Findings of ACL-IJCNLP 2021

HyKnow: End-to-End Task-Oriented Dialog Modeling with Hybrid Knowledge Management

Silin Gao[1*], Ryuichi Takanobu[1*], Wei Peng[2], Qun Liu[2], Minlie Huang[1†]
[1] CoAI Group, DCST, IAI, BNRIST, Tsinghua University, Beijing, China
[2] Huawei Technologies, Shenzhen, China
[1] gsl16@tsinghua.org.cn, gxly19@mails.tsinghua.edu.cn, aihuang@tsinghua.edu.cn
[2] peng.wei1@huawei.com, qun.liu@huawei.com

Natural Language Processing and Artificial Intelligence Lab

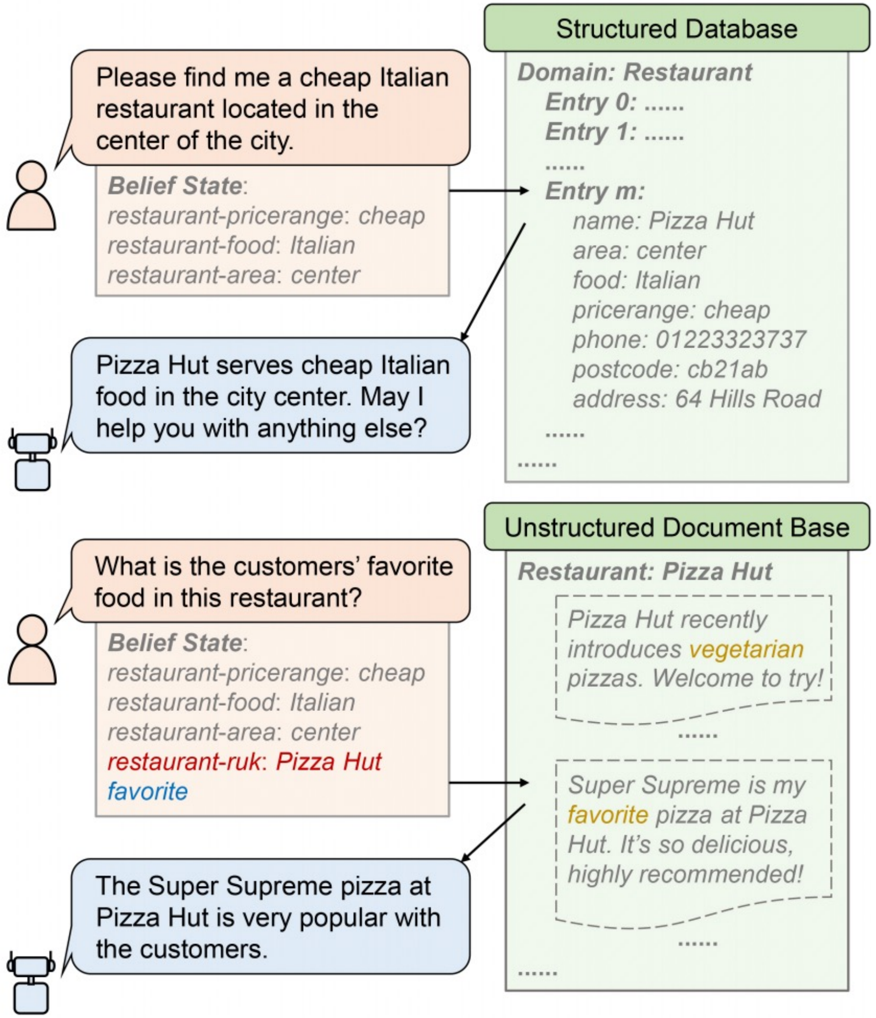
Suhyun Son

Natural Language Processing & Artificial Intelligence

KOREA UNIVERSITY 고려대학교 KOREA UNIVERSITY

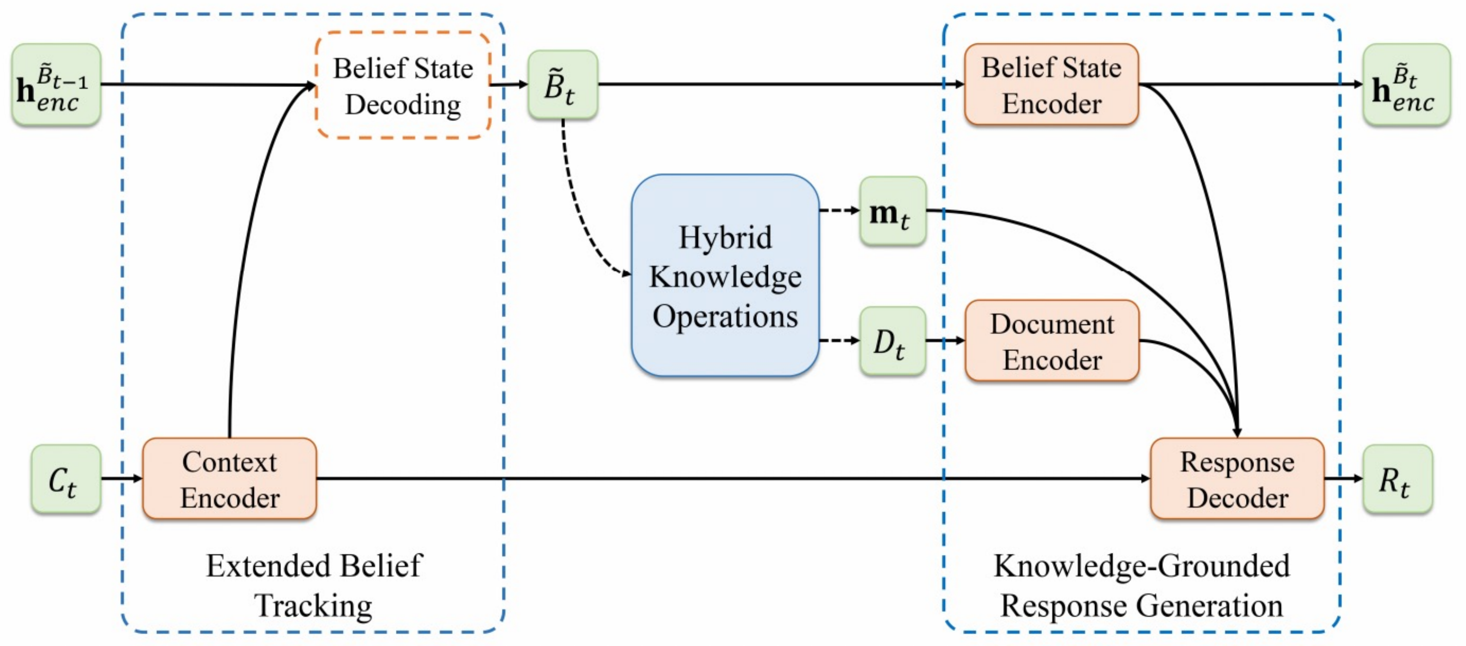
1. HyKnow

2. 적용할 수 있는 부분



- we formulate a task of modeling TOD grounded on both structured and unstructured knowledge.
- In turns involving structured knowledge,
  the system needs to track the user goals as triples and use them to perform database queries
- in turns involving unstructured knowledge,
  the system manages a document base to retrieve relevant references

- This model extends the belief state to handle TODs grounded on hybrid knowledge, and further uses the extended belief state to perform both database query and document retrieval
- We consider two implementations of our system, with different schemes of extended belief state decoding.
- HyKnow is the first end-to-end model to jointly optimize dialog modeling grounded on the two kinds of knowledge.



1. Extended Belief Tracking

- **define an extended belief state $\widetilde{B}_t$**

- Structured knowledge:

$\widetilde{B}_t$ = Origianl B_t

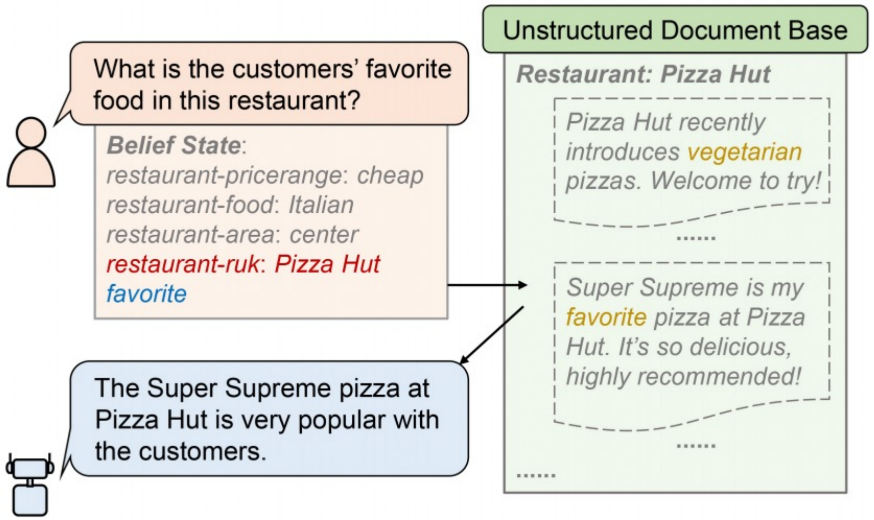
- Unstructured knowledge

1) additional slot *ruk* (requires unstructured knowledge)

→ Original state + additional triple = DSV_t

2) topic of U_t

→ As a word sequence T_t



1. Extended Belief Tracking

- **decode the current extended belief state** $\widetilde{B}_t$

dialog context C_t

previous extended belief state $\mathbf{h}_{enc}^{\widetilde{B}_{t-1}}$

- Single Decoder

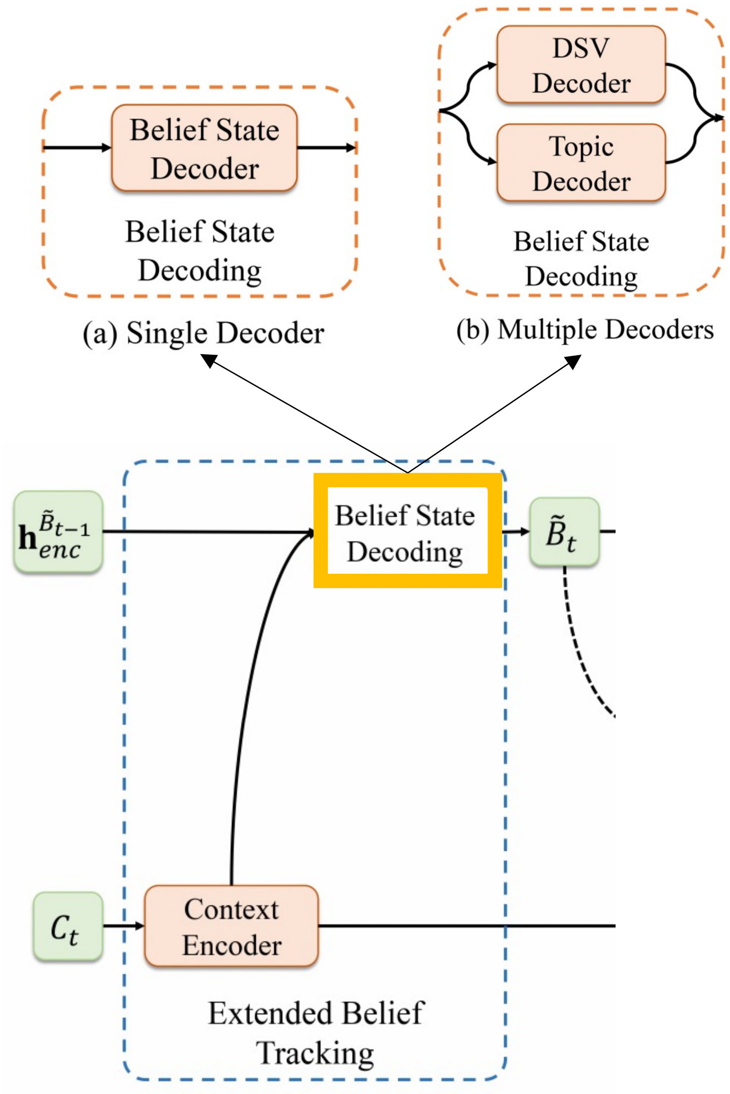
: generate the whole $\widetilde{B}_t$

$$\widetilde{B}_t = \text{Seq2Seq}^{(b)}(C_t|\mathbf{h}_{enc}^{\widetilde{B}_{t-1}})$$

- Multiple Decoder

: generate DSV_t and T_t separately

$$DSV_t = \text{Seq2Seq}^{(dsv)}(C_t|\mathbf{h}_{enc}^{\widetilde{B}_{t-1}}),$$
$$T_t = \text{Seq2Seq}^{(t)}(C_t|\mathbf{h}_{enc}^{\widetilde{B}_{t-1}}),$$
$$\widetilde{B}_t = [DSV_t, T_t].$$



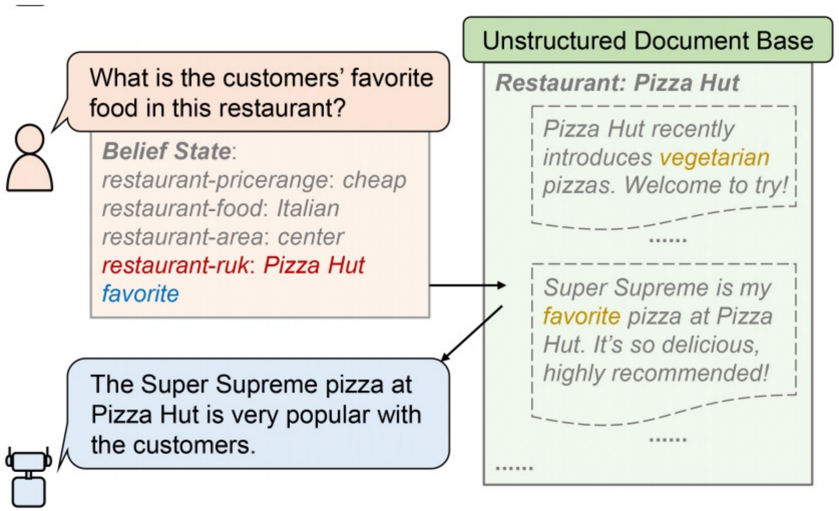
2. Hybrid Knowledge Operations

we conduct both DB query and document retrieval to get the query result mt and the relevant document

In the operation of DB query, we simply match the original triples in Bet with the DB entries

In the operation of document retrieval,

1) extract the topic of each document as its retrieval index
2) match the domain, entity and extracted topic of each document, and select the best-matched one

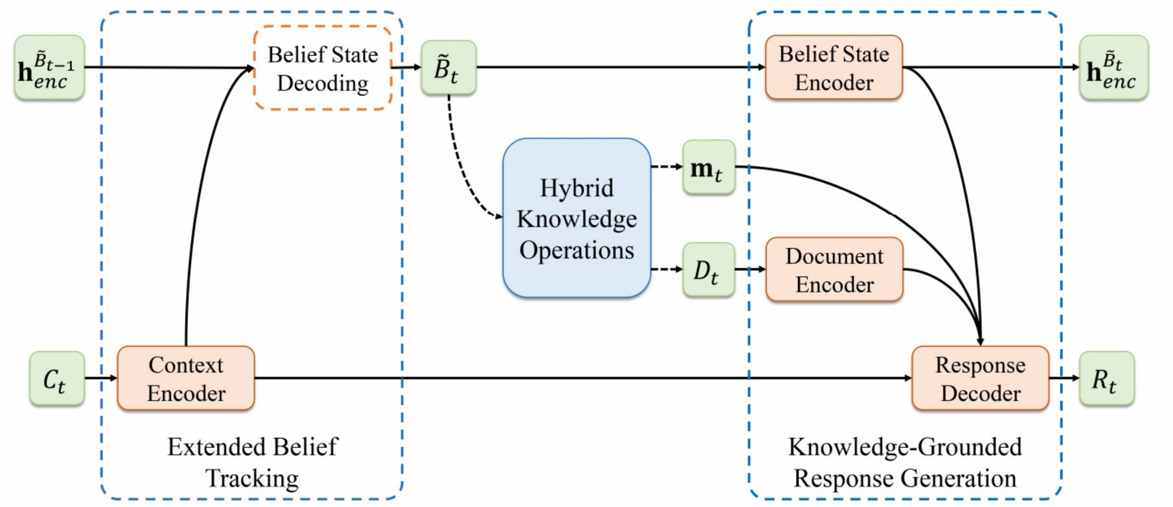


3. Knowledge-Grounded Response Generation

we use the belief state encoder and the document encoder → $\mathbf{h}_{enc}^{\widetilde{B}_t}$ $\mathbf{h}_{enc}^{D_t}$

we use the response decoder to generate the system response

$$\mathbf{h}_{enc}^{\widetilde{B}_t} = \text{Encoder}^{(\text{b})}(\widetilde{B}_t),$$
$$\mathbf{h}_{enc}^{D_t} = \text{Encoder}^{(\text{d})}(D_t),$$
$$R_t = \text{Seq2Seq}^{(r)}(C_t|\mathbf{h}_{enc}^{\widetilde{B}_t}, \mathbf{h}_{enc}^{D_t}, \mathbf{m}_t),$$



(Inform + Success) * 0.5 + BLEU

E2E TOD Models

Model	Pretrained LM	Inform	Success	BLEU	METEOR	ROUGE-L	Combined
UniConv	none	71.5	61.8	18.5	37.8	40.5	85.7
LABES-S2S	none	76.5	65.3	17.8	36.8	39.9	88.7
UniConv + BDA	-	72.0	62.6	16.9	35.7	38.9	84.2
LABES-S2S + BDA	-	77.1	66.2	15.7	33.8	37.8	87.4
HyKnow (Single)	none	**81.9**	**68.3**	**19.0**	**38.5**	40.9	**94.1**
- w/o Joint Optim	none	78.5	65.7	18.3	36.9	39.6	90.4 (-3.7)
HyKnow (Multiple)	none	79.1	67.6	18.7	38.1	**41.0**	92.1
- w/o Joint Optim	none	77.7	65.4	18.0	36.6	39.5	89.6 (-2.5)
SimpleTOD	GPT-2	81.7	67.9	14.5	34.2	37.0	89.3
SimpleTOD + BDA	-	83.3	68.6	14.8	33.6	36.5	90.8

Table 1: End-to-end evaluation results on modified MultiWOZ 2.1. "+" denotes the combination of Beyond Domain APIs (BDA) with E2E TOD models. Best results among light-weight systems (i.e. above internal dividing line) are marked in bold. Evaluation metrics are described and marked in bold in Sec. 6.1.

Model	Pretrained LM	Joint Goal
TRADE	none	42.9
UniConv	none	45.5
LABES-S2S	none	46.0
TRADE + BDA	-	43.8
UniConv + BDA	-	46.5
LABES-S2S + BDA	-	46.8
HyKnow (Single)	none	**48.0**
- w/o Joint Optim	none	46.2 (-1.8)
HyKnow (Multiple)	none	47.6
- w/o Joint Optim	none	45.6 (-2.0)
TripPy	BERT	50.4
SimpleTOD	GPT-2	48.4
TripPy + BDA	-	51.2
SimpleTOD + BDA	-	49.8

Table 3: Original turns' belief tracking results on modified MultiWOZ 2.1. "+" denotes the combination of BDA with DST/E2E models. The best result among light-weight systems (i.e. above internal dividing line) is marked in bold. The evaluation metric is described and marked in bold in Sec. 6.3.

Model	Type	MRR@5	R@1
TF-IDF	standard IR	68.7	54.1
BM25	standard IR	69.2	52.5
BDA	classification	80.6	69.8
HyKnow (Single)	topic match	**81.7**	**80.2**
- w/o Joint Optim	topic match	80.1 (-1.6)	77.8 (-2.4)
HyKnow (Multiple)	topic match	81.1	79.5
- w/o Joint Optim	topic match	79.7 (-1.4)	77.4 (-2.1)

Table 4: Newly inserted turns' document retrieval results on modified MultiWOZ 2.1. Best results are marked in bold. Evaluation metrics are described and marked in bold in Sec. 6.3.

Pros	Cons
• Structured knowledge(DB)뿐 만 아니라 unstructured knowledge도 함께 사용 가능 • Utterance에 따라 필요한 knowledge를 가져와서 사용 가능 • 코드가 공개되어 있고, 모델이 크게 복잡하지 않아서 비교적 쉽게 구축이 가능 • Structured knowledge에 필요한 정보가 존재하는지에 대한 slot을 추가함으로써 실제 inference때 확장 가능성 多	• Document retrieval을 위한 Document Knowledge base가 구축되어야 있어야 함

1. 업데이트가 되지 않은 DB에서 최신정보를 가져올 수 없을 때

Ex) user: 요즘 유행하는 노래 틀어줘

→ DB에 최신정보가 없을 때, unstructured document base에서 가장 적합한 문서를 가져와서 정보 활용 가능

2. DB에서 정확한 정보를 가져올 수 없을 때

Ex) user: 블핑 노래 틀어줘

→ 기존의 경우라면 "블핑"이 DB에 명시되어 있지 않으면 그 결과를 제대로 가져오지 못하지만, "블핑"과 "블랙핑크"의 연관성을 통해 유사한 document를 가져와서 정보 활용 가능

Thank you