

#Dialogue System #Task-oriented Dialogue #Chitchat

Exploring ToD ↔ Chitchat Switching

: Current Progress, Challenges, and Next Steps

Wed, Jan 24, 2025, 1:00PM @Online

HYU NLP Lab.

Exploring ToD ↔ Chitchat Switching

Contents

1

Background & Motivation

- Why This Research Matters
- Key Gaps in Existing Studies
- Research Objectives

2

Research Progress

- Dataset Construction
- Key Challenges
- Solutions

3

Questions & Discussion

- Questions for Discussion
- Next Steps in the Research

Background & Motivation

Why This Research Matters

Key Gaps in Existing Studies

Research Objectives

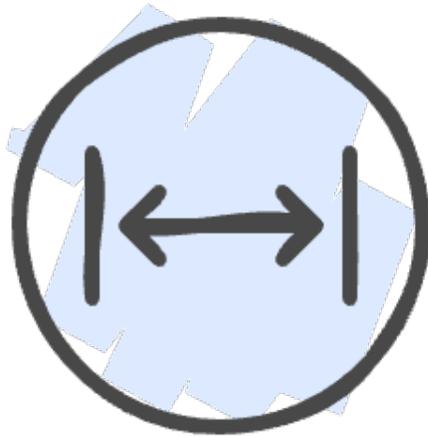
Why This Research Matters

• Background

- ToD와 Chitchat 전환이 원만히 이루어지는 (혹은 통합된) 대화 시스템을 위한 연구는 꾸준히 지속되어 옴
- 특히 최근 chat-based LLM의 General ability가 크게 향상되며 다양한 대화 유형의 통합에 대한 기대 향상
- LLM이 전통적인 ToD 시스템을 완전히 대체할 정도로 최적화되기는 어려움

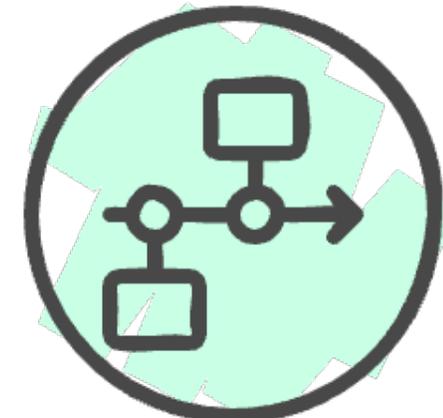
→ ToD System과 LLM 통합에 대한 연구 니즈 확대

end-2-end method



LM이 input으로 dialogue context를 받고,
알아서 적절한 response를 output

Router-based pipeline method



Input으로 받은 context에 대해
어떤 모델이 해결해야하는지 판단한 후,
판단 결과에 따라 해당 모델이 response

Key Gaps in Existing Studies

• Related Work – Summary #ToD #Chitchat

SalesAgent

1. Introduction

ToD과 Chitchat 분야의 발전은 개별적, 두 분야 간 통합 부족

2. Limitation

- 공개된 데이터셋 없음 (데이터 생성 코드만 공개)

3. Seed Dataset: SalesBot

- SGD intent space 활용 ToD
- 학습된 BlenderBot간 대화

4. Contributions

- SalesBot 2.0 (Dataset)
- SalesAgent : ReAct기반 CoT SFT 제안

Chitchat → ToD

FusedChat

1. Introduction

대부분 ToD와 ODD를 별도 취급, 두 유형의 대화 통합 연구 부족

2. Limitation

- ToD / Chitchat 발화 성격이 매우 달라서 구분이 쉬움

3. Seed Dataset: MultiWOZ

- Human(Agent)-Human(User) 대화

4. Contributions

- FusedChat (Dataset)
- Pipeline 형태 제안
 - 매 turn마다 대화 모드 분류
 - 각 모드별 SFT 모델이 답변

Chitchat → ToD | ToD → Chitchat

InterfereChat

1. Introduction

Chitchat + ToD 혼합 발화 처리 방식에 대한 연구 부족

2. Limitation

- Interfere된 user 발화는 항상 (원래의) ToD 발화 + backstory(chitchat) 형태로만 구성 → 별개 발화 이어 붙인 느낌

3. Seed Dataset: ← FusedChat

4. Contributions

- InterfereChat (Dataset)
- SimpleToD (Generative SFT) 기준 실험 결과 공유

ToD → Chitchat → ToD

Key Gaps in Existing Studies

• Related Work – SalesAgent

SalesAgent

1. Introduction

TOD과 Chitchat 분야의 발전은 개별적, 두 분야 간 통합 부족

2. Limitation

- 공개된 데이터셋 없음 (데이터 생성 코드로 확인 필요)

3. Seed Dataset: SalesBot

- SGD intent space 활용 ToD
- 학습된 BlenderBot간 대화

4. Contributions

- SalesBot 2.0 (Dataset)
- SalesAgent : ReAct기반 CoT SFT 제안

Chitchat → ToD

1. **Chitchat Dialogue Revision:** gpt-3.5 turbo 활용 대화 수정+확장 (수정 이유 생성 → 해당 이유기반 수정)
2. **Potential Intent Detection:** 대화 중 user가 암시적으로 표현하는 의도 탐지 (SGD Intent space)
3. **Dialogue Continuation:** 탐지된 의도 → 현재 주제-목표 의도 교차주제 확인 → 대화이어서 생성
4. **Transition Boundary Detection:** 의도가 명시적으로 나타나는 지점 감지 → ToD로 전환
 - 대화 탐지된 의도 전환에 적합한 턴 선택, 명시적 언급이 있는 경우 전환 시점으로 설정

SALESAGENT	SalesBot 1.0	SalesBot 2.0
<i>Turn-level</i>		
Intent Detection	31.62	39.61
Policy Selection	22.69	39.05
Both Match	17.31	32.06
<i>Dialogue-level</i>		
# Turns	7.40	12.64
Proceed TOD Rate	100.00	59.60
Naturalness ↑	55.62	78.24
Coherence ↑	55.20	79.02
Agent Consistency ↑	57.56	78.12
Agent Aggressiveness ↓	72.93	40.38
Smoothness ↑	42.93	68.58

ToD 평가지표
Chitchat 평가지표

SalesBot 1.0

U: Yes, we grew up together. We are both going to U of Michigan this fall.
 A: That's awesome! I'm hoping to start my own business soon. What do you do for a living?
 U: I work at a grocery store. What kind of business are you looking to start? I hope it works out for you.
 A: I'm thinking of starting a sports bar. I love sports and I think it would be a great place to start.

Chit-Chat Dialogue Revision



U: Yes, we grew up together. We are both going to U of Michigan this fall.
 A: That's great! What are you planning to study at U of Michigan?
 U: I haven't decided yet, but I'm leaning towards psychology. How about you?
 A: I actually graduated a few years ago, but now I'm thinking of starting my own business. I want to open a sports bar.



Potential Intent Detection

FINDRESTAURANT Triggered TOD Intent



Dialogue Continuation

U: That sounds exciting. Are you a big sports fan?
 A: Definitely! Watching sports is one of my favorite pastimes. Speaking of which, have you visited any good sports bars around here yet?
 U: No, I haven't had a chance to check any out yet. Do you have any recommendations?
 A: Definitely, there's a great place down the street called The Game Day Grill. They have amazing wings and a huge selection of beers on tap. You should definitely check it out sometime.

Transition Boundary Detector



SalesBot 2.0

U: Yes, we grew up together. ...
 ...
 U: That sounds exciting. ...
 A: Definitely! ...
 U: No, I haven't had a chance to check any out yet. Do you have any recommendations?
 A: Definitely, there's a great place ...

Key Gaps in Existing Studies

• Related Work – FusedChat

FusedChat

1. Introduction

대부분 TOD와 ODD를 별도 취급, 두 유형의 대화 통합 연구 부족

2. Limitation

- ToD / Chitchat 발화 성격이 매우 달라서 구분이 쉬움

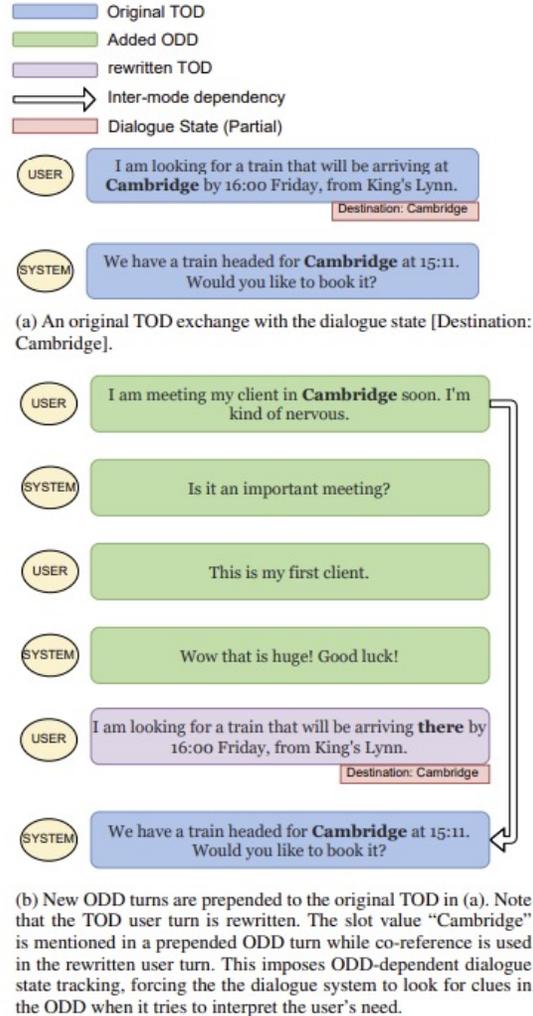
3. Seed Dataset: MultiWOZ

- Human(Agent)-Human(User) 대화

4. Contributions

- FusedChat (Dataset)
- Pipeline 형태 제안
 - 매 turn마다 대화 모드 분류
 - 각 모드별 SFT 모델이 답변

Chitchat → ToD | ToD → Chitchat



1. MultiWoZ를 Seed로 활용

2. ODD(=chitchat) 구축: human creator 활용

3. MultiWoZ 대화 전 또는 후에 Chitchat 추가

- Prepending: chitchat을 TOD 전에 추가
- Appending: chitchat을 TOD 후에 추가

4. 전체 대화 재작성

- chitchat에서 언급된 entity가 ToD에서 참조하도록 Rewrite
- Co-reference도 고려 (e.g. Cambridge → there)

Models	TOD metrics				
	Slot Accuracy	Joint SA	Inform	Success	BLEU
Two-in-one model	0.972	0.592	70.4	57.0	12.05
Classification-based model	0.973	0.600	75.1	60.9	12.17

Models	ODD metrics			
	PPL	Sensibleness	Specificity	SSA
Two-in-one model	10.49	0.52	0.47	0.50
Classification-based model	10.50	0.58	0.51	0.55

■ ToD 평가지표

■ Chitchat 평가지표

Key Gaps in Existing Studies

• Related Work – InterfereChat

InterfereChat

1. Introduction

Chitchat + ToD 혼합 발화 처리 방식에 대한 연구 부족

2. Limitation

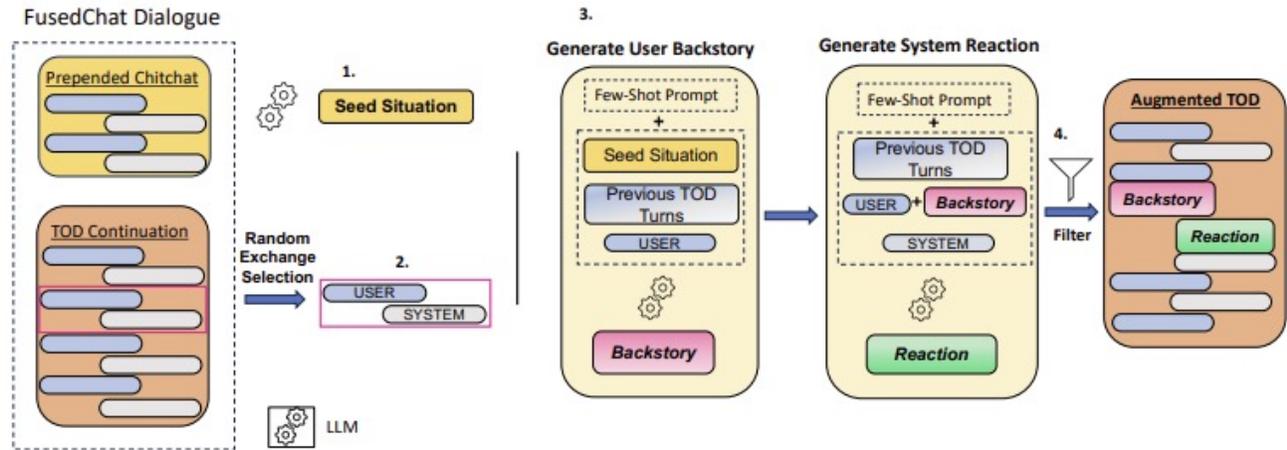
- Interfere된 user 발화는 항상 (원래의) ToD 발화 + backstory(chitchat) 형태로만 구성 → 별개 발화 이어 붙인 느낌

3. Seed Dataset: FusedChat

4. Contributions

- InterfereChat (Dataset)
- SimpleToD (Generative SFT) 기준 실험 결과 공유

CASE 2+3 (ToD → Chitchat → ToD)



1. **Seed situation 생성**: FusedChat에서 pretended 부분만 추출 → seed situation 생성
2. **Turn selection**: MultiWoZ대화에서 무작위로 1-턴 선택 → chitchat 추가
3. **Exchange augmentation**: Seed situation 기반 user backstory 발화 + 이어지는 system 발화 생성
4. **Filtering**: Rule-base로 낮은 품질의 데이터 제거 (e.g. user 원래 발화와 유사, 프롬프트 위반, ...)

	inform	success	JGA	CBE	unique tris.	BLEU-aug	BLEU-orig	BLEU-all
SimpleToD-ref	80.73	72.85	0.66	1.87	2870	-	-	-
SimpleToD	80.47	72.85	0.66	1.91	3184	0.11	0.24*	0.21
SimpleToD-fused	67.08	51.38	0.39	2.12	5674	0.04	0.2	0.17
SimpleToD-inter	82.07	73.2	0.64	2.21	7561*	0.25*	0.22	0.22

■ ToD 평가지표

■ Chitchat 평가지표

Key Gaps in Existing Studies

→ On Problem Set-up

A. *Only consider a limited subset of all switching cases*

SalesAgent: $C \rightarrow T$, FusedChat: $T \rightarrow C$ or $C \rightarrow T$, InterfereChat: $(T \rightarrow) C \rightarrow T, \dots$

B. *Only interested in response immediately after switching*

Need to think about Dialog Management perspective: the flow not recovered

- $T \rightarrow C \rightarrow T$: If agent go to C while T is unresolved, there's no way to get back to T

- $C \rightarrow T \rightarrow C$: If T is resolved, the conversation is cut off.

→ On Solution Approaches

C. *No specific solution for handling switching dialogue*

Research Objectives

→ On Problem Set-up

A. Consider all switching cases

SalesAgent: $C \rightarrow T$, FusedChat: $T \rightarrow C$ or $C \rightarrow T$, InterfereChat: $(T \rightarrow) C \rightarrow T, \dots$

B. Consider 'dialogue flow' before/after the switching

Need to think about Dialog Management perspective: the flow not recovered

- $T \rightarrow C \rightarrow T$: If agent go to C while T is unresolved, there's no way to get back to T
- $C \rightarrow T \rightarrow C$: If T is resolved, the conversation is cut off.

→ On Solution Approaches

C. (1) Detect Switching → (2) Find Recover Point → (3) Recover the Flow

Conversational agent that **switches naturally/freely** between ToD and chitchat

Research Questions #1

1. *Do the outlined cases adequately capture real-world scenarios?*

- U: 스타벅스에서 아메리카노 주문해줘 (메뉴 주문) (→ A: 주문이 완료되었습니다.) [ToD]
→ U: 스타벅스 새로운 메뉴 맛있다더라 → A: ~ 도넛 메뉴가 인기가 많다고 해요. → ... [chitchat]
→ A: 주문하신 매장으로 찾으러 갈까요? (switching) → U: 응(목적지 설정) [ToD]

2. *From a practical standpoint, does this approach make sense for designing dialogue management systems?*

3. *The solution framework we've proposed feel too **abstract**...*

Research Progress

Dataset Construction

Key Challenges

Solutions

Dataset Construction

• Trial-and-error (1/2)

1 각 전환 CASE를 모두 포함.

CASE 1 ToD → ToD	CASE 2 ToD → Chitchat
CASE 3 Chitchat → ToD	CASE 4 Chitchat → Chitchat

- CASE별 프롬프트 복잡성 증가
- 각 CASE별 생성된 대화는 자연스러움.
- 의도했던 CASE별 구분이 명확함(쉬움).

2 단순한 대화 생성 prompt.

2-1 매 턴마다 대화 모드 부여 (dynamic 전환 시도)

- 지나치게 부자연스러운 대화 흐름 전환
(e.g. 그런데, ...)

2-2 5-turn 단위 전환 시도

① 모델 자유도 ↑ ② 엄격한 지시

- ① 대화는 자연스럽게지만 추가로 연쇄적인 prompting 필요 + correction step 고려
- ② 부자연스러운 대화

3 CASE 간소화 (확장).

ToD → Chitchat → ToD

Chitchat → ToD → Chitchat

- ①에서의 4가지 CASE 포함되도록 데이터셋 설계
- 이상적인 Agent의 역할: 대화 전체 흐름을 인지하고 유지할 수 있는 대화 시스템 설계
- Multi-turn << Mode Switching
- 누가 전환을 시도하는가?

U: 스타벅스에서 아메리카노 주문해줘 (메뉴 주문) (→ A: 주문이 완료되었습니다.) [ToD]
 → U: 스타벅스 새로운 메뉴 맛있다더라 → A: ~ 도넛 메뉴가 인기가 많다고 해요. → ... [chitchat]
 → A: 주문하신 매장으로 찾으러 갈까요? (switching) → U: 응 (목적지 설정) [ToD]

Dataset Construction

• Trial-and-error (2/2)

Type	Seed data	Method	Result	Success rate
TCT	MultiWoZ 2.2	ToD 사이에 Chitchat 생성	<ul style="list-style-type: none"> 대화 모드 전환이 잘 되고, 전환 주체(C:user, T:agent) 지시를 잘 따름 <u>첫번째 ToD와 두번째 ToD의 Intent가 동일한 경우 성공적 Recover</u> C → T 전환 주체가 의도적으로 agent다 보니, user와 동일한 발화 반복 <p>TODO 실패 case에 대한 filtering 수행 (LangChain)</p>	샘플기준 약 80% 1.태그 정확한가 2.전환주체가 의도대로 반영되었는가 3.대화가 자연스러운가
	SLURP	SLURP에서 확장한 동일한 topic의 CASE1의 ToD 대화 두 개 사이에 Chitchat 생성	<ul style="list-style-type: none"> 확장 구축한 TCT를 다시 한 번 prompting하여 자연스러운 대화 생성 (전환 주체에 대한 지시를 따르도록 수정됨) <u>Recover를 위해 Intent space를 엄밀히 설계할 필요</u> <p>TODO Intent space 확정 후 재시도 TODO 후처리 (tag 정보) prompt 추가</p>	샘플기준 약 70% 1.태그 정확한가 2.전환주체가 의도대로 반영되었는가 3.대화가 자연스러운가
CTC	SLURP	SLURP에서 확장한 CASE2 대화에 선행하는 Chitchat 생성	<ul style="list-style-type: none"> <u>Chitchat스러운 intent space가 ToD같지 않은 대화 생성으로 이어짐</u> 이전에 만든 데이터(CASE2 or CASE3)로 recover flow를 포함하도록 확장하는 게 부적절 <p>TODO Intent space 확정 및 Prompt 재설계 TODO TCT와 같이 보정 Step 설계 필요</p>	-

ToD (or chitchat) only Multi-turn Dialogue ← Interrupt chitchat (or ToD)

Dataset Construction

• T → C → T Dialogue Generation (1/2)

Prompt

****Instruction****: ... Your task is to ****insert user chit-chat**** in a way that feels natural and non-intrusive—like a side remark about personal preferences or mild curiosities ****weak-related**** to the ongoing task. ...

You have to follow these **guidelines**: ...

1. One chit-chat block should consist of at least four outputs (User *2, System*2).
3. The system should ****acknowledge or briefly elaborate**** on at least one of these chit-chat turns, keeping the response light and relevant.
4. ****After**** the chit-chat, the system should ****use**** the newly revealed user preference(s) in the subsequent steps of the task-oriented dialogue—demonstrating that the side remark influenced the final recommendation or solution.

...

****FormatRequirements****

****Key Requirements for Reflection in ToD****

****Example****

****input**** \n[input_example]

****output**** \n[output_example]

<Dialogue>

← MultiWoZ 2.2 활용 (앞/뒤 2개 발화 제외)

2 [USER] [find_hotel] Don't care. But I'd like a place that's moderately priced and offers free WiFi.
 3 [SYSTEM] There is the Acorn Guesthouse in the moderate price range and 4 stars. It also has free Wi-Fi and parking.
 4 [USER] [chitchat] That sounds good! I always prefer places with parking, especially when I'm traveling with a group.
 5 [SYSTEM] Parking is definitely a great convenience, especially for group trips! It's good that the Acorn Guesthouse offers that.
 6 [USER] [chitchat] Yes, it's always nice not to worry about finding space for multiple cars. Plus, it'll be easier for our group to move around.
 7 [SYSTEM] [Transition to ToD] Exactly! It will definitely add to the convenience. Now, let's proceed with booking the Acorn Guesthouse for you.
 8 [USER] [find_hotel] Great, let's book that one. Please send me the address, postcode, and reference number.

Dataset Construction

• T → C → T Dialogue Generation (2/2)

Output

0 [USER] [find_hotel] Hi, I'm looking for a hotel in the center of Cambridge with free Wi-Fi.

1 [SYSTEM] Sure, I can help with that. Do you have a budget in mind?

2 [USER] [find_hotel] Not really, but something moderately priced would be good.

3 [USER] [chitchat] You know, I love staying in places with a cozy reading area. Do you know if any hotels like that are around?

4 [SYSTEM] Oh, that's a wonderful detail! Some boutique hotels do have reading lounges or libraries.

5 [USER] [chitchat] Yeah, it makes the stay so much more relaxing. I guess I just love the charm of quiet spaces.

6 [SYSTEM] I completely understand. I'll check if any hotels fit that vibe—it would really enhance your stay.

7 [SYSTEM] **[Transition to ToD]** Alright, let me find a moderately priced hotel in central Cambridge with free Wi-Fi. I'll prioritize places that also offer a cozy or quiet lounge for reading.

8 [USER] [find_hotel] Thank you, please let me know if you find a hotel.

9 [SYSTEM] I've found a few options! One that stands out is The Cambridge Stay. It's moderately priced, centrally located, and even has a small library for guests. Would you like more details or to proceed with booking?

Dataset Construction

• T → C → T Dialogue Samples (1/2) : SLURP

	Topic	Intent	Result
Good	email	email_query	<p>user: Hi, I was wondering if you could help me with something related to my email. agent: Of course! What do you need help with? user: Check my email for new emails during the last hour. agent: Sure, I can do that. Anything in particular you're expecting, or just a general check? user: Not really, just keeping an eye on things. agent: Got it, I'll check for any updates right away and let you know if anything stands out. user: By the way, how's the weather been around your area lately? agent: It's been pretty chilly, actually! A perfect time for a hot drink. How about you? user: Same here, I've been enjoying some tea. agent: Sounds cozy! Alright, let me check your emails. Would you like me to also look for any recent updates from specific senders, or just a general overview? user: Actually, could you check for any recent emails from Anna? agent: Absolutely! I'll take a look at those right away. Just a moment.</p>
Bad	weather	weather_query	<p>user: Hi there! Can you tell me what the weather is like today? agent: Of course! Do you have a specific location in mind for the weather update? user: What's the current weather? agent: Let me check that for you. Could you tell me your location? user: It's a bit chilly today, isn't it? I was hoping for a warmer day. agent: It does feel a bit cooler than usual. Hopefully, it'll warm up soon! user: Hi, can you tell me what the weather is like today? ← 대화 안에 새로운 session처럼 발화 진행 agent: Sure! Today is mostly sunny with a high of 45 degrees. user: When is it supposed to be above fifty again? agent: I can check that for you! Let me look up the forecast for the next few days.</p>

Dataset Construction

• T → C → T Dialogue Samples (2/2) : MultiWoZ 2.2

	Intent (Topic)	Result
Good	find_train (hotel) find_hotel (hotel)	<p>0 [USER] [find_train] I am looking for information in Cambridge.</p> <p>1 [SYSTEM] I can certainly help you with that! What do you want to know?</p> <p>2 [USER] [find_hotel] I'd like to stay at a guesthouse that offers free wifi.</p> <p>3 [SYSTEM] Do you have a price range or area preference? Because if not, how about A and B Guest House located in the east and moderately priced?</p> <p>...</p> <p>6 [USER] [chitchat] Yeah, I feel like hotels have a bit more to offer, especially when it comes to service and facilities.</p> <p>7 [SYSTEM] [Transition to ToD] I totally agree. Let me find a hotel that offers free Wi-Fi and fits your preferences.</p> <p>8 [USER] [find_hotel] The hotel should have a star of 4 and should include free parking.</p> <p>9 [SYSTEM] The Alexander Bed and Breakfast is a great place. Do you need a reservation?</p>
Bad	find_attraction (hotel) find_train (train)	<p>0 [USER] [find_hotel] I'm trying to find a hotel called Worth House. Can you help me?</p> <p>1 [SYSTEM] Sure. Worth House is located in the east at 152 Chesterton Road. Its phone number is 1223316074. Would you like me to help you book it?</p> <p>2 [USER] [book_hotel] Great! Yes, please book it for 8 people and for 3 nights.</p> <p>...</p> <p>7 [SYSTEM] Where will you be departing from and going to?</p> <p>8 [USER] [book_train] I'll be leaving Cambridge and going to Bishops Stortford. I want the train to leave after 10:45.</p> <p>9 [SYSTEM] There are 7 options. Shall I book the earliest one?</p> <p>10 [USER] [chitchat] I always prefer a train that gives me a bit of time in the morning, so I can relax a bit.</p> <p>11 [SYSTEM] That's a great approach! It's always nice to have a little extra time to ease into the day.</p> <p>12 [USER] [chitchat] Exactly! I feel like that extra time makes the trip much more enjoyable.</p> <p>13 [SYSTEM] [Transition to ToD] I completely agree! Let me proceed with booking the earliest train to make sure everything fits perfectly for you.</p> <p>...</p>

← User는 빠르지 않은 시간대의 기차를 원하지만, Agent는 (User의 의도와 다르게) 빠른 시간대 예매

Dataset Construction

• ToD → Chitchat → ToD

- MultiWoZ 2.2

- 전처리: Intent annotation 수정

1. 4-turn 이상 대화선별
2. 대화 사이에 적절한 chitchat 생성 요청
3. 검수

- SLURP

- 전처리: Intent 병합 / 이름 변경 + 인텐트 시나리오 설계

1. single utterance → multi-turn 확장
2. 1에서 확장한 대화 2개를 자연스럽게 연결
3. 2의 대화 사이에 적절한 chitchat 생성 요청
4. 검수

set_event related with:
recommend_events,
remove_event, check_weather,
send_email, play_music, ...

• Chitchat → ToD → Chitchat

- MultiWoZ 2.2

1. 4-turn 미만 대화선별
2. 대화 앞, 뒤에 적절한 chitchat 생성 요청
3. 검수

- SLURP

1. single utterance → multi-turn 확장
2. 1에서 확장한 대화 앞, 뒤에 적절한 chitchat 생성 요청
3. 검수

< 검수 기준 >

(Human evaluation 기반으로 검토 후 LangChain 활용 예정)

1. 대화가 자연스러운가?
2. 의도한 대로 전환이 진행되었는가? (TCT, CTC)
3. Speaker, intent, transition 태그가 정확히 붙었는가?

Research Question #2

Dataset construction process has been determined... but it heavily depended on **empirical practices**.

1. *What measures can be introduced to **standardize or formalize** this process?*
2. *What **alternative** methods or frameworks could ensure a more **objective and replicable decision-making process**?*

Key Challenges

(1) Detect Switching → (2) Find Recover Point → (3) Recover the Flow

1

Agent는 매 turn마다 대화 성격을 파악해야 한다.

- User가 intention이 있는 request를 했는가? (ToD)
- User가 intention 없이 대화를 이어 나가고자 하는가? (chitchat, LLM Chat, ...)

2

Agent는 대화 흐름이 변경되는 이력 정보를 유지해야 한다.

- Agent는 대화의 switching flow를 파악하고 있는가?
- 해당 정보를 어떻게 저장/관리해야 하는가?

3

Agent는 원래 대화 흐름으로 되돌아올 수 있어야 한다.

- Agent는 대화 흐름을 회복하기 위해 적절한 시점을 판별할 수 있는가?
- User가 회복한 대화 흐름에 적절히 대처할 수 있는가?

Solutions

(1) Detect Switching → (2) Find Recover Point → (3) Recover the Flow

1

Agent는 매 turn마다 대화 성격을 파악해야 한다.

- User가 intention이 있는 request를 했는가? (ToD)
- User가 intention 없이 대화를 이어 나가고자 하는가? (chitchat, LLM Chat, ...)

1. Pipeline style

- model 1 (switching detection)
→ model 2 (intent detection) → model 3 (response generation)

2. e2e style (on ToD-side)

- joint-learning: switching detection + intent detection + ?
- Generative model로 switching 처리하는 special token 학습

Solutions

(1) Detect Switching → (2) Find Recover Point → (3) Recover the Flow

2

Agent는 대화 흐름이 변경되는 이력 정보를 유지해야 한다.

- Agent는 대화의 switching flow를 파악하고 있는가?
- 해당 정보를 어떻게 저장/관리해야 하는가?

1. Detect Switching = [T → C]

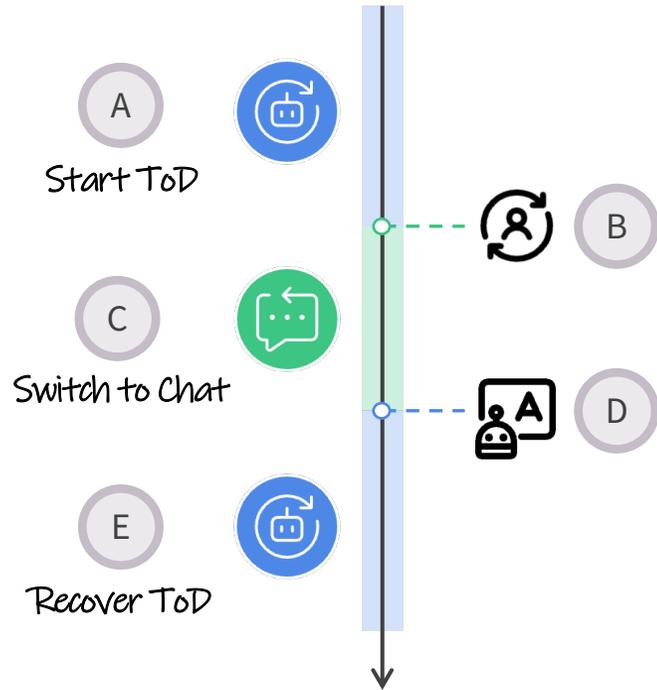
- + is it enough? (DST) = [o] → Need to Stay “C”
- + is it enough? (DST) = [x] → Need to Return ”T”

2. Detect Switching = [C → T]

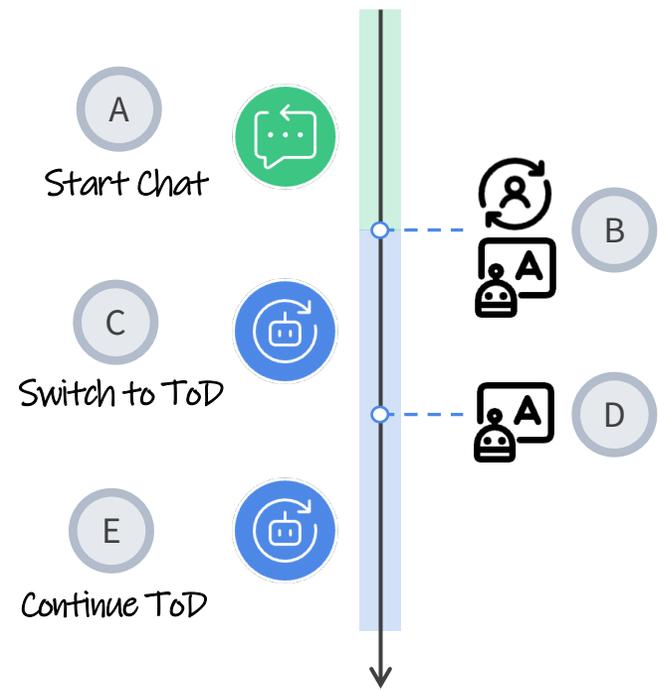
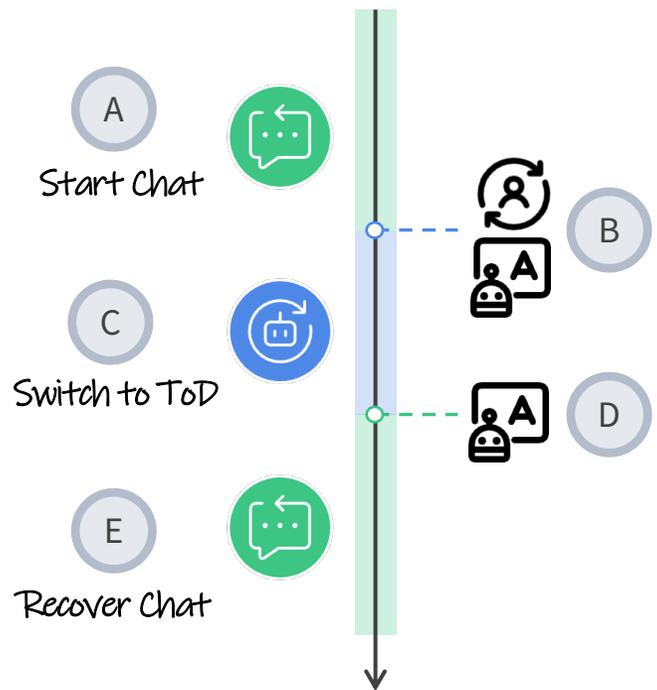
- + is it enough? (DST) = [o] → Need to Return “C”
- + is it enough? (DST) = [x] → Need to Stay “T” (until [o])

Evaluation (Metrics)

• ToD → Chitchat → ToD



• Chitchat → ToD → Chitchat



- A. (ToD) User Intent Detection
- B. (Switching) Policy Selection
- C. (Chitchat) Proper Chat Response
- D. (Switching) Suggest Intent
- E. (ToD) User Intent Detection

D = E & Overall Success

- A. (Chitchat) Proper Chat Response
- B. (Switching)
 - a. User : Policy Selection
 - b. Agent: Suggest Intent
- C. (ToD) User Intent Detection

B-b = C & Overall Success

- D. (Switching)
 - i. Address previous chat topic
 - ii. Anything else? (or suggest intent)
- E.
 - i. (Chitchat) Proper Chat Response
 - ii. (ToD) User Intent Detection

Questions & Discussion

Next Steps in the Research

Questions for Discussion

Progress Update

• Dataset Construction

Type	Seed data	Method	Result
TCT	MultiWoZ 2.2	<ol style="list-style-type: none"> 4-turn 이상 대화선별 대화 사이에 적절한 chitchat 생성 요청 	<p>TODO LangChain 검수 로직 도입</p> <p>TODO MultiWoZ 2.2 intent 오류 수정</p>
	SLURP	<p>Intent space 재정의</p> <ul style="list-style-type: none"> Intent 병합 이름 변경 시나리오 정의 <ol style="list-style-type: none"> single utterance → multi-turn 확장 1에서 확장한 대화 2개를 연결하는 자연스러운 대화 생성 2의 대화 사이에 적절한 chitchat 생성 요청 	<p>TODO LangChain 검수 로직 도입</p> <p>TODO 전수 데이터 생성</p>
CTC	MultiWoZ 2.2	<ol style="list-style-type: none"> 4-turn 미만 대화 선별 해당 대화의 전, 후에 대한 ToD 생성 요청 	<p>TODO LangChain 검수 로직 도입</p> <p>TODO MultiWoZ 2.2 intent 오류 수정</p>
	SLURP	<p>Intent space 재정의</p> <ul style="list-style-type: none"> Intent 병합 이름 변경 시나리오 정의 <ol style="list-style-type: none"> single utterance → multi-turn 확장 해당 대화의 전, 후에 대한 ToD 생성 요청 	<p>TODO sample 구축 및 검토</p>

• TODO

- 공통 검수 로직 도입
- MultiWoZ 2.4를 사용하거나, 수정 step 진행

[find_restaurant | find_hospital] Can I have the address, phone number, and type of food served by Restaurant Alimentum?

Progress Update

- **Baseline Experiment**

- e2e
 - SimpleToD : InterfereChat Baseline, SFT
 - ~~- Function Calling for DST : SIGDIAL 2024 SOTA, IntefereChat / FusedChat, fs ICL~~
 - ~~- FCToD : ACL 2024 Main SOTA~~
 - SalesAgent(?) : SFT
- pipeline
 - ToD (ToD-only trained e2e baseline) + LLM
 - LLM (ToD prompt ICL) + LLM

- **TODO**

- Baseline 구현 후 inference 확인

Research Questions #1

1. *Do the outlined cases adequately capture real-world scenarios?*

- U: 스타벅스에서 아메리카노 주문해줘 (메뉴 주문) (→ A: 주문이 완료되었습니다.) [ToD]
→ U: 스타벅스 새로운 메뉴 맛있다더라 → A: ~ 도넛 메뉴가 인기가 많다고 해요. → ... [chitchat]
→ A: 주문하신 매장으로 찾으러 갈까요? (switching) → U: 응(목적지 설정) [ToD]

2. *From a practical standpoint, does this approach make sense for designing dialogue management systems?*

3. *The solution framework we've proposed feel too **abstract**...*

Research Question #2

Dataset construction process has been determined... but it heavily depended on **empirical practices**.

1. *What measures can be introduced to **standardize or formalize** this process?*
2. *What **alternative** methods or frameworks could ensure a more **objective and replicable decision-making process**?*

Next Steps

- **Baseline 구축**

- **경험적 개선에 의존하고 있는 데이터 생성 prompting**

- *MultiWoZ 2.2*

1. 4-turn 이상 대화선별
2. 대화 사이에 적절한 chitchat 생성 요청
3. 검수

- *SLURP*

- Preprocess: Intent 병합 / 이름 변경 + **인텐트 시나리오 설계**

1. single utterance → multi-turn 확장
2. 1에서 확장한 대화 2개를 연결하는 자연스러운 대화 생성
3. 2의 대화 사이에 적절한 chitchat 생성 요청
4. 검수

Thank You

Yejin Yoon

HYU NLP Lab.
Hanyang University, South Korea

stillwithyou@hanyang.ac.kr