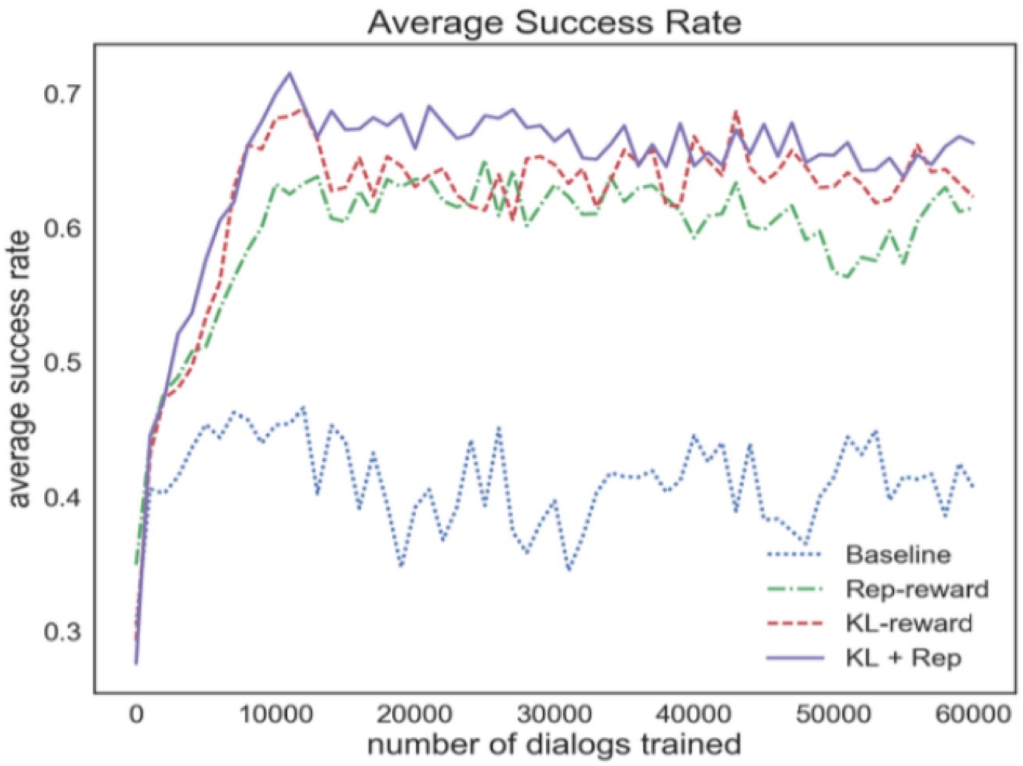
calculate KL divergence as the reward to force the RL model to be closer to the real distribution

Ptrans from state1 to state2 from D-VRNN, which summarizes the transition pattern in the real data

Algorithm 1 Reward function

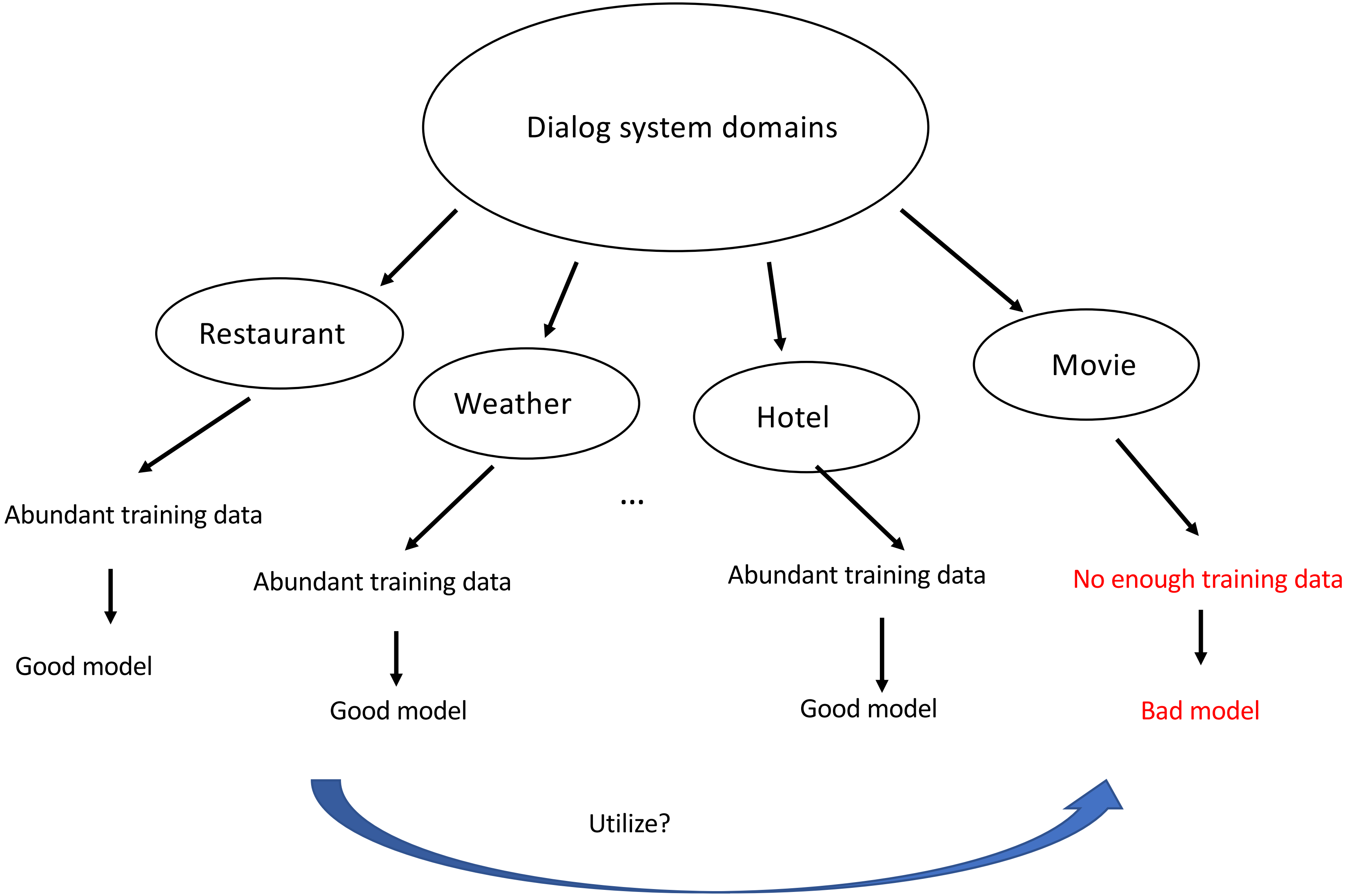
if success **then**
  $R = 20$
else if failure **then**
  $R = -10$
else if repeated question **then**
  $R = f_{\text{reward}}(p_{\text{trans}}, p_{\text{pred}})$
else if each proceeding turn **then**
  $R = -1$
end if

$$f_{\text{reward}}(p_{\text{trans}}, p_{\text{pred}}) = \begin{cases} -1 & \text{Baseline} \\ -5 & \text{Rep-reward} \\ -\mathrm{KL}(p_{\text{trans}}, p_{\text{pred}}) & \text{KL-reward} \\ -\mathrm{KL}(p_{\text{trans}}, p_{\text{pred}}) - 2 & \text{KL+Rep} \end{cases} \tag{10}$$



Meta Learning for Dialog Domain Adaptation

ACL 2019



- Model-Agnostic Meta-Learning(MAML)

$$\theta \xrightarrow{(x_j^1, y_j^1) in D_1} \theta_1' = \theta - \alpha \nabla_\theta \mathcal{L}_{D_1}(\theta) \xrightarrow{(x_j^1, y_j^1)} \mathcal{L}_{D_1}(\theta_1')$$

$$+$$

$$\theta \xrightarrow{(x_j^2, y_j^2) in D_2} \theta_2' = \theta - \alpha \nabla_\theta \mathcal{L}_{D_2}(\theta) \xrightarrow{(x_j^2, y_j^2)} \mathcal{L}_{D_2}(\theta_2')$$

$$+$$

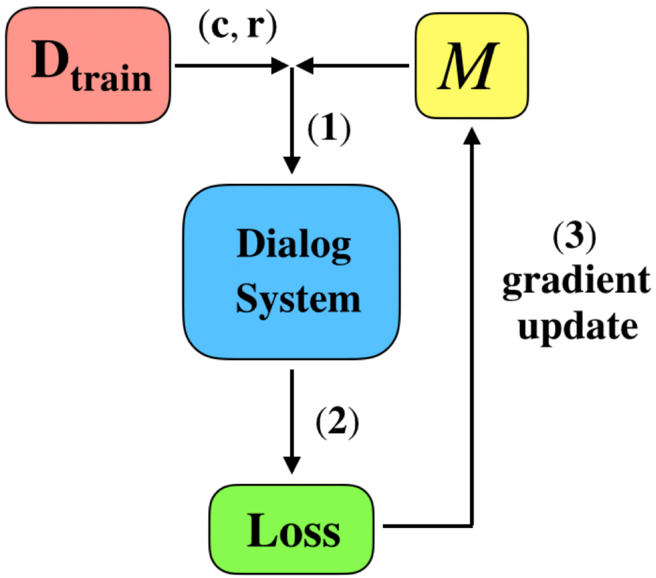
$$\theta \xrightarrow{(x_j^3, y_j^3) in D_3} \theta_3' = \theta - \alpha \nabla_\theta \mathcal{L}_{D_3}(\theta) \xrightarrow{(x_j^3, y_j^3)} \mathcal{L}_{D_3}(\theta_3')$$

$$\sum \mathcal{L}_{D_i}(\theta_i')$$

$$\hat{\theta} = \theta - \beta \nabla_\theta \left(\sum \mathcal{L}_{D_i}(\theta_i')\right)$$

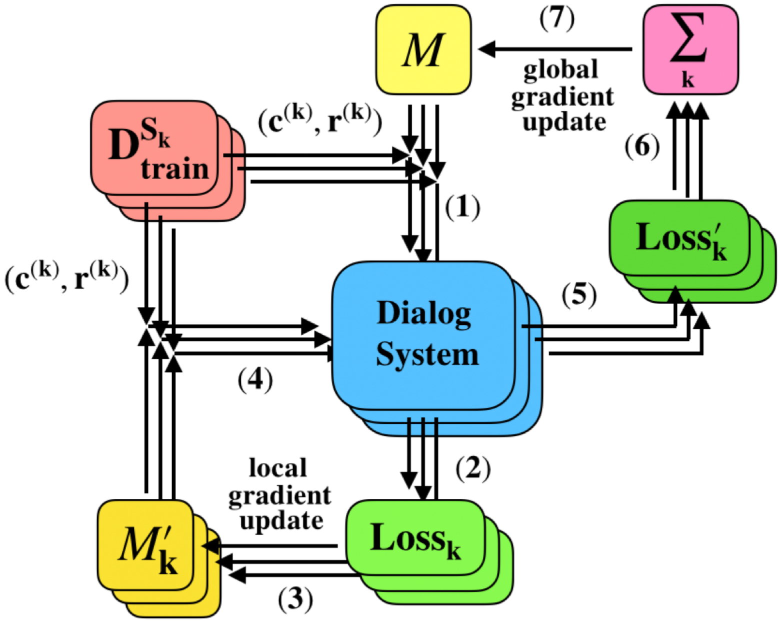
- Dialog System with Meta-learning

(a) Classic gradient update



$$\mathcal{M} \leftarrow \mathcal{M} - \alpha \nabla_{\mathcal{M}} \mathcal{L}(\mathcal{M}, c^{(k)})$$

(b) Meta-learning update



$$\mathcal{M} \leftarrow \mathcal{M} - \beta \nabla_{\mathcal{M}} \sum_{D_k} \mathcal{L}_{D_k}(\mathcal{M}_k', c^{(k)})$$

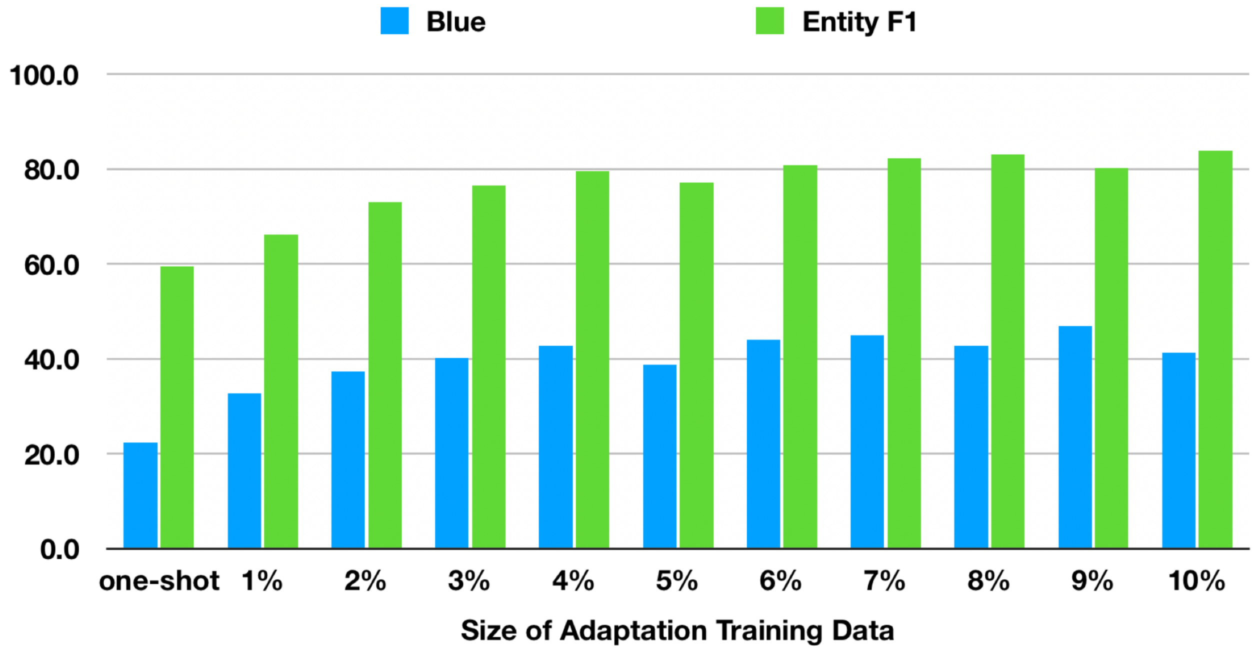
- Source Domains (900 training, 100 validation dialogs for each domain):
  Restaurant, Bus, Weather

- Target Domains (500 testing dialogs for each domain):
  Restaurant (in-domain)
  Restaurant-slot (unseen slot): introduce new slot values
  Restaurant-style (unseen NLG): same slot values but different NLG templates
  Movie (new-domain): completely new domains

- Metric:
  - BLEU score evaluate the quality of generated response sentence
  - **Entity F1 score** evaluate the completeness of tasks

In domain	ZSDG	Transfer	DAML	Transfer-oneshot	DAML-oneshot
BLEU	70.1	51.8	51.8	51.1	53.7
Entity F1	79.9	88.5	**91.4**	87.6	91.2
Epoch	-	2.7	1.4	2.2	1.0
Unseen Slot	ZSDG	Transfer	DAML	Transfer-oneshot	DAML-oneshot
BLEU	68.5	43.3 (46.3)	41.7 (46.3)	40.8 (43.9)	40.0 (41.8)
Entity F1	74.6	78.7 (78.5)	75 (**79.2**)	70.1 (67.7)	72.0 (73.0)
Epoch	-	2.6 (2.4)	4.8 (3.4)	3.2 (2.6)	5.0 (3.0)
Unseen NLG	ZSDG	Transfer	DAML	Transfer-oneshot	DAML-oneshot
BLEU	70.1	30.6 (32.4)	21.5 (26.0)	20.0 (21.5)	19.1 (19.1)
Entity F1	72.9	82.2 (85.0)	77.5 (82.4)	82.8 (86.2)	69.0 (**86.4**)
Epoch	-	3.2 (3.0)	3.2 (2.1)	12.3 (20.3)	4.7 (5.7)
New Domain	ZSDG	Transfer	DAML	Transfer-oneshot	DAML-oneshot
BLEU	54.6	30.1	32.7	21.5	22.4
Entity F1	52.6	64.0	**66.2**	55.9	59.5
Epoch	-	5.6	4.5	14.2	5.8

movie	Transfer	DAML
Entity F1	64.0	**66.2**
BLEU	30.1	**32.7**
restaurant	Transfer	DAML
Entity F1	80.7	**82.1**
BLEU	46.1	**47.9**
bus	Transfer	DAML
Entity F1	60.0	**61.9**
BLEU	32.0	**35.9**
weather	Transfer	DAML
Entity F1	79.1	**80.4**
BLEU	38.9	**43.3**



- DAML outperforms transfer learning in different target domains
- DAML almost converges with only 4% data from the target domain

Customer Service Application: Sentiment Adaptive

Shi & Yu ACL 2018

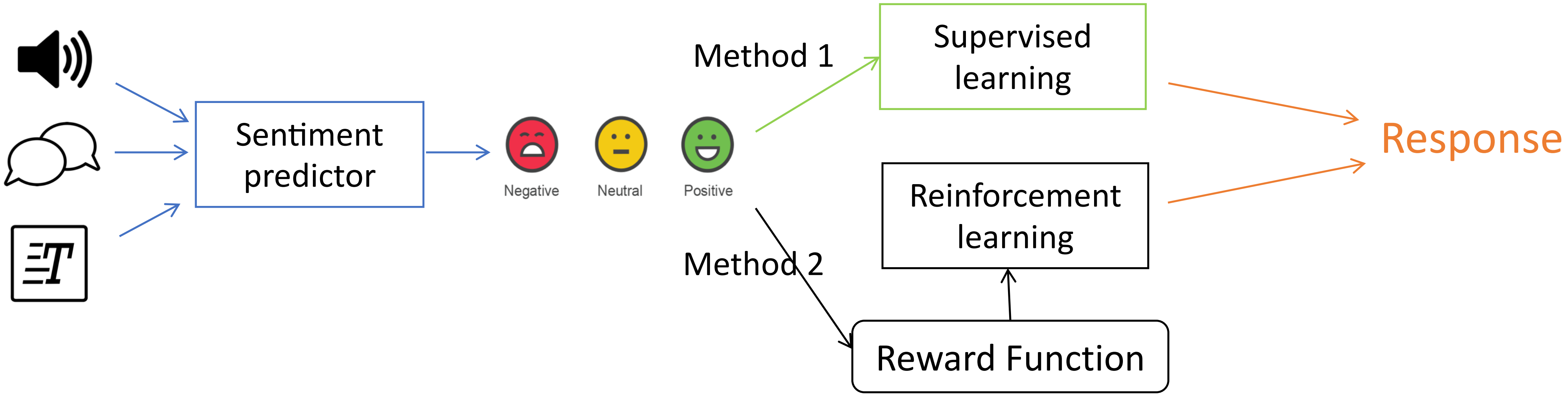
Goal: Incorporate user sentiment information into dialog system training for better task success and user experience.

Method: Supervised learning and reinforcement learning.

Sentiment predictor: 1) acoustic 2) dialogic and 3) textual

Supervised learning: sentiment as a direct feature

Reinforcement learning: sentiment as immediate reward



Personalized Persuasion Dialog

ACL 2019

Donate to Save the Children Charity, based on personality, decision making style and value systems use different persuasion strategies

Persuader: Would you like to donate to a children’s charity called Save the Children?
Persuadee: What is this charity about?
Persuader: a non-governmental organization that promotes children's right, supports kids in developing countries.
Persuadee: Developing countries is a broad category. Is there any countries that they focus?
Persuader: Syria, Thousands of kids die in just a quarter of a year over there and you could help save them with just some pocket change
Persuadee: Where is the organisation's headquarters?
Persuader: London, United Kingdom. If you care about children, even a little bit, then you should be outraged and willing to help make an impact
Persuadee: I get mails and calls from different charities. I just want to make sure that this one is genuine

Anti-Scam Dialog Systems

Waste attackers time and also elicit their personal information

Role	Sentences	Dialog Act
Attacker	Hi.	*other_statement*
Attacker	I'm INCMP_NAME with Amazon's Distribution Center.	*providing_info*
User	Hello NAME I'm REAL_NAME how are you today?	*other_question*
Attacker	I'm doing very well, thank you for asking.	*thanking*
Attacker	How did you enjoy your recent Amazon purchase?	*other_question*
User	Well I'm very excited to use it, it hasnt seemed to arrive just yet	*explanation*
Attacker	May I please have you verify a few pieces of account information to better assist you?	*other_question*
User	Yes but first can you tell me where the package was shipped to.	*other_question*
Attacker	What is the name and billing address on the account for this order?	*elicitation_question*
User	REAL_NAME.	*providing_info*
User	I forgot the address, I don't know which vacation home its for	*refusal*
Attacker	I just need the billing address for now.	*elicitation_command*
Attacker	Then I can verify the address of the shipment.	*explanation*
User	I can't remember.	*refusal*
User	Can I call you back with it?	*elicitation_question*
Attacker	I can look this information up with your payment info if that's easier.	*other_statement*
User	No, but I'll call you after I ask my wife.	*refusal*
User	What's the best number to reach you?	*elicitation_question*
Attacker	Sure.	*hold*
Attacker	It is REAL_PHONE.	*providing_info*
Attacker	Thank you Mr. NAME.	*thanking*

Targeted Movie Promotion

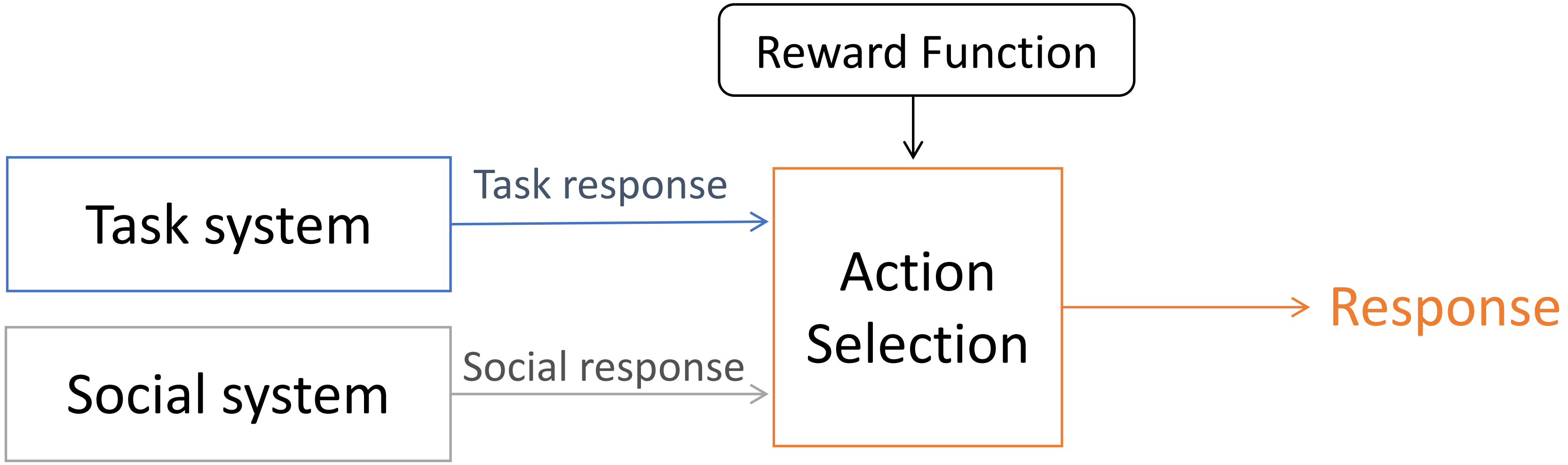
Yu et al., IJCAI, 2017

Goal: Elicit users' opinion about movies and then promote a movie based on their preferences.

Method: Interleave social content with task contents.

Task contents: A set of slots, such as: preferred movie type.

Reinforcement learning policy: Transition between social conversation and task conversation.

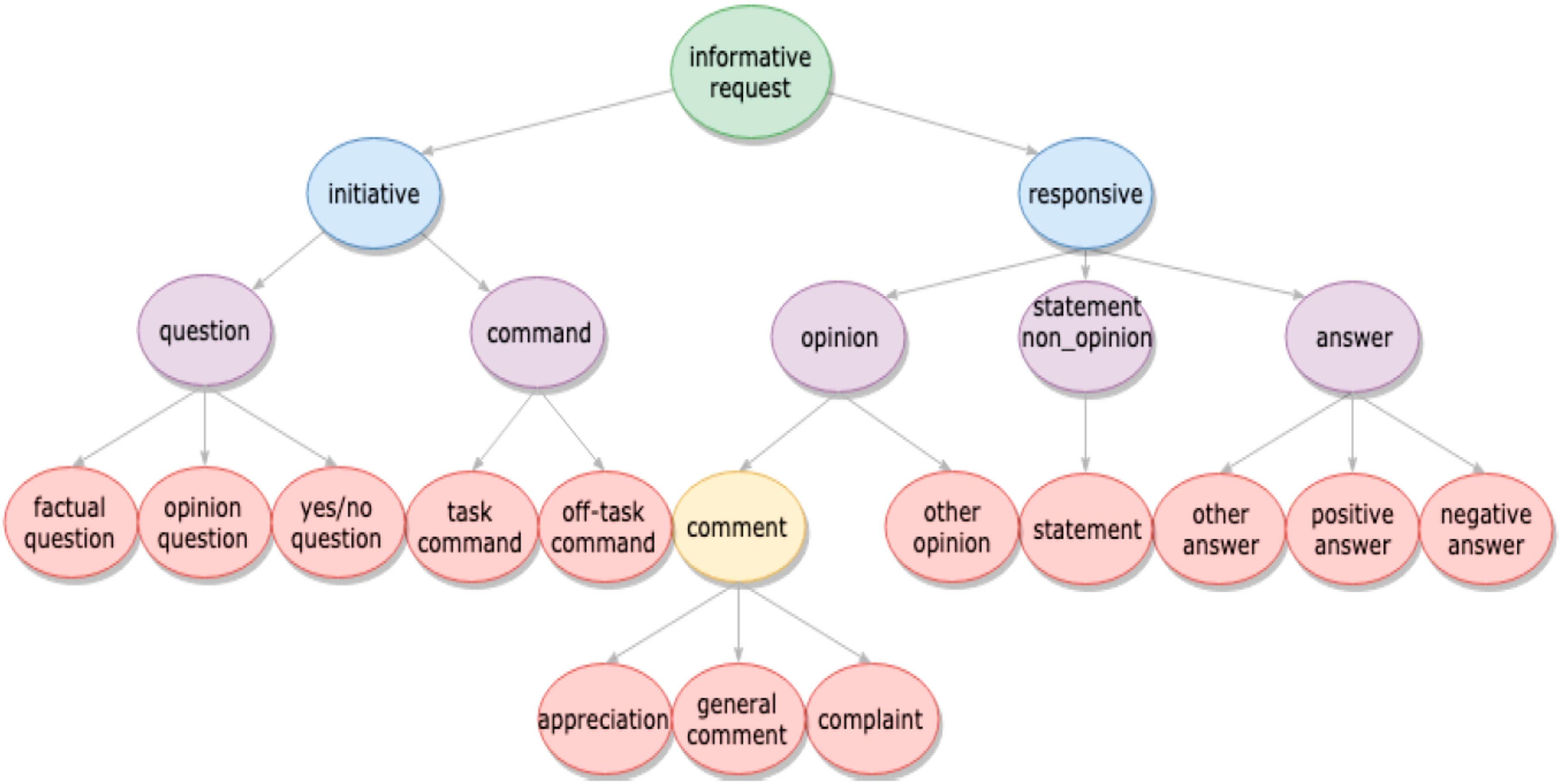


Visual Task-Oriented Dialogs

Zhang, Zhao & Yu
SIGDIAL 2018

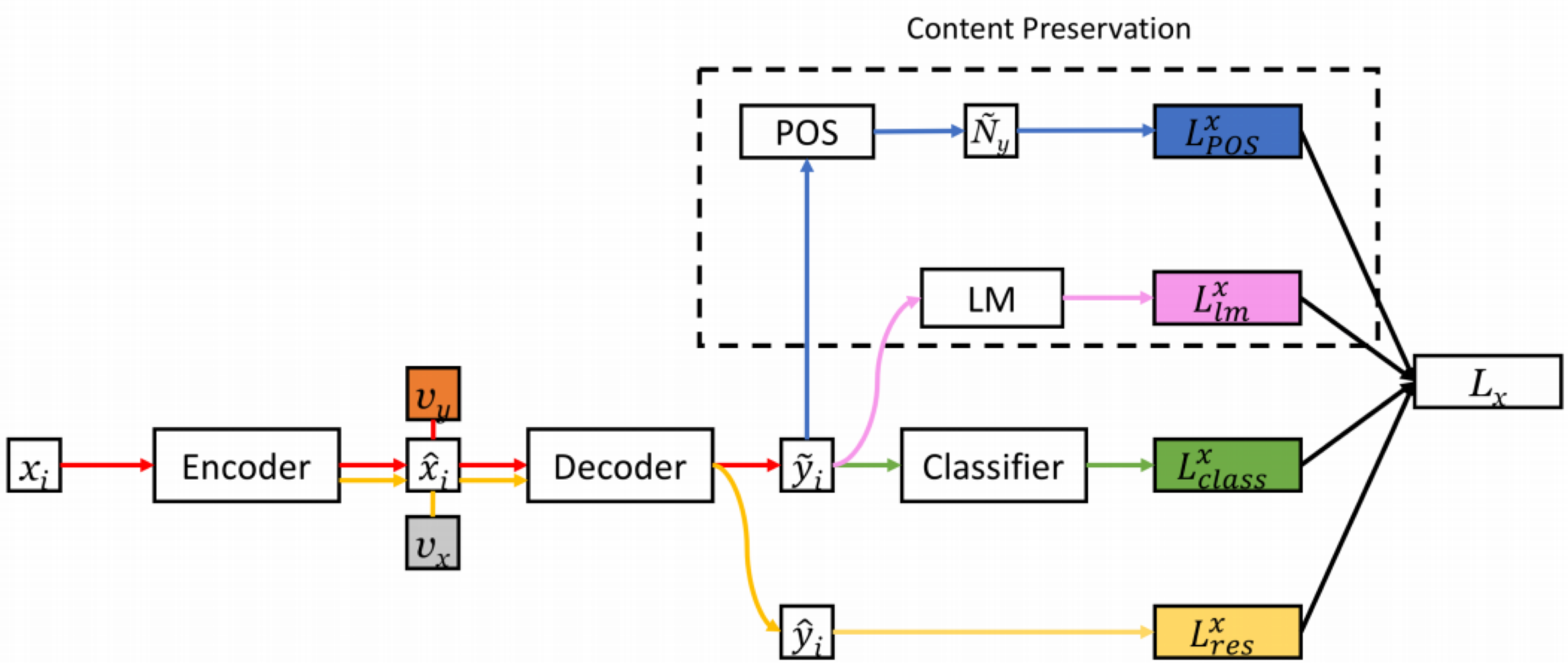
Das et al., 2017a

Open Domain Dialog Act



Text Style Transfer

Submitted to ACL



Story Completion with Common Sense Knowledge

AAAI 2019