

Recent Advances and Challenges in Task-oriented Dialogue Systems

집현전 9조

김도은, 이지현, 권다영, 이강희, 최윤진

2022.05.15

목차

1. What is Task-oriented Dialogue System (TOD)?
2. Modules of TOD
 - a. Natural Language Understanding (NLU)
 - b. Dialogue State Tracking (DST)
 - c. Dialogue Policy Learning
 - d. Natural Language Generation (NLG)
 - e. Evaluation of TOD
3. Dataset
4. Challenges & Future Trends

1. What is TOD?

집현전 시즌 2 TOD 영상

Task-oriented Dialogue System에 대한 넓고 얇은 지식들



집현전 증급반 4조 - 금빛나, 김도은, 김종우

발표자: 김도은, 금빛나 (NLG)

2021.07.04

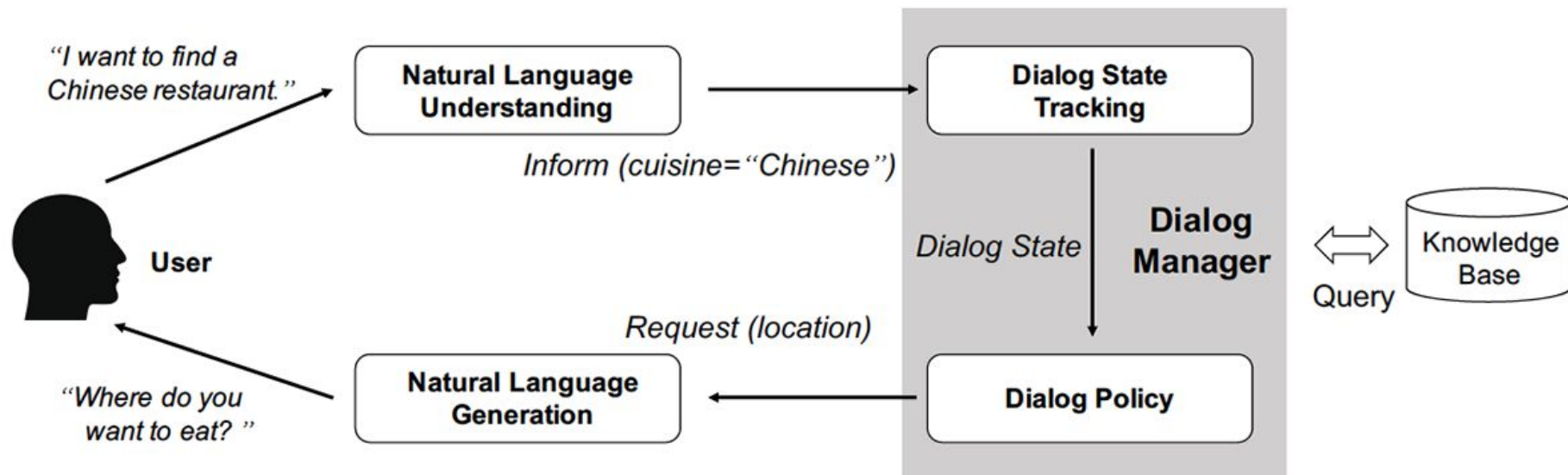
https://youtu.be/_KngLJQj2T0

1. What is TOD?

TOD(Task Oriented Dialogue System): 특정 Task를 달성하기 위한 Dialogue System



1. What is TOD



1. What is TOD?

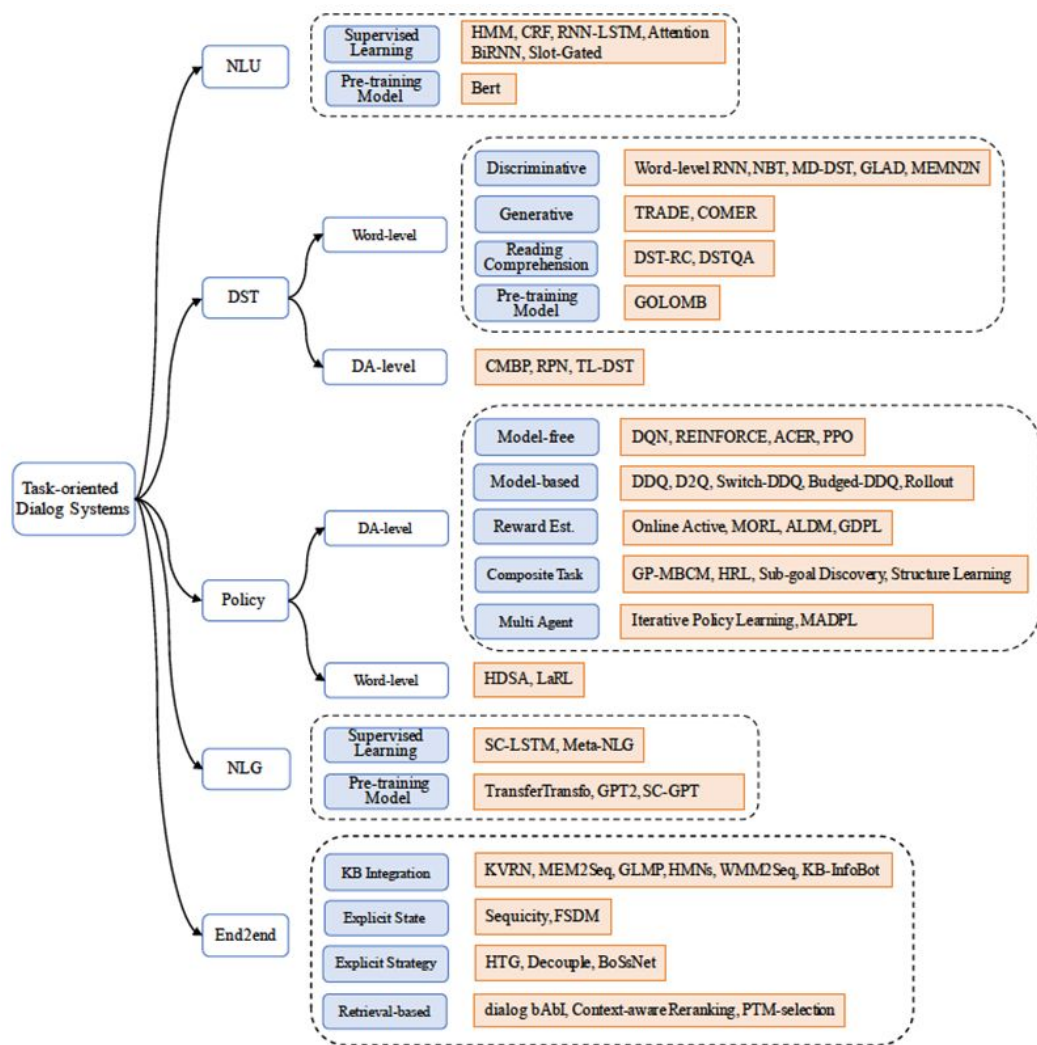
Challenges of TOD

- Data Efficiency:
 - Open-Domain보다 어려운 Data 수집.
 - Low-resource training의 필요성.
- Multi-turn Dynamics:
 - 여러 turns의 발화 속에서 자연스럽게 사용자의 목적을 달성해야 함
- Ontology Integration:
 - 다양한 task에서도 활용이 가능한 모델
- Evaluation:
 - 더 효율적인 평가 방법

2. Modules of TOD

2. Modules of TOD

Overview of TOD Modules



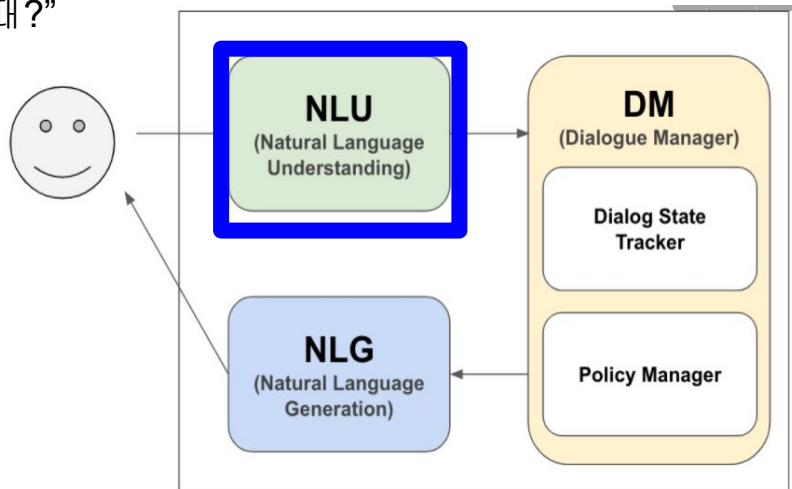
2. Modules of TOD

Natural Language Understanding

- 문장 입력에 대한 의미를 해석하는 과정
- 문장의 의도 (Intent)와 문장을 이해하기 위한 주요 값 (Slot) 을 추출
- [In] Text, Voice 형태의 문장 → [Out] 문장의 의미 (Intent, Slot, domain)

{intent : SearchWeather, Slots : (“내일”, SearchDate)}

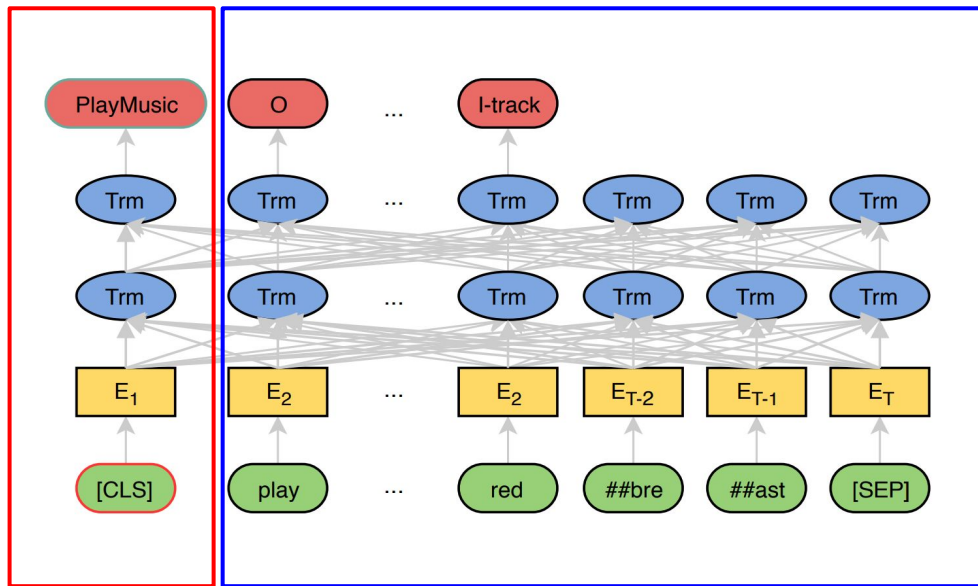
“내일 날씨는 어때?”



2. Modules of TOD

Natural Language Understanding

BERT for Joint Intent Classification and Slot Filling



- Intent Classification과 Slot filling은 Goal-oriented Dialogue System 에서 주요한 Task
- BERT Pretrained Model을 사용하여 두 task를 Joint learning하여 큰 성능 향상을 보여줌

Experiment

- Dataset :
 - snips (식당 예약) : 72개 slot label, 7개 intent types
 - atis (항공권 예약) : 120개 slot label, 21개 intent types
- Metrics :
 - Intent - Accuray / Slot - F1 score / Sentence-level semantic frame accuracy
- Model :
 - uncased BERT-base (12 layers, 768 hidden states, 12 heads)
- 결과 :

| Models | Snips | | | ATIS | | |
|-------------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | Intent | Slot | Sent | Intent | Slot | Sent |
| RNN-LSTM (Hakkani-Tür et al., 2016) | 96.9 | 87.3 | 73.2 | 92.6 | 94.3 | 80.7 |
| Atten.-BiRNN (Liu and Lane, 2016) | 96.7 | 87.8 | 74.1 | 91.1 | 94.2 | 78.9 |
| Slot-Gated (Goo et al., 2018) | 97.0 | 88.8 | 75.5 | 94.1 | 95.2 | 82.6 |
| Joint BERT | 98.6 | 97.0 | 92.8 | 97.5 | 96.1 | 88.2 |
| Joint BERT + CRF | 98.4 | 96.7 | 92.6 | 97.9 | 96.0 | 88.6 |

2. Modules of TOD

DST (Dialogue State Tracking)

- 문장이 여러번 반복되는 Multi-Turn 대화 시스템에서 State (slot:value)를 추적하는 과정
- 즉, 여러번의 대화를 통해 최종 목적지로 갈 수 있도록 Slot들의 State를 관리하는 역할
- [in] dialogue, turn + previous states → [out] dialogue state (slot:value)

user 안녕하세요.

B_1 []

sys 네. 안녕하세요. 무엇을 도와드릴까요?

user 서울 중앙에 위치한 호텔을 찾고 있습니다. 외국인 친구도 함께 갈 예정이라서 원활하게 인터넷을 사용할 수 있는 곳이었으면 좋겠어요.

B_2 [(숙소-지역, 서울 중앙), (숙소-종류, 호텔), (숙소-인터넷 가능, yes)]

sys 네. 확인해보겠습니다. 혹시 추가로 필요하신 사항이 있으실까요?

user 음.. 예약 인원은 총 8명 이고요. 아. 가격대는 크게 상관 없습니다.

B_3 [(숙소-지역, 서울 중앙), (숙소-종류, 호텔), (숙소-인터넷 가능, yes), (숙소-예약 명수, 8), (숙소-가격대, dontcare)]

sys 네. 확인 감사합니다. 숙박을 원하시는 요일과 기간 같이 확인 부탁드립니다.

user 아. 중요한 걸 깜빡했네요. 일요일에 2일간 예약하고 싶습니다.

B_4 [(숙소-지역, 서울 중앙), (숙소-종류, 호텔), (숙소-인터넷 가능, yes), (숙소-예약 명수, 8), (숙소-가격대, dontcare), (숙소-예약 요일, 일요일), (숙소-예약 기간, 2)]

✓ Turn = System - User 가 한번씩 주고받는 대화 Pair

2. Modules of TOD

DST (Dialogue State Tracking)

DST 접근법

Predefined Ontology



- 미리 dialogue ontology를 구축해두고 상황에 맞는 답을 찾아내는 방식
- 모든 경우의수를 고려할 수 없어
Unseen domain, slot, value가 주어졌을 때 아웃풋을 만들 수 없음

Open Vocabulary



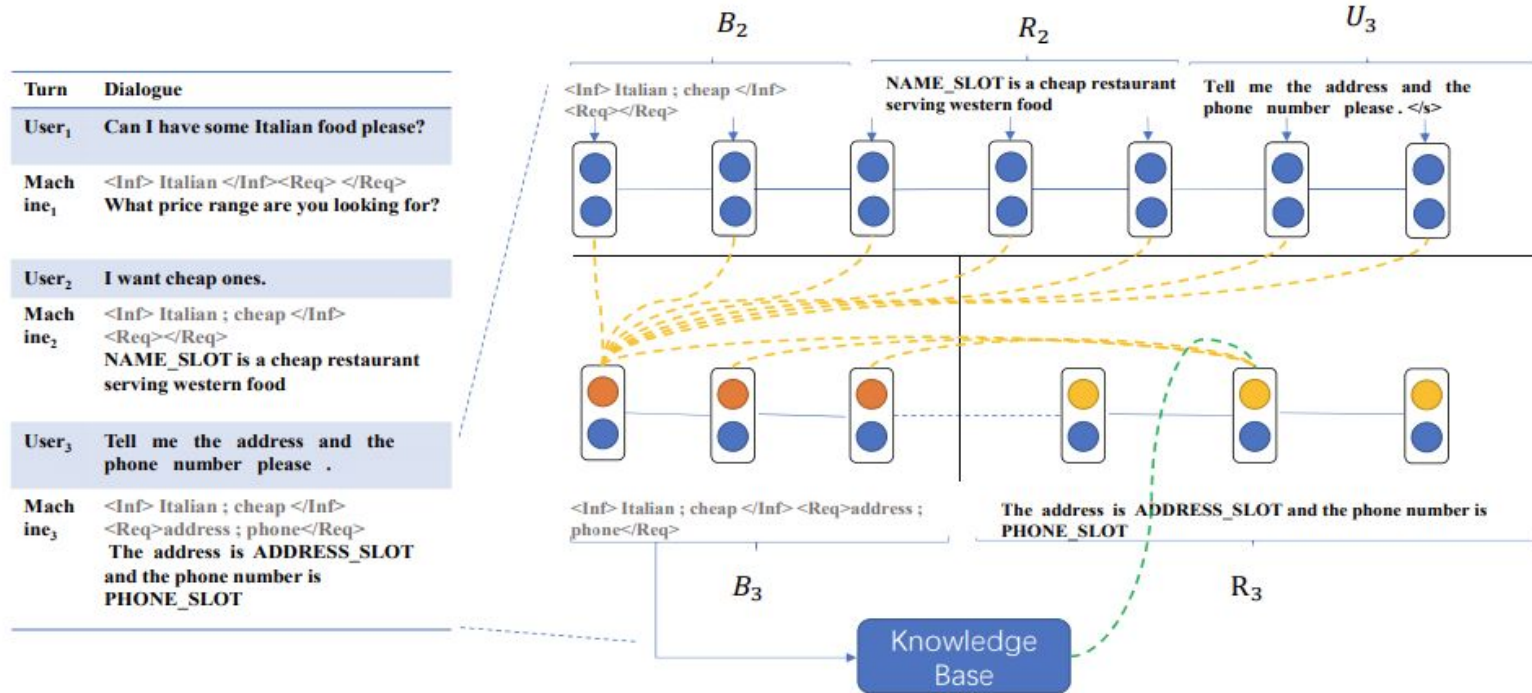
- Input에 따라 state를 즉각적으로 생성하는 방식
- 비교적 유연한 방식 (실제 대화에 더 적합방식)
- Unseen domain, slot, value가 주어졌을 때 아웃풋 생성이 가능함

2. Modules of TOD

DST (Dialogue State Tracking)

Open Vocab DST Model

B : Dialogue State
R : System 발화
U : User 발화



2. Modules of TOD

Dialogue Policy Learning - [Markov Property]

- Policy Learning: DST 에서 받은 State 정보를 바탕으로 System Action을 정하는 ‘policy’를 학습
- ‘Dialogue States들은 **Markov Property**를 지닌다’

user 안녕하세요.

B_1 []

sys 네. 안녕하세요. 무엇을 도와드릴까요?
user 서울 중앙에 위치한 호텔을 찾고 있습니다. 외국인 친구도 함께 갈 예정이라서 원활하게 인터넷을 사용할 수 있는 곳이었으면 좋겠어요.

B_2 [(숙소-지역, 서울 중앙), (숙소-종류, 호텔), (숙소-인터넷 가능, yes)]

sys 네. 확인해보겠습니다. 혹시 추가로 필요하신 사항이 있으실까요?
user 음.. 예약 인원은 총 8명 이고요. 아. 가격대는 크게 상관 없습니다.

B_3 [(숙소-지역, 서울 중앙), (숙소-종류, 호텔), (숙소-인터넷 가능, yes), (숙소-예약 명수, 8), (숙소-가격대, dontcare)]

sys 네. 확인 감사합니다. 숙박을 원하시는 요일과 기간 같이 확인 부탁 드립니다.
user 아. 중요한 걸 깜빡했네요. 일요일에 2일간 예약하고 싶습니다.

B_4 [(숙소-지역, 서울 중앙), (숙소-종류, 호텔), (숙소-인터넷 가능, yes), (숙소-예약 명수, 8), (숙소-가격대, dontcare), (숙소-예약 요일, 일요일), (숙소-예약 기간, 2)]

☑ Turn = System - User 가 한번씩 주고받는 대화 Pair

State Tracking 예시 (revisited)

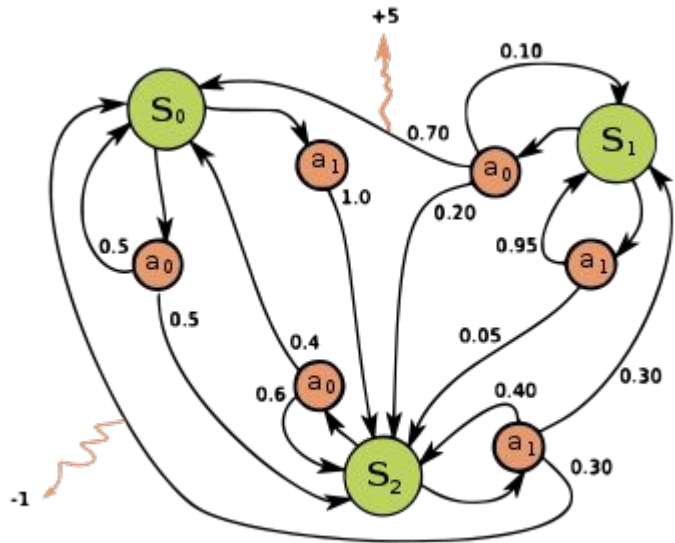
Markov Property

- 시점 t 에서의 state 정보는 이전 state들의 정보를 모두 포함하고 있다.
- 시점 t 에서의 state은 dialogue history의 representation 이다.
- 시점 t 에서의 state을 기반으로 다음 system action을 예측하기 위해서 이전 시점의 states들을 참조할 필요가 없다.

2. Modules of TOD

Dialogue Policy Learning - [MDP-based RL]

- Policy Learning: DST 에서 받은 State 정보를 바탕으로 System Action을 정하는 'policy'를 학습
- 'Dialogue States들은 **Markov Property**를 지닌다' 라는 전제로 **Markov Decision Process (MDP)** 기반의 **강화학습**이 가능하다



States, Actions, and Rewards in MDP

- Policy Learning is a mapping of $\pi : S \times A \rightarrow [0, 1]$, where S is the set of states and A is set of actions
- For every state $s \in S$, policy can assign the probability of taking the action to each $a \in A$, i.e., $\pi(a|s)$
- There will be always one or more deterministic optimal policy for any MDP

2. Modules of TOD

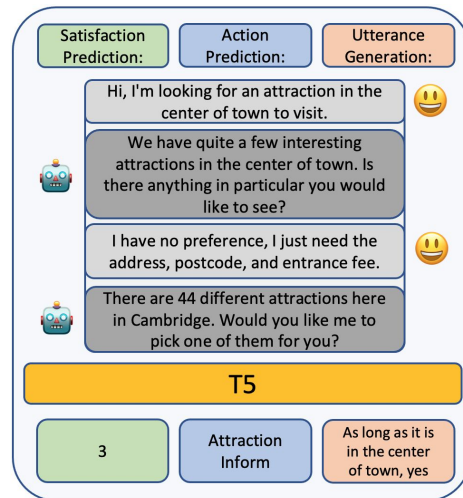
Dialogue Policy Learning - [User Simulation]

- Policy Learning: DST 에서 받은 State 정보를 바탕으로 System Action을 정하는 ‘policy’를 학습
- ‘Dialogue States들은 **Markov Property**를 지닌다’ 라는 전제로 **Markov Decision Process (MDP)** 기반의 **강화학습**이 가능하다
- 실제 사용자가 action에 대한 reward를 준다. 하지만 TOD 데이터는 비싸고 양이 적다. → User Simulator 의 필요성.
- User Simulator가 강화학습 훈련시, Environment가 되어줄 수 있다.

A Multi-Task Based Neural Model to Simulate Users in Goal-Oriented Dialogue Systems

To Eun Kim
University College London
London, United Kingdom
to.kim.17@ucl.ac.uk

Aldo Lipani
University College London
London, United Kingdom
aldo.lipani@ucl.ac.uk



2. Modules of TOD

Dialogue Policy Learning - [Deep Dyna-Q (2018)]

Deep Dyna-Q: Integrating Planning for Task-Completion Dialogue Policy Learning

Baolin Peng* Xiujun Li† Jianfeng Gao† Jingjing Liu† Kam-Fai Wong*‡ Shang-Yu Su§

†Microsoft Research, Redmond, WA, USA

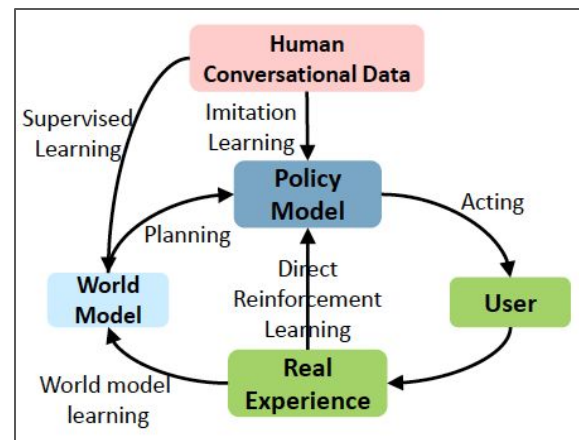
*The Chinese University of Hong Kong, Hong Kong

‡MoE Key Lab of High Confidence Software Technologies, China

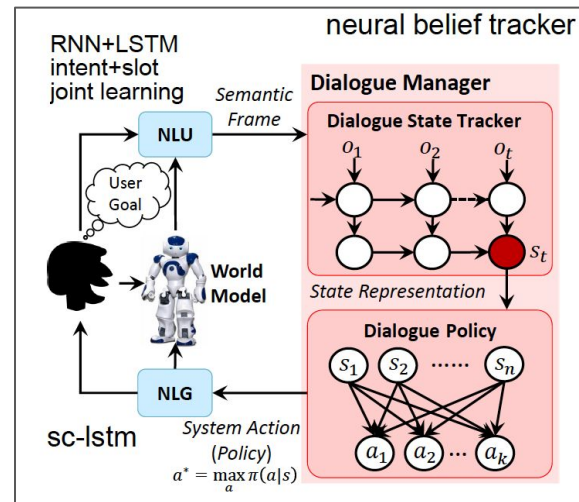
§National Taiwan University, Taipei, Taiwan

훈련 방법

- Direct Reinforcement Learning
 - MDP 기반의 epsilon greedy policy
- World Model Learning
 - World Model은 environment를 simulate 하기 위한 모듈
 - 실 사용자 데이터를 이용하여 World Model을 online으로 구축
- Planning
 - World Model로 부터 simulated response 생성
 - DQN 알고리즘으로 planning



Learning with real users via DDQ



Task completion DDQ Agent

2. Modules of TOD

Dialogue Policy Learning - [Deep Dyna-Q (2018)]

Deep Dyna-Q: Integrating Planning for Task-Completion Dialogue Policy Learning

Baolin Peng* Xiujun Li† Jianfeng Gao† Jingjing Liu† Kam-Fai Wong*‡ Shang-Yu Su§

†Microsoft Research, Redmond, WA, USA

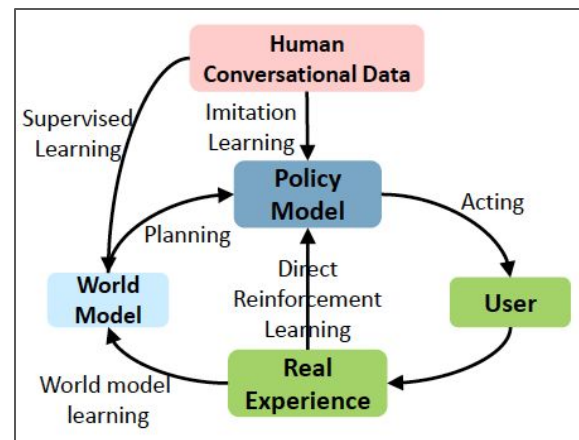
*The Chinese University of Hong Kong, Hong Kong

‡MoE Key Lab of High Confidence Software Technologies, China

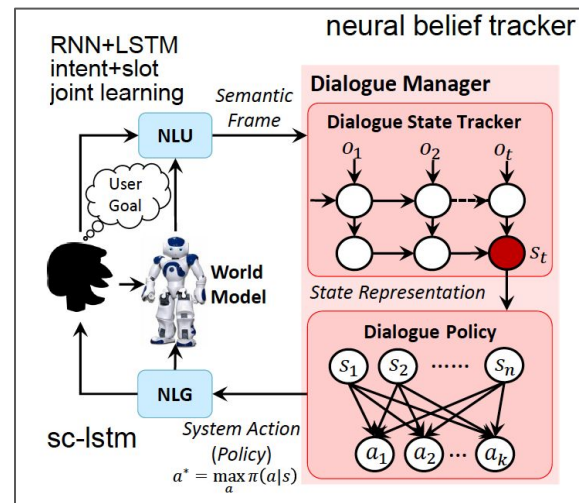
§National Taiwan University, Taipei, Taiwan

보상 체계

- Session-level Reward
 - Simulator가 Agent에게 매 turn마다 simulated response를 보내주고 대화 세션이 끝났을 때 Task의 성공 여부를 보상으로써 알려준다.
 - Task 성공 시: $2 * L$, Task 실패 시: $-L$ 의 보상 (L 은 turn의 길이)
- Turn-level Reward
 - 매 turn마다 Agent는 -1 의 보상을 받는다.
- 정리하자면, 짧은 turn 내에 Task를 달성하는 것이 보상을 최대화 하는 전략.



Learning with real users via DDQ

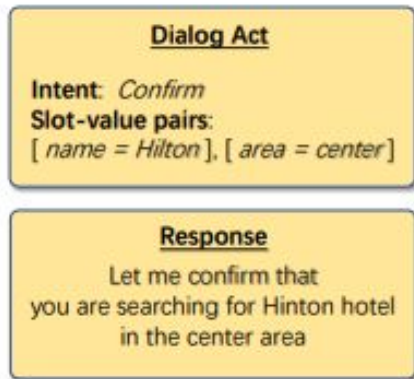


Task completion DDQ Agent

2. Modules of TOD

Natural Language Generation (NLG)

- Task-oriented dialog (TOD) 시스템에서 NLG는 dialogue act (+ DST에서 얻어진 정보) 를 natural language로 바꿔주는 모듈

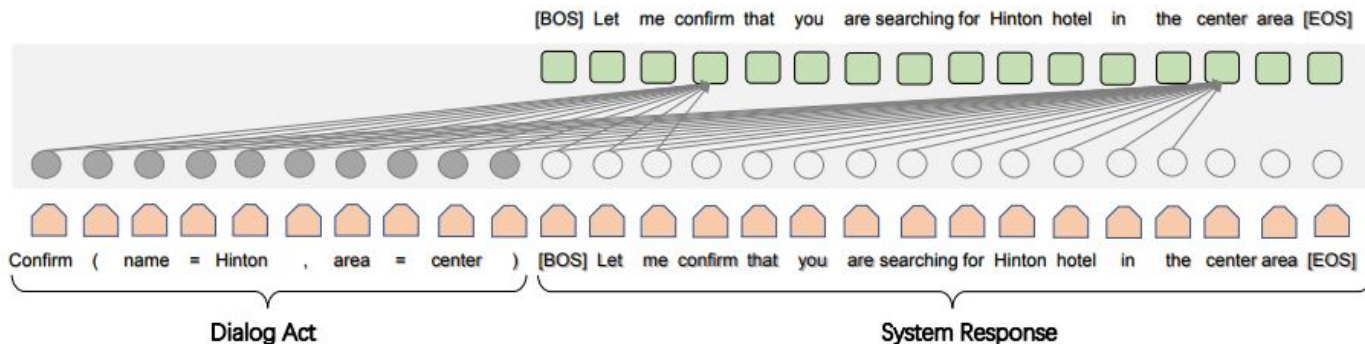


- NLG 모델의 종류
 - Template based
 - 미리 답변 템플릿을 만들어 두고, 이 템플릿에 정보를 채워넣는 방식
 - Statistical language model
 - 템플릿 없이 응답을 생성하는 방식

2. Modules of TOD

Natural Language Generation (NLG)

SC-GPT



- Semantically-Conditioned Generative Pre-Training의 줄임말.
- GPT2를 기반으로 “원하는 방향으로 대화를 생성”해 낼 수 있도록 한 모델
- Pre training을 통해 기존 모델보다 적은 도메인의 데이터, 또는 **unseen domain**의 환경에서도 잘 작동함

SC-GPT

- Pre-training

- 1 Massive Plain Language Pre-training
 - OpenWebText
- 2 Dialog-Act Controlled Pre-training
 - act-response의 형태를 포함하는 데이터 셋으로 act가 주어지면 response를 생성하는 사전학습 (Schema-Guided Dialog, MultiWOZ, Frame, Facebook Multilingual Dialogue 사용)

- Fine-tuning

- 두번째 pre-training과 동일한 방식 사용

$$p_{\theta}(x|\mathcal{A}) = \prod_{t=1}^T p_{\theta}(x_t|x_{<t}, \mathcal{A})$$

Auto-regressive
generation process

$$\mathcal{L}_{\theta}(\mathcal{D}) = \sum_{n=1}^{|\mathcal{D}|} \sum_{t=1}^{T_n} \log p_{\theta}(x_{t,n}|x_{<t,n}, \mathcal{A}_n)$$

Maximize Log likelihood

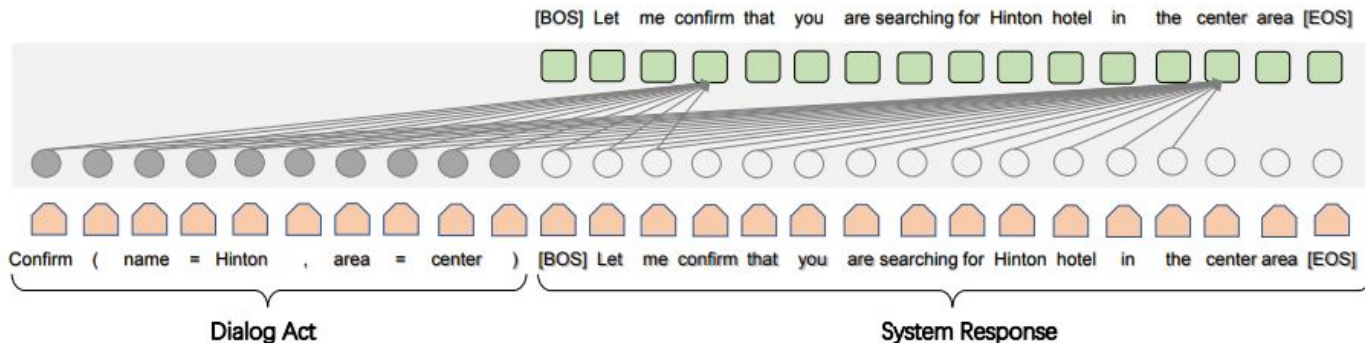
$$\mathcal{A}' = [\mathbf{I} \ (s_1 = v_1, \dots, s_P = v_P)]$$

Inform (hotel = north)

2. Modules of TOD

Natural Language Generation (NLG)

SC-GPT



- 구조의 장점

- **Flexibility** : 모델의 행동(Act)이 미리 정의되어 있을 필요가 없다. 즉 학습 때 데이터에 없었던 행동(Act)도 input으로 넣을 수 있다.
- **Controllability** : GPT2는 문장 생성에 가이드가 없는 반면, SC-GPT는 intent와 slot-value를 통해 high level guide가 있다.
- **Generalizability** : 대량의 코퍼스로 사전학습이 되어있어 일반화 성능이 좋다

2. Modules of TOD

Evaluation

Automatic Evaluation

- NLU: Slot F1-score, Intent Accuracy
- DST: Slot Accuracy, Join state Accuracy
- DPO: Inform rate(F1-score), Match rate, Task success rate
- NLG: BLEU, Perplexity

Simulated Evaluation

- Rule-based or can be learned data driven
- Task success rate, Dialogue length, Average rewards

Human Evaluation

- Task success rate, Irrelevant turn rate, Redundant turn rate, User satisfaction score
- Direct vs. Indirect

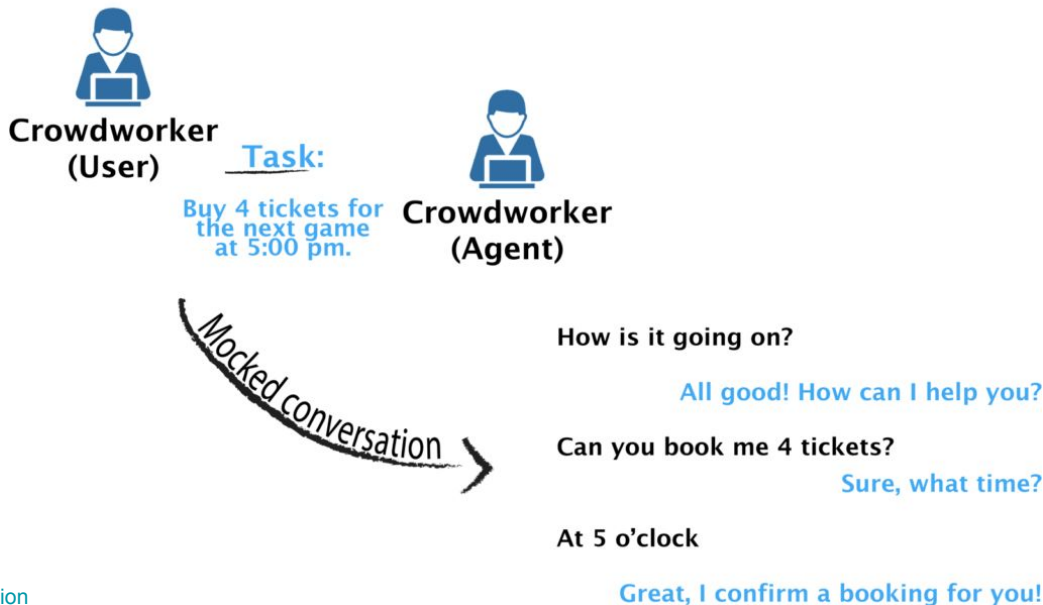
3. Dataset

3. Dataset

Human-to-Human(H2H)

- 2명의 User에게 역할(시스템/사용자)을 부여하고 대화를 진행
- 대표적인 Dataset: CamRest, WOZ, KVRiT, MultiWOZ, ETC....

Human-to-Human



3. Dataset

Human-to-Human Dataset(WoZ)

Wizard-of-Oz data collection websites

Task 02004: You are looking for and it should serve **gastropub food**. You don't care about the **price range**. You want to know the **address**.

Info Desk : Hello , welcome to the Cambridge restaurant system . You can ask for restaurants by area, price range or food type . How may I help you ?

Customer : I want a gastropub food

Info Desk : There are 4 restaurants serving gastropub food, what price range do you want ?

Next turn!

Customer : (Your response)

I don't care about the price range, just give me the address please.

Submit the HIT

3. Dataset

Human-to-Human Dataset(WoZ)

Info Desk : Hello , welcome to the Cambridge restaurant system . You can ask for restaurants by area, price range or food type . How may I help you ?
Customer : i want a gastropub food
Info Desk : There are 4 restaurants serving gastropub food, what price range do you want ?
Customer : i dont care

Next turn!

Please **modify** the following answers based on the latest customer response:

- What does user want?
What is the **food type** the user wants?
What is the **area** the user wants?
What is the **price range** the user wants?
- What does user ask?
Is the user asking for **food type** of an offered venue?
Is the user asking for **price range** of an offered venue?
Is the user asking for **area** of an offered venue?
Is the user asking for **postcode** of an offered venue?
Is the user asking for **phone number** of an offered venue?
Is the user asking for **address** of an offered venue?
Is the user mentioning any **restaurant names**?

finish

Info Desk : (Your response)

I would recommend backstreet bistro, a great gastropub restaurant in the centre. do you want their phone number ? ☐ end-of-dialogue?

3. Dataset

Human-to-Machine(H2M)

- 사전에 공개된 시스템과 사용자간의 대화를 수집
- 대표적인 Dataset: DSTC ~ DSTC10

Human-to-Machine



How is it going on?

All good! How can I help you?

Can you book me 4 tickets?

Sure, what time?

At 5 o'clock

Great, I confirm a booking for you!



**Dialogue
system**

3. Dataset

Human-to-Machine Dataset(DSTC 8)

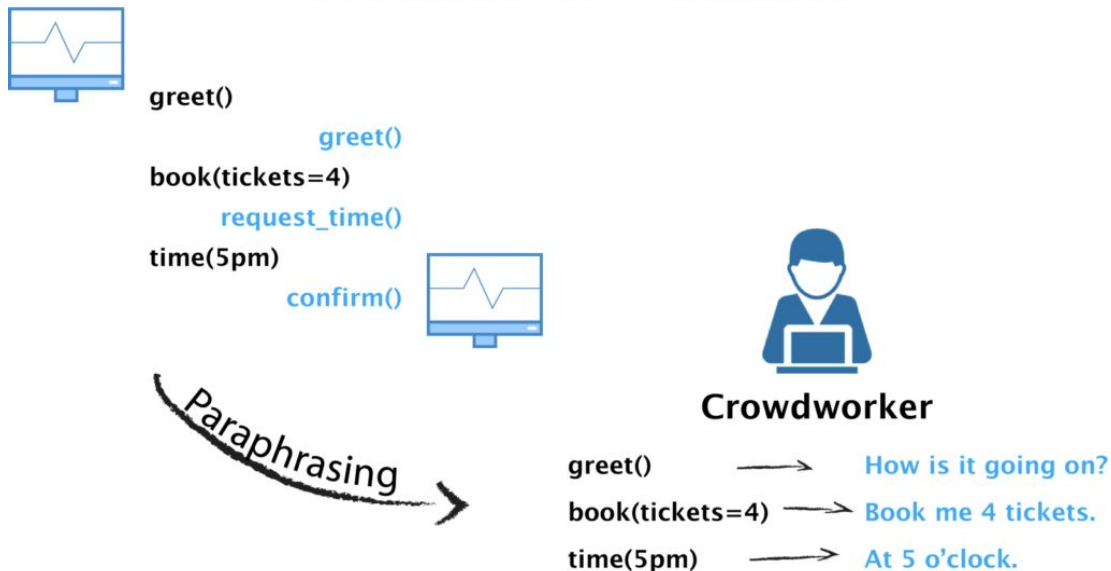


3. Dataset

Machine-to-Machine(M2M)

- 공개된 2개의 시스템에게 역할(시스템/유저)을 부여하고 대화를 수집
- 대표적인 Dataset: bAbI, SimD, AirD, SGD

Machine-to-Machine



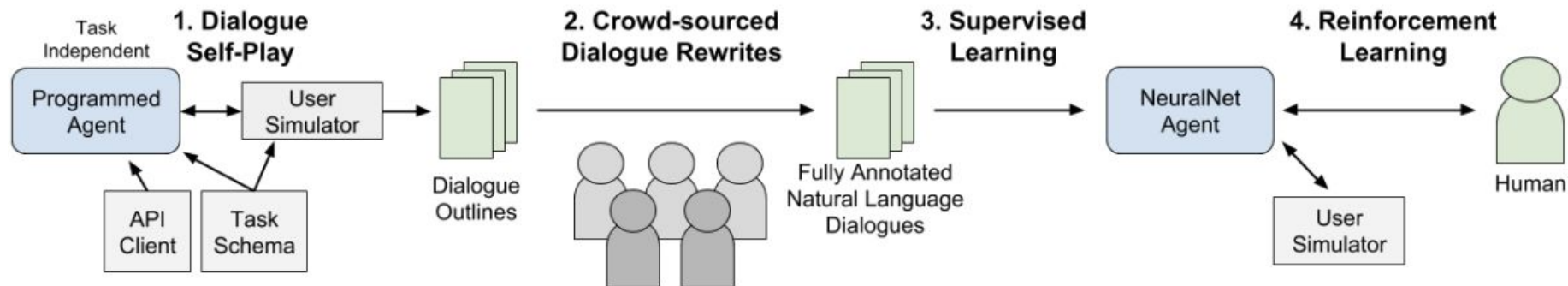
3. Dataset

Machine-to-Machine Dataset(AirD)

| Outline | | Rewrite |
|---|---|--|
| Annotations | Template utterances | NL utterances |
| S: greeting() | Greeting. | Hi, how can I help you? |
| U: inform(intent=book_movie, name=Inside Out, date=tomorrow, num_tickets=2) | Book movie with name is Inside Out and date is tomorrow and num tickets is 2. | I want to buy 2 tickets for Inside Out for tomorrow. |
| S: ack() request(time) | OK. Provide time. | Alright. What time would you like to see the movie? |
| U: inform(time=evening) | Time is evening. | Anytime during the evening works for me. |
| S: offer(theatre=Cinemark 16, time=6pm) | Offer theatre is Cinemark 16 and time is 6pm. | How about the 6pm show at Cinemark 16? |
| U: affirm() | Agree. | That sounds good. |
| S: notify_success() | Reservation confirmed. | Your tickets have been booked! |

3. Dataset

Machine-to-Machine Dataset(AirD)



4. Challenges & Future Trends

4. Challenges & Future Trends

Data Efficiency

- Issue
 - Open-domain 데이터 보다 수집이 어려움
 - User를 활용한 수집은 시간과 비용이 비쌈
- Future trends → low-resource training
 - Transfer learning
 - Unsupervised methods
 - User Simulation
 - Multi-task Learning

4. Challenges & Future Trends

Multi-turn Dynamics

- Issue
 - Open-domain과 같이 대화하는 것도 중요하지만 목표한 작업의 성공이 중요
 - 인간과 같은 방식으로 대화의 상태와 흐름을 파악이 중요(Dialog Management)
- Future trends
 - Generative DST with free-form slots
 - Dialog Planning
 - User Goal Estimation

4. Challenges & Future Trends

Ontology Integration

- Issue

- 파이프 라인 방식에서는 대부분의 Domain Schema가 사전에 정의되어 있으며 Corpus에 대한 의존도가 높음
- End-to-End 방식에서는 Context information과 Knowledge base를 학습하기 어려움

- Future trends

- Schema Integration
- Knowledge base integration

4. Challenges & Future Trends

Evaluation

- Issue
 - Single-turn evaluation 혹은 single-path evaluation만 가능
 - online evaluation의 어려움
- Future trends
 - User simulation을 활용한 evaluation
 - Automatic Metrics 외의 다양한 평가 지표 구축
 - Task Success Rate
 - Dialogue Length
 - Average Rewards
 - User Satisfaction Level

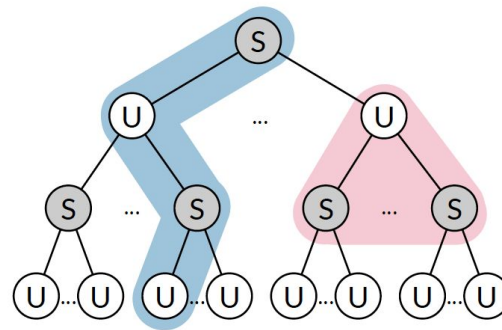


Figure 2: The space of possible dialogue states increases exponentially with the number of turns between system (S) and user (U). Evaluation is currently limited to either a single path (blue area) or a single turn (red area).

Thank YOU 😊