GPT-2 Based End to End Task-Oriented Dialogue System

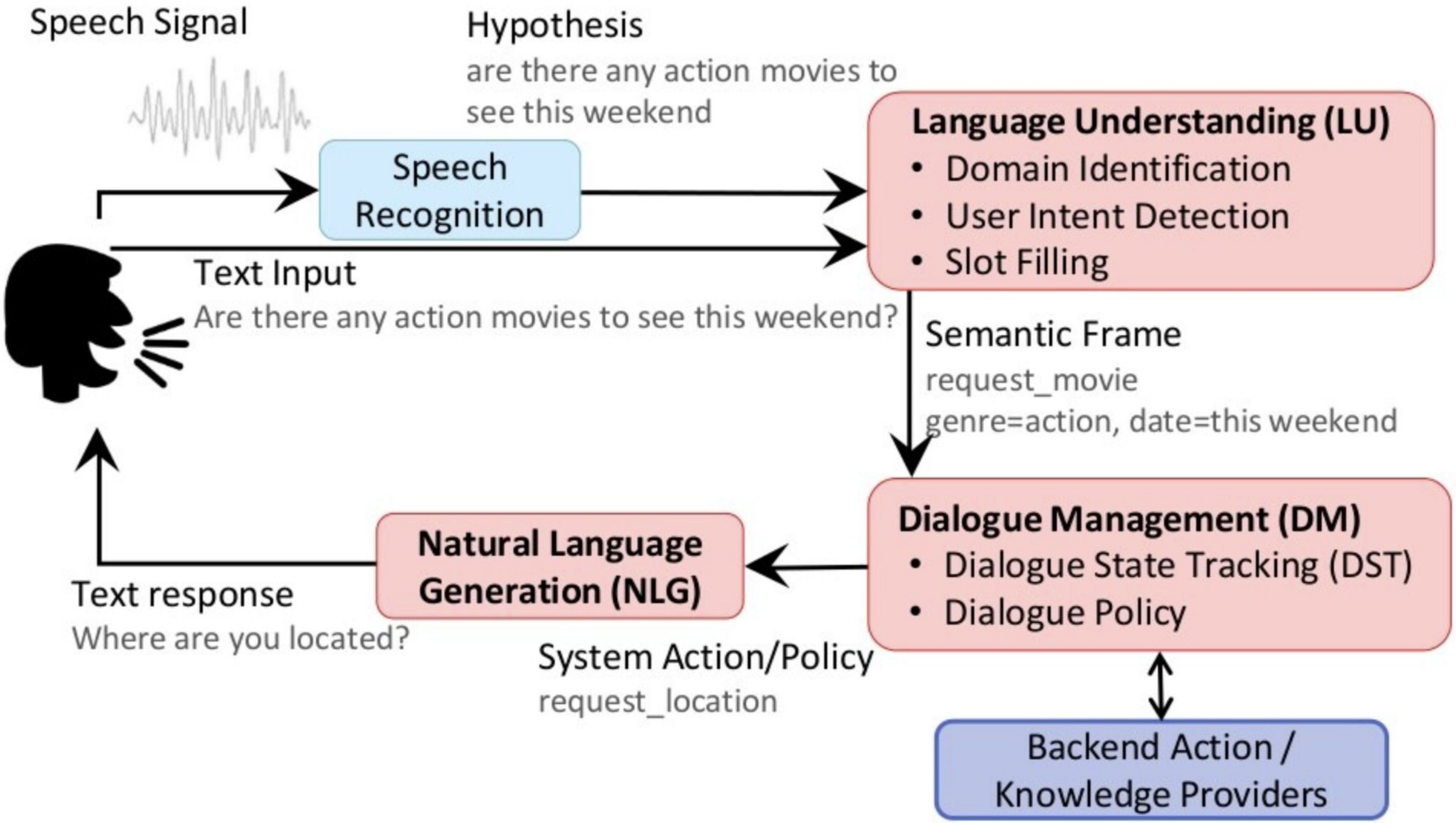
2021.10.03

집현전 2기 5조: 모윤호, 김대규, 원혜진

발표자 : 모윤호

https://arxiv.org/pdf/2012.03539.pdf

Task-Oriented Dialogue System



Overview

- **SimpleTOD**
  - GPT-2 based The first **end-to-end TOD model**
- **SOLOIST**
  - **SimpleTOD** + loss + negative sampling + pre-training
- **UBAR**
  - **SimpleTOD** + session level training + domain sharing

Dataset

Multi-domain Wizard-of-Oz (MultiWOZ) 2.0, 2.1

- 대규모의 다중 도메인의 대화 데이터셋
- 7개의 도메인 (레스토랑, 기차, 명소, 호텔, 택시, 병원, 경찰)
- 경찰, 병원도메인은 validation, test set ❌ -> 학습 제외
- 다중 도메인의 대화와 Dialogue State(belief State), DB데이터로 구성

2.1 = 2.0 + 데이터양 10배 + noise labeling

Multi–WOZ dataset Example

👨‍💻 : Help me find a moderate priced british food place please.

- {resturant–info = food, british}, {price = moderate}

🤖 : Restaurant one seven is a nice place.
Do you want to book?

👨‍💻 : Not at this time. Could I just get the phone number?

- {Restaurant–Request = phone, ?}

🤖 : Their phone number is 01223337766.
Do you need assistance with anything else?

Multi–WOZ 2.1 PMUL3125.json

Evaluation

아래 논문들에서는 automatic evaluation metrics를 따른다.

- Inform - 시스템이 올바른 entity를 제공했는지 여부를 측정
- Success - 요청된 모든 정보에 응답을 했는지 여부를 측정
- BLEU score - 생성된 응답이 얼마나 잘 생성됐는지 여부를 측정
- Combined score - (Inform+Success)x0.5+BLEU
- joint goal accuracy (JGA) - DST 성능평가

A Simple Language Model for Task-Oriented Dialogue (SimpleTOD)

Ehsan Hosseini-Asl, Bryan McCann, Chien-Sheng Wu, Semih Yavuz, Richard Socher

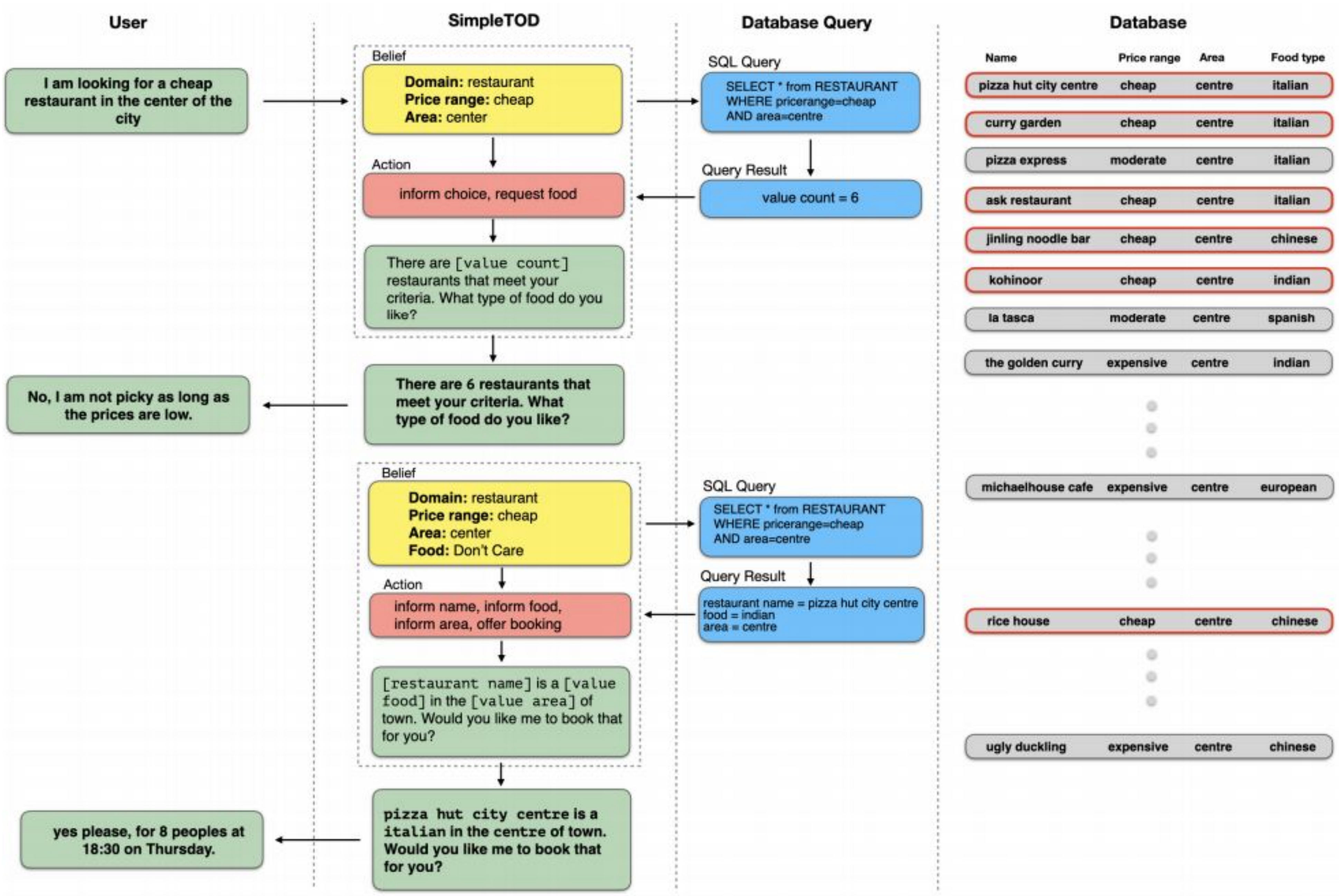
Salesforce Research

https://arxiv.org/pdf/2005.00796.pdf

Introduction

- dialogue state tracking, action decisions, response generation 을 함께 수행하는 최초의 End-to-End 모델
- Special Token의 중요성 확인
- GPT-2 based, used MultiWOZ data for fine-tuning

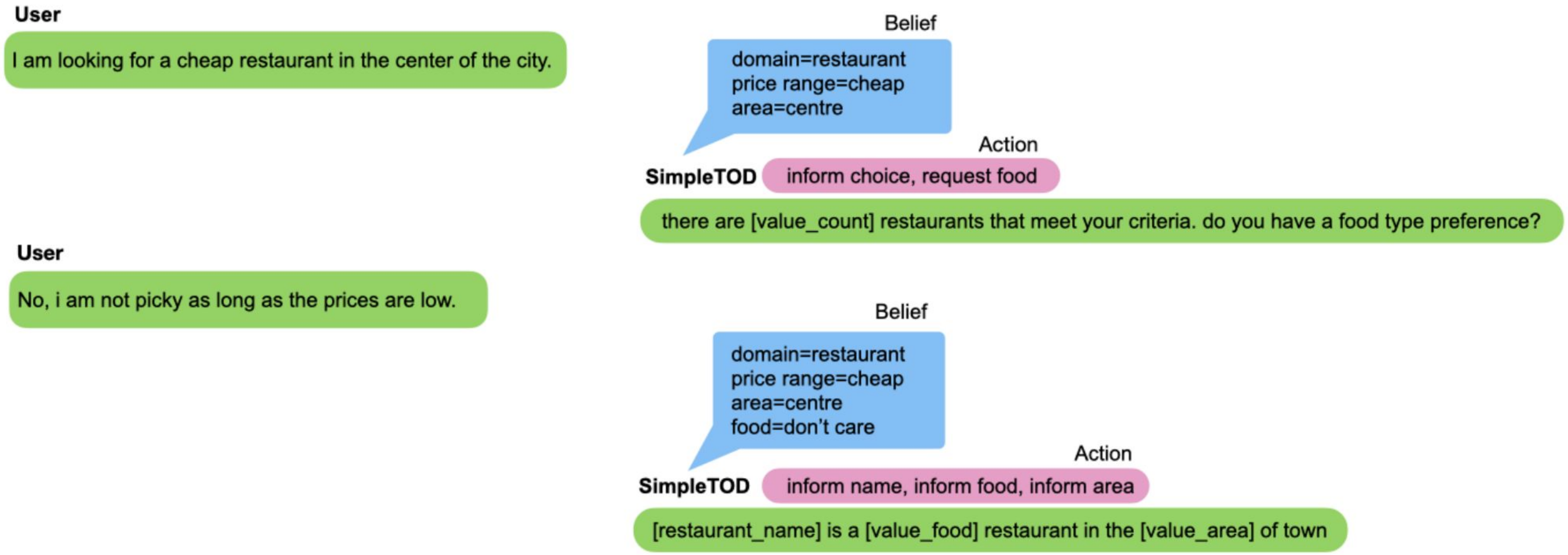
A Simple Language Model for Task-Oriented Dialogue (SimpleTOD)



Delexicalization

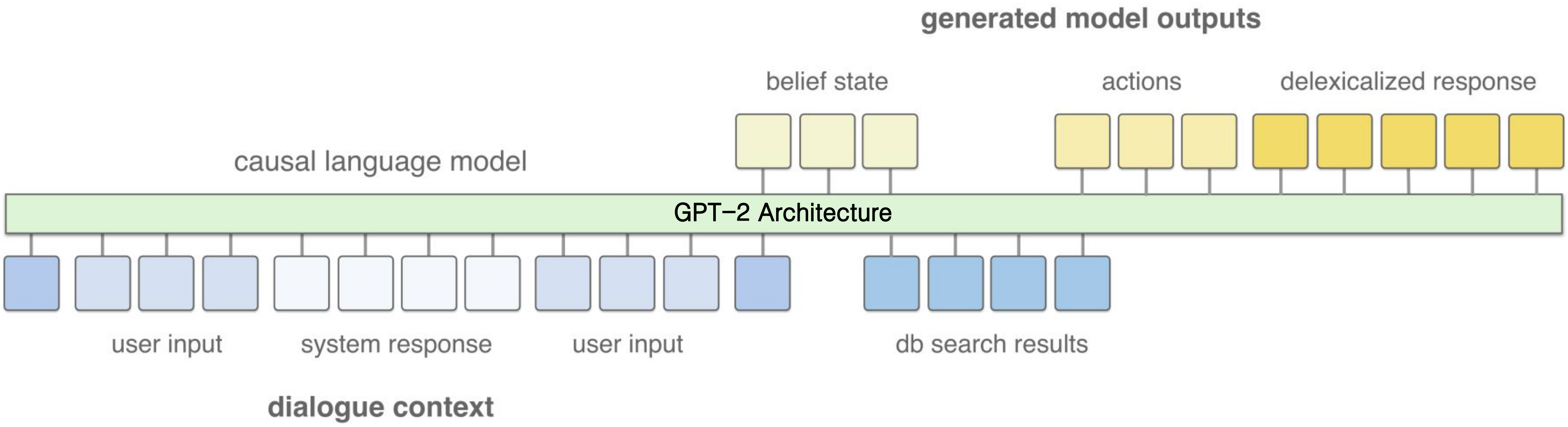
: 특정 slot값을 카테고리에 넣는 방법

: response generation할 때 카테고리가 출력됨



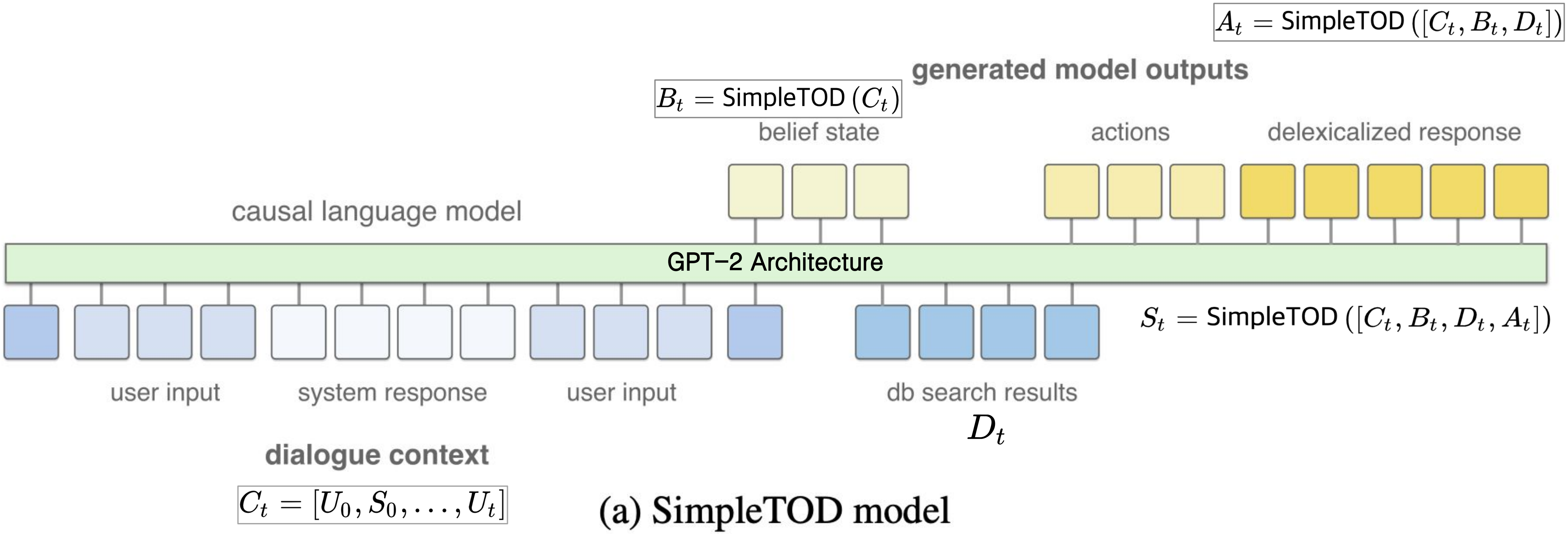
https://arxiv.org/pdf/1508.01745.pdf

Method



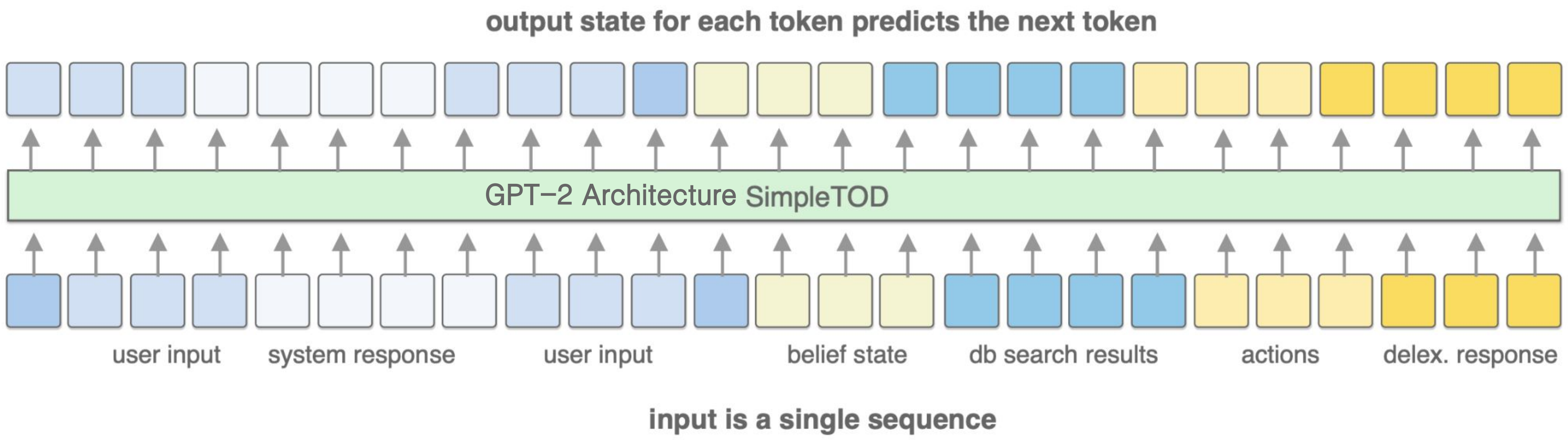
(a) SimpleTOD model

Method



(a) SimpleTOD model

Method



$$x^t = [C_t; B_t; D_t; A_t; S_t]$$

$$x = (x_1, \ldots, x_n)$$

$$x^t = n_t$$

$$D = \{x^1, \ldots, x^{|D|}\}$$

$$\mathcal{L}(D) = -\sum_{t=1}^{|D|} \sum_{i=1}^{n_t} \log p_\theta(x_i^t | x_{<i}^t)$$

Experiments

End-to-End 성능평가

Model	Belief State	DB Search	Action	Inform	Success	BLEU	Combined
DAMD+augmentation	generated	oracle	generated	76.3	60.4	16.6	85
SimpleTOD (ours)	generated	oracle	generated	78.1	63.4	16.91	87.66
SimpleTOD (ours)	generated	dynamic	generated	81.4	69.7	16.11	91.66
SimpleTOD (ours)	generated	-	generated	**84.4**	**70.1**	15.01	**92.26**

Table 2: Action and response generation on MultiWOZ 2.0 reveals that SimpleTOD, a single, causal language model, is sufficient to surpass prior work.

Belief State	DB Search	Action	Inform	Success	BLEU	Combined
generated	oracle	generated	79.3	65.4	16.01	87.36
generated	dynamic	generated	83.4	67.1	14.99	90.24
generated	-	generated	85	70.5	15.23	92.98

Table 3: Action and response generation on MultiWOZ 2.1 for SimpleTOD.

Experiments

DST성능평가

Model	Decoder	Context Encoder	Extra Supervision	Joint Accuracy
TRADE*	Generative + Classifier	Bidirectional	-	45.6
DSTQA**	Classifier	Bidirectional	knowledge graph	51.17
DST-Picklist*	Classifier	Bidirectional	-	53.3
SST*	Generative	Bidirectional	schema graph	55.23
TripPy†	Classifier	Bidirectional	action decision	55.3
SimpleTOD^o	Generative	Unidirectional	-	55.72
SimpleTOD*	Generative	Unidirectional	-	**55.76**
SimpleTOD+	Generative	Unidirectional	-	57.47

O – MultiWOZ 2.1 no label-cleaning
*** – MultiWOZ 2.1 test label cleaning**
+ – MultiWOZ 2.1 partial label cleaning

Experiments

Context	[context] [user] *user input* [system] *system response* . . . [user] *user input* [endofcontext]
Belief State	[belief] *domain slot_name value, domain slot_name value,* . . . [endofbelief]
DB Search	[db] *#_matches, booking_status* [endofdb]
Action	[action] *domain action_type slot_name, domain action_type slot_name,* . . . [endofaction]
Response	[response] *system delexicalized response* [endofresponse]

End token	User/System token	Joint Acc	Inform	Success	BLEU	Combined
No	No	16.79	33.8	10.6	4.53	26.73
Yes	No	21.5	54.5	41.2	9.48	57.33
No	Yes	22.22	61.9	52.7	9.57	66.87
Yes	Yes	55.76	85	70.5	15.23	92.98

Conclusion

- SimpleTOD는 dialogue state tracking, action decisions, response generation을 함께 수행하는 최초의 End to End방식 모델

- 당시 DST분야에서는 SOTA를 찍은 모델

- 다른 sub-task를 분류하는 Special Token을 사용하니 성능 향상

SOLOIST: Building Task Bots at Scale with Transfer Learning and Machine Teaching

Baolin Peng, Chunyuan Li, Jinchao Li Shahin Shayandeh,
Lars Liden, Jianfeng Gao

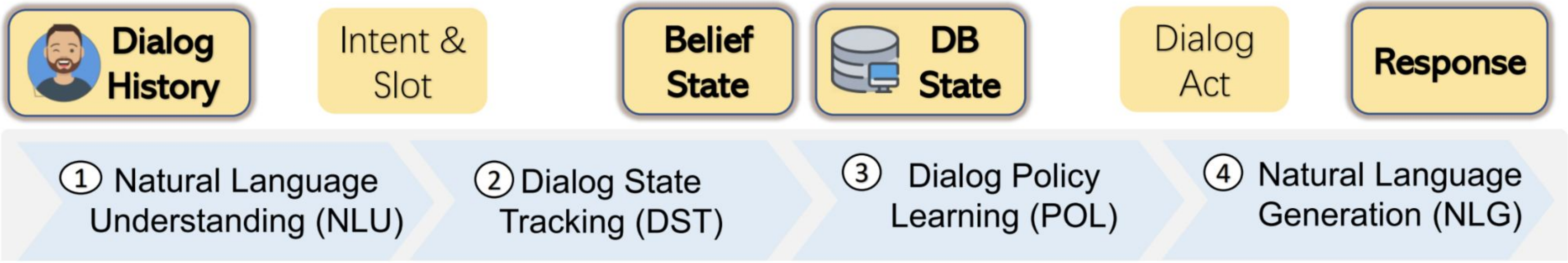
Microsoft Research, Redmond

https://arxiv.org/pdf/2005.05298.pdf

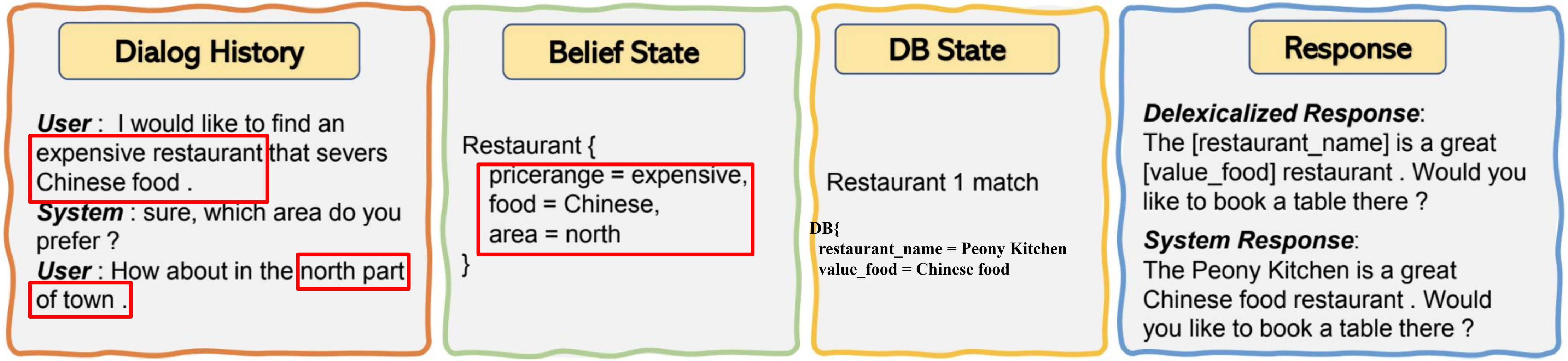
Introduction

- Transfer learning을 사용한 Task-oriented dialog system
- Transformer 기반의 Auto-regressive language model -> GPT-2 사용
- 데이터가 적은 Task-oriented dialog system 안에서 해결 방법을 제시

- Negative Sampling 을 통한 Data Augmentation
- Task-Grounded pre-training

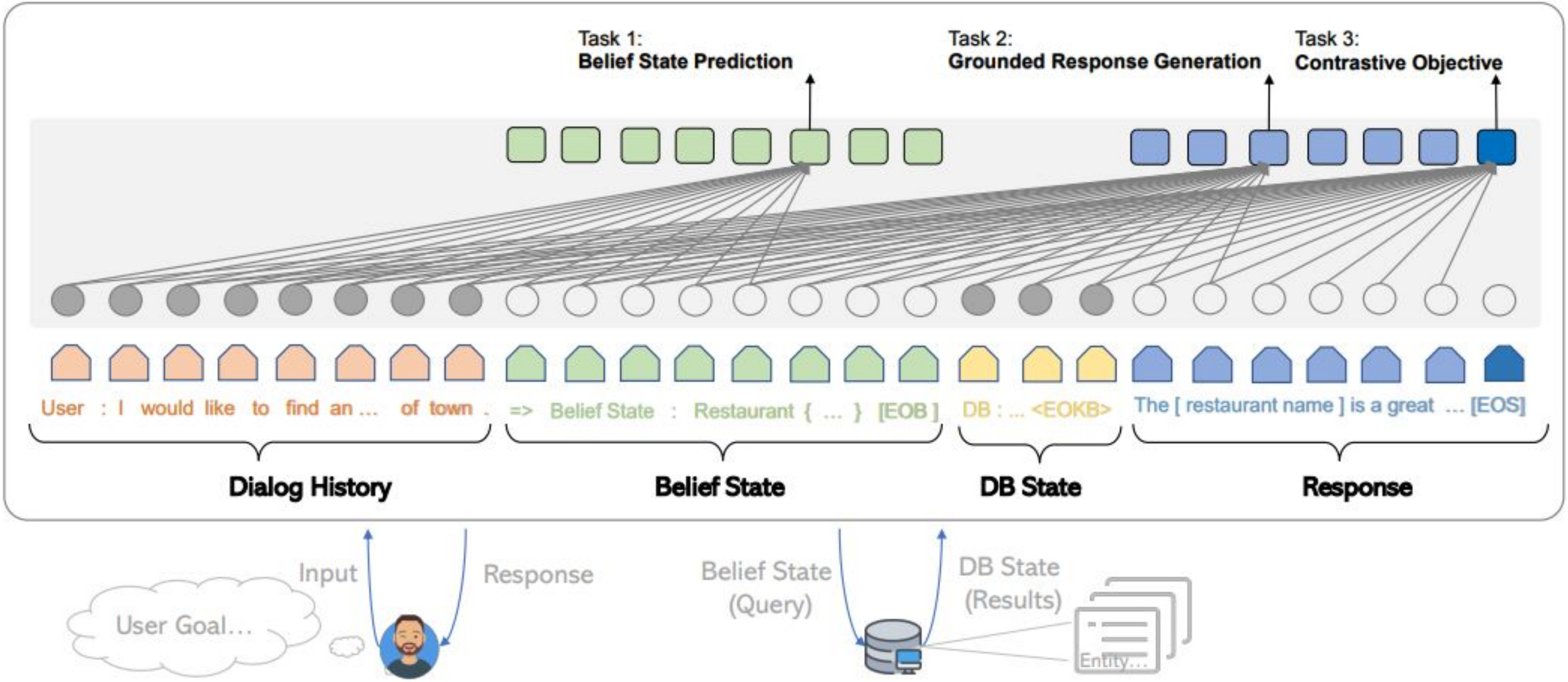


(a) A typical task-oriented dialog system pipeline.



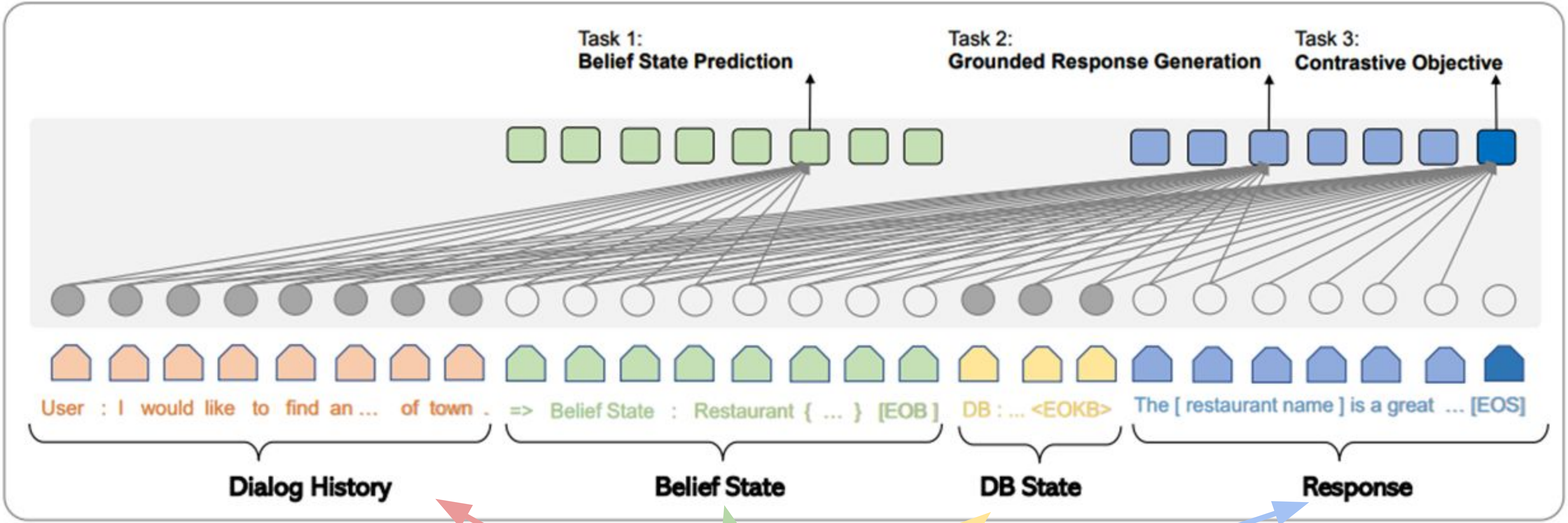
(b) Example snippets for the items compounding the input of SOLOIST model.

Architecture



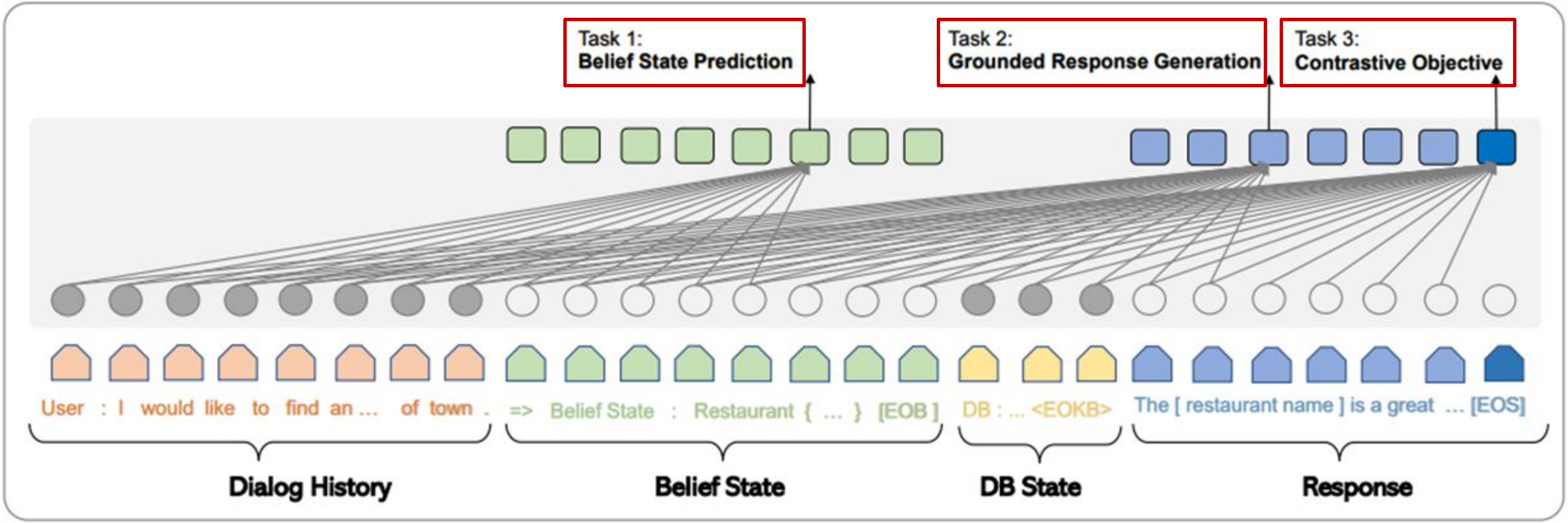
(c) The proposed SOLOIST model architecture and training objectives.

Architecture



$$x = (s, b, c, r)$$

Architecture



Method

Dialog History | Belief State | DB State | Response

$$x = (s, b, c, r)$$

$$p(x) = p(r, c, b, s)$$

$$p(x) = p(r, c, b, s) \cdot p(c|b, s) \cdot p(b|s) \cdot p(s)$$

$$p(x) = p(r, c, b, s) \cdot \cancel{p(c|b, s)} \cdot p(b|s) \cdot p(s)$$

$$= \underbrace{p(r, c, b, s)}_{\text{Grounded Response Generation}} \cdot \underbrace{p(b|s)}_{\text{Belief Prediction}} \cdot p(s)$$

Method

Task 1: Belief Prediction

$$\mathcal{L}_{\mathrm{B}} = \log p(\boldsymbol{b}|\boldsymbol{s}) = \sum_{t=1}^{T_b} \log p_{\boldsymbol{\theta}}(b_t|b_{<t}, \boldsymbol{s})$$

Task 2: Grounded Response Generation

$$\mathcal{L}_{\mathrm{R}} = \log p(\boldsymbol{r}|\boldsymbol{c}, \boldsymbol{b}, \boldsymbol{s}) = \sum_{t=1}^{T_r} \log p_{\boldsymbol{\theta}}(r_t|r_{<t}, \boldsymbol{c}, \boldsymbol{b}, \boldsymbol{s})$$

Task 3 : Contrastive Objective

$$\mathcal{L}_{\mathrm{C}} = y \log(p_{\boldsymbol{\theta}}(\boldsymbol{x})) + (1-y) \log(1 - p_{\boldsymbol{\theta}}(\boldsymbol{x}'))$$

Full Pre-Training Objective

$$\mathcal{L}_{\boldsymbol{\theta}}(\mathcal{D}) = \sum_{n=1}^{|\mathcal{D}|} (\mathcal{L}_{\mathrm{B}}(\boldsymbol{x}_n) + \mathcal{L}_{\mathrm{R}}(\boldsymbol{x}_n) + \mathcal{L}_{\mathrm{C}}(\boldsymbol{x}_n))$$

Task 3: Contrastive Objective

$$\mathcal{L}_C = y \log(p_{\boldsymbol{\theta}}(\boldsymbol{x})) + (1-y)\log(1 - p_{\boldsymbol{\theta}}(\boldsymbol{x}'))$$

- Negative Sampling 을 위한 Loss Function
- Y = 1(POS), 0 (NEG) Cross Entropy사용
    - POS : NEG = 1 : 4 로 생성

- Negative Sampling은 1/3의 확률로 생성

  - Negative Belief State

  - Negative Response

  - Negative Belief State + Response

slot { domain = restaurant
cost = cheap}   이게 정답이면 label 1

slot { domain = restaurant
cost = expensive}   label 0

Dataset

Name	#Dialog	#Utterance	Avg. Turn	#Domain
task-grounded pre-training:				
Schema	22,825	463,284	20.3	17
Taskmaster	13,215	303,066	22.9	6
fine-tuning:				
MultiWOZ2.0	10,420	71,410	6.9	7
CamRest676	676	2,744	4.1	1
Banking77	-	25,716	-	21
Restaurant-8k	-	8,198	-	1

Table 1: Dialog corpora. The datasets in the upper block are used for task-grounded pre-training, and the datasets in the lower block are for fine-tuning.

Experiments

Model	Annotations		Evaluation Metrics			
	Belif State	Policy	Inform ↑	Success ↑	BLEU ↑	Combined ↑
Sequicity (Lei et al., 2018)	✓	✓	66.41	45.32	15.54	71.41
HRED-TS (Peng et al., 2019)	✓	✓	70.00	58.00	17.50	81.50
Structured Fusion (Mehri et al., 2019)	✓	✓	73.80	58.60	16.90	83.10
DSTC8 Track 1 Winner [1] (Ham et al., 2020)	✓	✓	73.00	62.40	16.00	83.50
DAMD (Zhang et al., 2019a)	✓	✓	76.40	60.40	16.60	85.00
SOLOIST	✓		**85.50**	**72.90**	**16.54**	**95.74**

[1]The result of DSTC8 Track 1 Winner is produced by adapting their code to our current setting.

Table 4: End-to-end evaluation on MultiWOZ.

Experiments

Model	Attraction			Train			Hotel			Restaurant		
	Inform ↑	Success ↑	BLEU ↑	Inform ↑	Success ↑	BLEU ↑	Inform ↑	Success ↑	BLEU ↑	Inform ↑	Success ↑	BLEU ↑
DAMD (Zhang et al., 2020b)	70.00	15.00	6.90	75.00	39.50	6.20	62.50	20.50	7.60	68.00	19.50	10.50
SOLOIST w/o pre-training	65.66	46.97	5.85	59.00	44.00	7.07	62.50	40.00	7.70	75.50	44.50	11.00
SOLOIST	**86.00**	65.00	12.90	80.81	64.65	9.96	74.50	43.50	8.12	81.00	55.50	12.80
SOLOIST_L	**86.00**	**68.00**	**14.60**	**81.31**	**74.24**	**11.90**	**75.00**	**51.50**	**10.09**	**84.00**	**62.50**	**13.17**

Table 6: End-to-end evaluation on MultiWOZ in the few-shot fine-tuning setting.

Model	1%			5%			10%			20%		
	Inform ↑	Success ↑	BLEU ↑	Inform ↑	Success ↑	BLEU ↑	Inform ↑	Success ↑	BLEU ↑	Inform ↑	Success ↑	BLEU ↑
DAMD (Zhang et al., 2020b)	34.40	9.10	8.10	52.50	31.80	11.60	55.30	30.30	13.00	62.60	44.10	14.90
SOLOIST w/o pre-training	46.10	24.40	10.39	63.40	38.70	11.19	64.90	44.50	13.57	70.10	52.20	14.72
SOLOIST	**58.40**	**35.30**	**10.58**	**69.30**	**52.30**	**11.80**	**69.90**	**51.90**	**14.60**	**74.00**	**60.10**	**15.24**

Table 7: End-to-end evaluation on MultiWOZ with varying sizes of task-specific training data for fine-tuning.

Experiments

Model	Inform ↑	Success ↑	BLEU ↑	Combined ↑
Full objective	85.50	72.90	16.54	95.74
- w/o belief	81.50	69.30	16.82	92.22
- w/o belief & response	82.50	67.30	16.28	91.18

Table 10: Ablation study on different negative samples in the contrastive objective on MultiWOZ in the end-to-end evaluation setup; The 2nd and 3rd row indicate removing individual belief only and individual belief & response, respectively.

Negative sampling이 효과적이다.
POS : NEG = 1 : 4 로 샘플링

Model	Success ↑	Under. ↑	Appr. ↑	Turns ↓
SOLOIST	**91.67**	**4.29**	**4.43**	18.97
DSTC8 Track 1 Winner	68.32	4.15	4.29	19.51
DSTC8 2nd Place	65.81	3.54	3.63	15.48
DSTC8 3rd Place	65.09	3.54	3.84	**13.88**
DSTC8 Baseline	56.45	3.10	3.56	17.54

Table 13: Human evaluation results. The results except SOLOIST are quoted from Li et al. (2020b).

Conclusion

- GPT-2모델을 바로 Task-Oriented Dialogue data에 fine-tuning하지 않음
- 비슷한 task의 large size corpus에 pre-training -> fine-tuning
- 성능 향상에 효과적
- Negative sampling 으로 Task-oriented Dialogue의 데이터 늘림

UBAR: Towards Fully End-to-End Task-Oriented Dialog System with GPT-2

Yunyi Yang,Yunhao Li, Xiaojun Quan

https://arxiv.org/pdf/2012.03539.pdf

Introduction

이전의 연구의 3가지 문제점 지적

1. 이전 연구들은 dialog session level 대신 **dialog turn level**에서 train, evaluation이 진행
   a. 이전 턴의 belief state 및 system act 와 같은 정보는 제외됨
   b. 제외된 정보는 현재 턴 생성에 유용한 정보가 될 수 있음

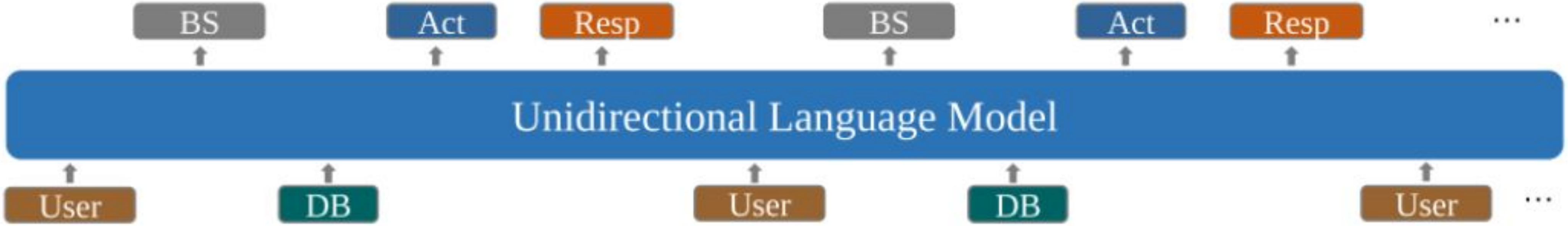
2. 다른 turn의 label 사용하여 현재 turn의 dialog generation할 수 있다.

3. 실제 대화에서 정답 시스템 응답에 액세스할 수 있다는 가정은 효과적이지 않음

-> dialog session level 에서 Task-oriented dialog 모델링

dialog session level vs dialog turn level

dialog turn level (SimpleTOD, SOLOIST)



dialog session level(UBAR)

Domain-Adaptive Pre-processing : delexicalizing

... It is also [value_price], and rated [value_stars] stars. <eos_r> <sos_u> great. can you book that 1 for 4 nights checking in on tuesday? there will be 6 people. <eos_u> <sos_b> [hotel] internet yes type guest house parking yes stars 4 name alexander b&d and breakfast stay 4 day tuesday people 6 <eos_b> <sos_db> <db_1> <eos_db> <sos_a> [hotel] [offerbooked] reference [general] [reqmore] <eos_a> <sos_r> ok , the booking was successful . your reference number is [value_reference] . is there anything else i can help you with today? <eos_r> <sos_u> no that will be all . Thanks <eos_u> <sos_b> …

Response delexicalizing : Belief state, DB result, system Act의 정보를 span으로 변환

Domain-Adaptive Pre-processing : special token

... It is also [value_price], and rated [value_stars] stars. <eos_r> <sos_u> great. can you book that 1 for 4 nights checking in on tuesday? there will be 6 people. <eos_u> <sos_b> [hotel] internet yes type guest house parking yes stars 4 name alexander b&d and breakfast stay 4 day tuesday people 6 <eos_b> <sos_db> <db_1> <eos_db> <sos_a> [hotel] [offerbooked] reference [general] [reqmore] <eos_a> <sos_r> ok , the booking was successful . your reference number is [value_reference] . is there anything else i can help you with today? <eos_r> <sos_u> no that will be all . Thanks <eos_u> <sos_b> …

sequence 구성요소(U, B, D, A, R)에 시작(sos_?)과 끝(eos_?) special token으로 감쌈

Belief State and System Act Spans

- Belief State : (Domain-Slot, value) pair
- UBAR에서는 domain과 slot의 name을 분리
  - IF. Slot이 일치한다면 다른 Domain에서도 사용가능
- Domain, Act, Slot-values는 Special Token으로 괄호를 사용해서 학습

ex) {[domain] [inform] slot … [request] slot …}

Result

Model	Belief State	System Act	Inform	Success	BLEU	Combined
HDSA	oracle	oracle	87.9	78.0	**30.4**	113.4
DAMD	oracle	oracle	95.4	87.2	27.3	118.5
SimpleTOD	oracle	oracle	92.3	85.8	18.67	107.7
UBAR (ours)	oracle	oracle	**96.9**	**92.2**	28.6	**123.2**
SFN+RL	oracle	generated	82.7	72.1	16.3	93.7
HDSA	oracle	generated	82.9	68.9	**23.6**	99.5
ARDM	oracle	-	87.4	72.8	20.6	100.7
DAMD	oracle	generated	89.2	77.9	18.6	102.2
SimpleTOD	oracle	generated	88.9	67.1	16.9	94.9
SOLOIST	oracle	-	89.6	79.3	18.0	102.5
UBAR (ours)	oracle	generated	**94.0**	**83.6**	17.2	**106.0**
SFN+RL	generated	generated	73.8	58.6	16.9	83.0
DAMD	generated	generated	76.3	60.4	16.6	85.0
SimpleTOD	generated	generated	84.4	70.1	15.0	92.3
SOLOIST	generated	-	85.5	72.9	16.5	95.7
UBAR (ours)	generated	generated	**95.4**	**80.7**	**17.0**	**105.1**

Belief	Act	Inf.	Succ.	BLEU	Comb.
oracle	oracle	95.4	91.4	28.8	122.2
oracle	generated	92.7	81.0	16.7	103.6
generated	generated	95.7	81.8	16.5	105.7

Table 2: UBAR in different settings on MultiWOZ 2.1.

다른 End-to-End Model인 SimpleTOD, SOLOIST보다 좋은 성능을 보인다.

DST 성능측정

Model	Joint Accuracy (%)	
	MultiWOZ 2.0	MultiWOZ 2.1
TRADE	48.62	45.60
DSTQA	51.44	51.17
DST-Picklist	-	53.3
SST	-	55.23
SimpleTOD	-	55.72
DST-UBAR	**52.59**	**56.20**

Table 3: Comparison of Dialog state tracking (DST) on MultiWOZ 2.0 and 2.1.

Analysis and Discussion

어떤 종류의 dialog cotext를 사용했는지

#Turns	Belief	Act	Inf.	Succ.	BLEU	Comb.
All	GT	GT	88.4	76.6	**17.6**	100.1
All	GT	Gen	95.4	**82.3**	17.2	**106.1**
All	Gen	Gen	**95.4**	80.7	17.0	105.1
Prev	GT	GT	87.2	75.3	**16.8**	98.0
Prev	GT	Gen	92.7	**79.0**	16.6	**102.5**
Prev	Gen	Gen	**92.7**	77.7	16.4	101.6

Model	Belief State	System Act	Inform	Success	BLEU	Combined
SimpleTOD	oracle	generated	88.9	67.1	16.9	94.9
SOLOIST	oracle	-	89.6	79.3	18.0	102.5

Analysis and Discussion

Trun-level VS Session-level

Model	Context	Inf.	Succ.	BLEU	Comb.
URUR	GT	82.6	73.1	**17.0**	94.8
URUR	Gen	**91.2**	**79.5**	16.5	**101.8**
UBAR	U&R	92.5	70.8	14.3	95.9
UBAR	B&A	**94.1**	**77.1**	**16.3**	**101.9**

Table 5: Results in the end-to-end setting. URUR is trained in turn-level. *GT* or *Gen* means it uses ground truth or generated responses in its context. *U&R* denotes the context of UBAR only consists of user utterances and generated responses. *B&A* denotes the context of UBAR only consists of belief states, database results and system acts.

Analysis and Discussion

unseen domain을 얼마나 잘 detection 하는지

Evaluation on 4 Domains	Except Hotel	Except Train	Except Attraction	Except Restaurant	Except Taxi
Base Model trained in-domain	99.03	99.40	99.68	**101.87**	95.30
Few-shot BM on new domain	84.88	84.66	96.96	**100.79**	88.14
Evaluation on New Domain	Hotel	Train	Attraction	Restaurant	Taxi
Zero-shot BM	58.40	**64.19**	49.96	40.55	59.03
Few-shot BM on new domain	78.48	70.52	**79.51**	74.19	74.02
UBAR on all domains	103.06	**106.40**	102.24	104.44	103.20

Table 6: Results of domain transfer. The first row is the base model trained on the four domains and evaluated in-domain. The second row is the results of the base model fine-tuned with 100 new domain examples on the four domains. The last three rows are evaluations on the new domains with zero-shot or few-shot BM or UBAR trained on full data, respectively.

Conclusion

현재 기준 End-to-End GPT-2 based Task-oriented dialog system SOTA, DST 5위 모델

본 논문은 이전 연구들인 SimpleTOD, SOLOIST들의 문제점을 지적 후 해결방안 제안

Task-oriented dialog system의 연구를 진행할 때 Turn-level이 아닌 Session-level이 중요

Dialog history보다 사이사이에 숨어있는 Belief State, System act 정보들이 중요

감사합니다