Hanyang Univ, NLP Lab.

BlendX

Complex Multi-Intent Detection
with Blended Patterns

2023/10/31 Yejin Yoon

CONTENTS

BlendX: Complex Multi-Intent Detection with Blended Patterns

Anonymous submission

Abstract

Task-Oriented Dialogue (TOD) systems typically suppose that a user utterance corresponds to a single intent. This assumption may be misaligned with real-world scenarios where users often express multiple intents simultaneously. While there is an emerging interest in Multi-Intent Detection (MID), existing in-domain datasets such as MixATIS and MixSNIPS have limitations in their formulation. To address these issues, we present BlendX, a suite of refined datasets featuring more diverse patterns than their predecessors, elevating both its complexity and difficulty. For dataset construction, we utilize both rule-based heuristics as well as a generative tool–OpenAI's ChatGPT—which is augmented with a similarity-driven strategy for utterance selection. To ensure the quality of the proposed datasets, we also introduce three novel metrics that assess statistical properties of an utterance related to word count, conjunction use, and pronoun usage. Extensive experiments on BlendX reveal that state-of-the-art MID models struggle with the challenges posed by the new datasets, highlighting the need to reexamine the current state of the MID field.

Keywords: Multi-Intent Detection, Task-Oriented Dialogue

1. Introduction

The successful implementation of Task-Oriented Dialogue (TOD) systems begins with the precise recognition of user intents. By accurately discerning the queries embedded in user inputs and routing them to the relevant components, the systems can adeptly respond, thereby effectively fulfilling user requests. In general, such systems are constructed on the assumption that each user utterance correlates exclusively with a single intent, which often diverges from practical scenarios.

Contrary to the conventional setting, the task of **Multi-Intent Detection (MID)** presents a more nuanced and comprehensive challenge for TOD
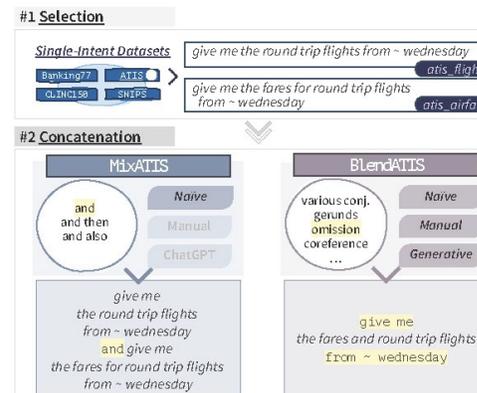
Introduction

Problem States
Previous works' naivety
Ex-/Implicit concatenation

Definition of Multi-intent Detection

- **Input instance**에 의도가 여러 개 할당될 수 있는 **classification task**

1 utterance   multi-intent

- Input: 1 utterance
- Output: 1 or more label(s)

Intent 1
Intent 2
Intent 3
Intent 4
Intent 5
Intent 6

Sigmoid

[CLS] T_1 T_2 ... T_N

Pretrained Language Model (**PLM**)
(e.g. BERT, RoBERTa, ELECTRA, …)

[CLS] Tok_1 Tok_2 ... Tok_N

Tokenizer

에어컨 켜고 집으로 목적지 설정해줘. (Raw Utterance)
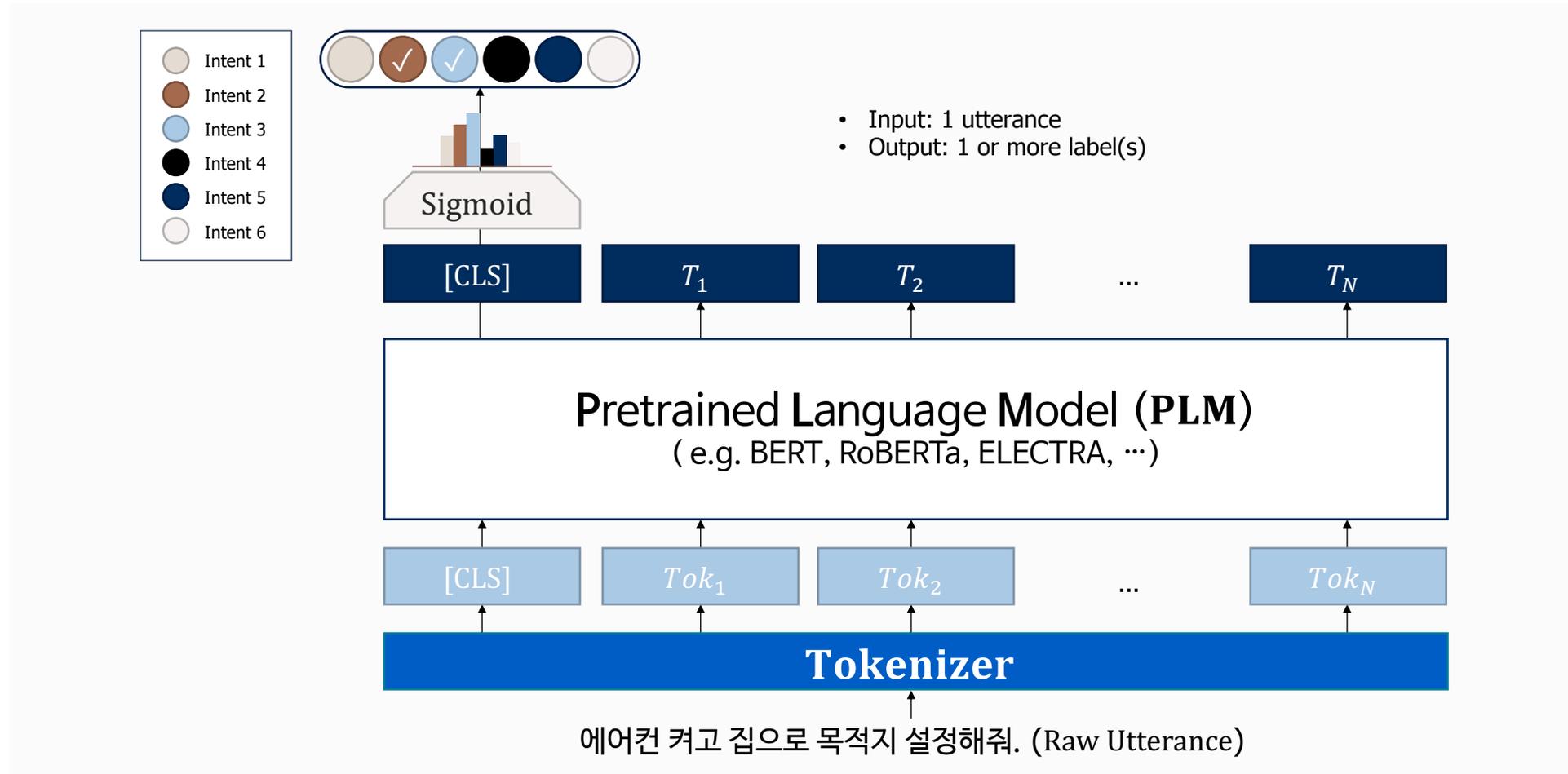
The Critical Role of MID in Task-Oriented Dialog System

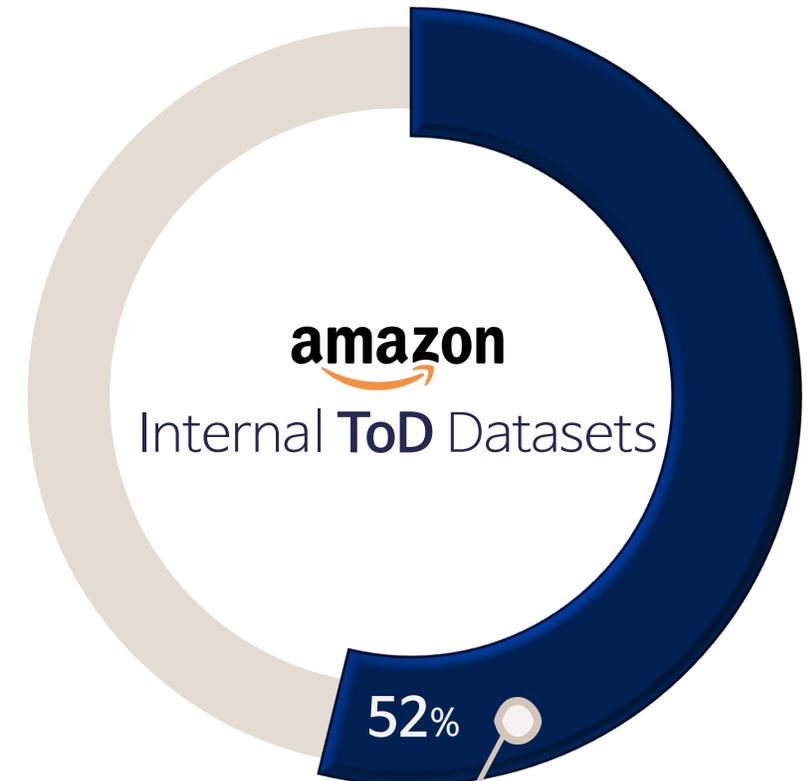According to a 2019 paper published by AWS AI, Amazon,

52%
more than half of its internal data utterances had multiple intentions.

amazon
Internal **ToD** Datasets

52%

Rashmi Gangadharaiah and Balakrishnan Narayanaswamy. *Joint Multiple Intent Detection and Slot Labeling for Goal-Oriented Dialog*. NAACL. 2019

Naivety of MixSNIPS / ATIS

Single-intent utterances

⌄

1 Multi-intent utterance

ATIS

give me the round trip flights
from cleveland to miami next wednesday

atis_flight

+

give me the fares for round trip flights
from cleveland to miami next wednesday

atis_airfare

⌄

MixATIS

give me the round trip flights
from cleveland to miami next wednesday
and give me the fares for round trip flights
from cleveland to miami next wednesday

Libo Qin, Xiao Xu et al. AGIF: An Adaptive Graph-Interactive Framework for Joint Multiple Intent Detection and Slot Filling. EMNLP 2020, Findings.

HYU 한양대학교 HANYANG UNIVERSITY

Naivety of MixSNIPS / ATIS

Single-intent utterances

⌄

1 Multi-intent utterance

ATIS

give me the round trip flights
from cleveland to miami next wednesday   atis_flight

+

give me the fares for round trip flights
from cleveland to miami next wednesday   atis_airfare

MixATIS

👎 give me the round trip flights
from cleveland to miami next wednesday
and give me the...
from clevelan...

👍 give me the fares and round trip flights
from cleveland to miami next wednesday

Libo Qin, Xiao Xu et al. AGIF: An Adaptive Graph-Interactive Framework for Joint Multiple Intent Detection and Slot Filling. EMNLP 2020, Findings.

HYU 한양대학교 HANYANG UNIVERSITY

Discussion #1

Data Construction
Concatenation
Selection
3-Metric

Multi-intent set up
on construction-side

shared 05/24/2023

(* Concatenation side)

Naïve Approach

- 상관성을 고려하지 않은 **대등관계** 문장
- **Template-based** Concatenation
  - **접속사** 활용: and then, and, also, next, for, yet, however, nevertheless, after, because, than, where, while, if, unless, first of all, finally, I also would like to know, additionally, …
- Sentence Selection
  : Simple Random Sampling
- Intent 조합
  : 의미적 연관성 낮음
  : 동시 발생 가능성도 낮음

  * MixSNIPS, MixATIS setting

---

e.g. A음악 켜고 B영화 켜줘
  1) A음악 켜줘 (#music.play:A음악)
  2) B영화 켜줘 (#movie.play:B영화)

Soft

Complex Approach

- 보다 상관성 있는 **대등관계** 문장
- **Template-based** Concatenation
  - **접속사** 활용: (좌동)
- Sentence Selection
  : **Utterance Representation** 활용
  1) Encode Utterance: Vector space에 intent별로 clustering되도록 utterance embedding 학습
  2) 서로 다른 intent이지만 utterance간 거리가 가까운 2개 utterance 선택
  3) Template-based merge → multi-intent utterance 생성
- Intent 조합
  : 의미적 연관성 높음
  : 동시 발생 가능성 모호

---

e.g. A음악 켜고
  A음악 플레이리스트에 넣어줘
  1) A음악 켜줘 (#music.play:A음악)
  2) A음악 플레이리스트에 넣어줘
    (#music.add_playlist:A음악)

Hard

Complex Approach

- 상관성을 고려한 **종속관계** 문장
- **생략, 치환, 상호참조, 순서** 등 활용
  - 동일한 주어/동사 생략
  - 대명사, 지시어 등 사용
  - 순서가 있는 intent 조합
  - …
- Sentence Selection
  : 연구적 고민 필요
- Intent 조합
  : 의미적 연관성(종속성) 높음
  : 동시 발생 가능성도 높음

---

e.g. A음악 켜고
  플레이리스트에 넣어줘
  1) A음악 켜줘 (#music.play:A음악)
  2) ~~A음악~~ 플레이리스트에 넣어줘
    (#music.add_playlist:A음악)

Approach of Multi-intent Concatenation

- Complexity side
- Methodology side

- **Complexity side**
  - Explicit Concatenation: use connectors during concatenation
  - Implicit Concatenation: do NOT use connectors during concatenation

- **Methodology side**
  - Manual Concatenation: rule-based concatenation approach
  - Generative Concatenation: concatenation by using generative language model

Complexity side

Explicit Concatenation
- Only use connectors to concatenate utterances
  - and, and then, and also
  - before, additionally, …

Implicit Concatenation
- Concatenate utterances without connectors
  - Inherent ambiguity
  - Omissions
  - Coreferences
  - Gerund phrase

- AND variants
- Various Conjunctions
- Inherent Ambiguity
- Gerund Phrase
- Omission
- Coreference

Methodology side

Naïve Approach
- Approach employed by MixX

Manual Approach
- Rule-based concatenation technique

Generative Approach
- Concatenate utterances by using generative language model, ChatGPT
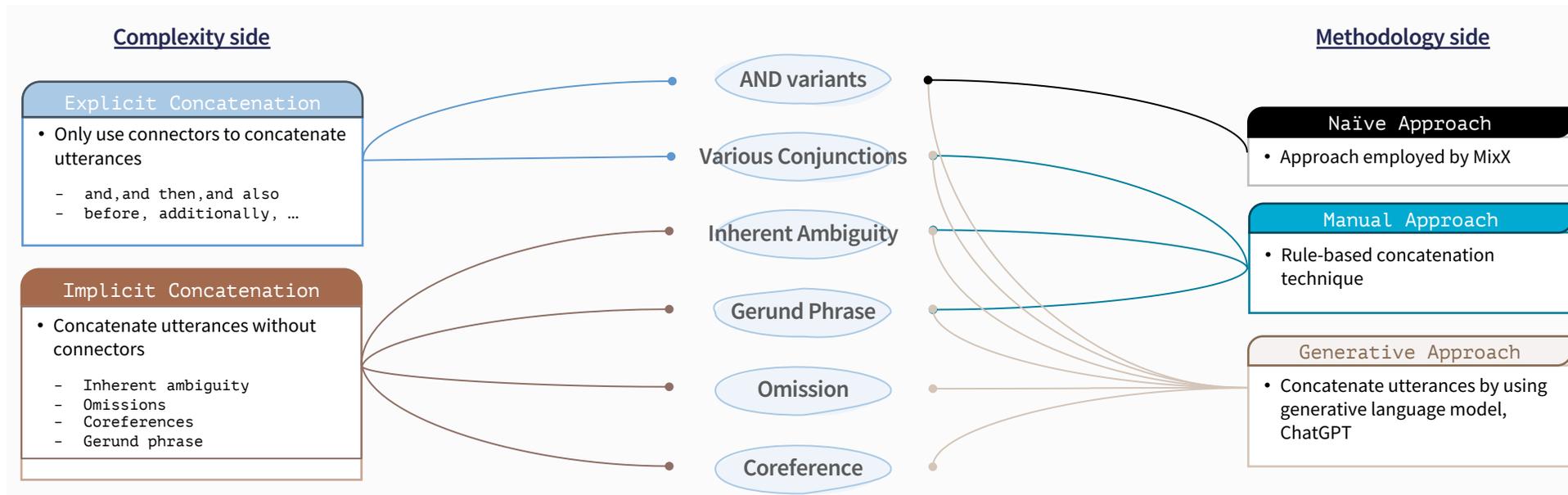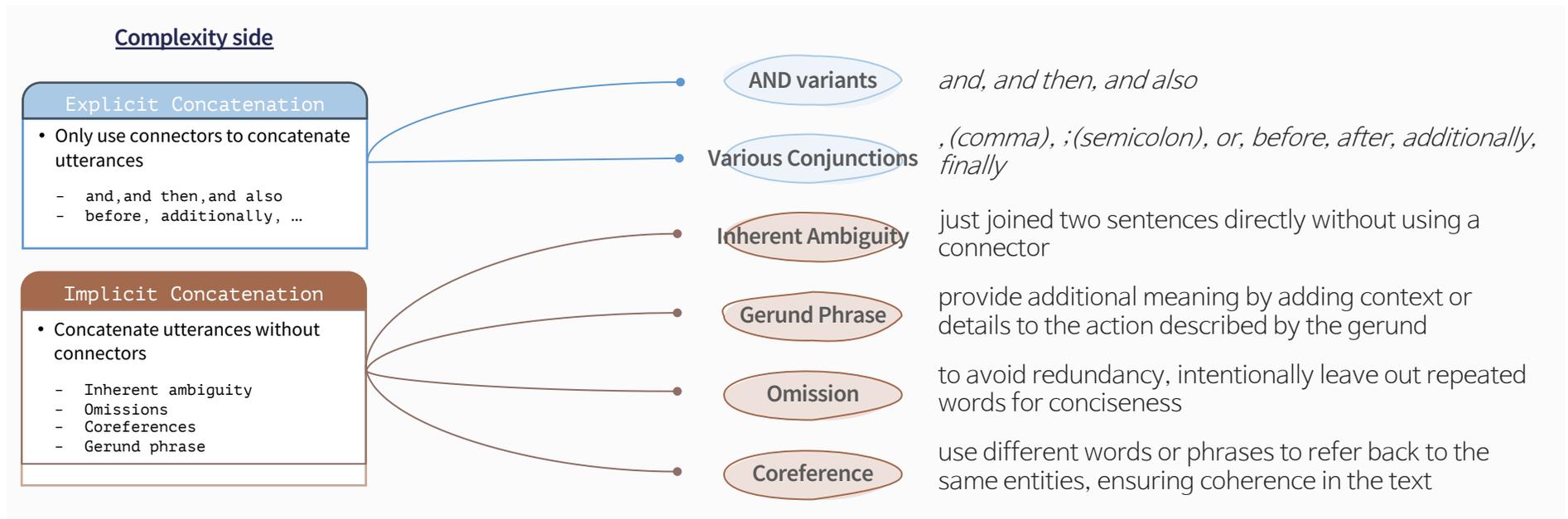
Approach of Multi-intent Concatenation

- Complexity side
- Methodology side

- **Complexity side**
  - Explicit Concatenation: use connectors during concatenation
    - → AND variants / Various Conjunctions
  - Implicit Concatenation: do NOT use connectors during concatenation
    - → Inherent Ambiguity / Gerund Phrase / Omission / Coreference

Complexity side

Explicit Concatenation
- Only use connectors to concatenate utterances
  - and, and then, and also
  - before, additionally, …

Implicit Concatenation
- Concatenate utterances without connectors
  - Inherent ambiguity
  - Omissions
  - Coreferences
  - Gerund phrase

AND variants — *and, and then, and also*

Various Conjunctions — *, (comma), ; (semicolon), or, before, after, additionally, finally*

Inherent Ambiguity — just joined two sentences directly without using a connector

Gerund Phrase — provide additional meaning by adding context or details to the action described by the gerund

Omission — to avoid redundancy, intentionally leave out repeated words for conciseness

Coreference — use different words or phrases to refer back to the same entities, ensuring coherence in the text
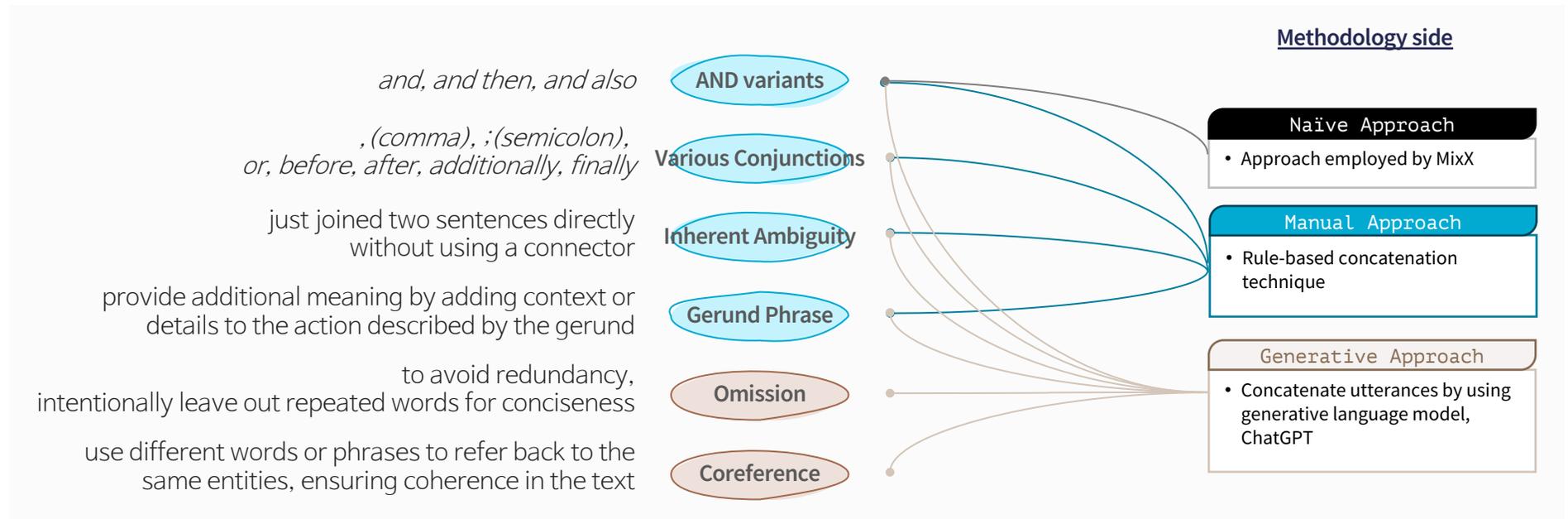
Approach of Multi-intent Concatenation

- Complexity side
- Methodology side

- **Methodology side**

  - Manual Concatenation: rule-based concatenation approach
    → AND variants / Various Conjunctions / Inherent Ambiguity / Gerund Phrase
  - Generative Concatenation: concatenation by using generative language model
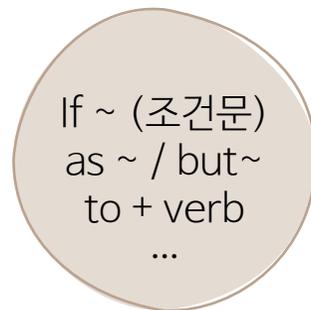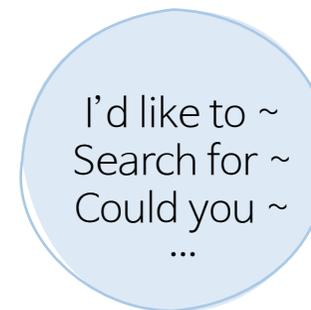   → Omission / Coreference

13

Data Construction
- Concatenation

- Complexity side

- Methodology side
  - Manual approach
  - Generative approach

- **Example for each concatenation approach**

	utt1	utt2	intent1	intent2	concatenation
Naïve	i want to put this song in my new boots playlist	what films are going to be playing at harkins theatres at zero a m	AddToPlaylist	SearchScreening Event	i want to put this song in my new boots playlist and what films are going to be playing at harkins theatres at zero am
Manual	please show me all airports in denver	can you list costs of denver rental cars	atis_airport	atis_ground_fare	please show me all airports in denver listing costs of denver rental cars
Generative	play some theme songs from 1974	play the movie white christmas	PlayMusic	SearchCreative Work	play some theme songs from 1974 and the movie white Christmas
	clear my to do list	repeat my to do list	todo_list_update	todo_list	i need to clear my to-do list and then repeat it

- Connectors excluded from Manual Concatenation

If ~ (조건문)
as ~ / but~
to + verb
…

→ intent 훼손 가능성

I'd like to ~
Search for ~
Could you ~
…

→ paraphrasing

14

Data Construction
- Concatenation

- Complexity side
- Methodology side
  - Manual approach
  - Generative approach

- **ChatGPT concatenation prompt**

```
You are a native English speaker.
[Task Definition] Combine 2 or 3 sentences as one single sentence.
[Goal] The focus is on creating a single sentence that captures the essence of both ideas without unnecessary
redundancy.
[Instructions] - Avoid adding just punctuation.
               - Don't paraphrase.
               - Don't compromise the meaning of each sentence.
               - Don't capitalize all characters.
               - Don't replace numbers with radix.
               - Maintain the intent of each sentence.
               - Don't forget that if a sentence starts with a verb, it's a statement.
               - Do NOT use conjunctions like 'and'
               - Don't print '[Good Answer]'
               - Don't print intent directly.
[Example 1]
        My dog is playful (dog's feature) + My dog loves chasing balls (dog's feature)
         [Good Answer] My playful dog loves chasing balls
         [Bad Answer] My dog is playful, and my dog loves chasing balls
         [Bad Answer] My dog is playful, and also loves chasing balls.
[Example 2]
        They finished the project(project done) + They had time(taking time)
         [Good Answer] With time on their hands, they finished the project
         [Bad Answer] They finished the project, and they had time
         [Bad Answer] They finished the project and had time
[Example 3]
        She answered the phone (answering)) + She was making dinner (cooking)
         [Good Answer] While answering the phone, she was making dinner
         [Bad Answer] She answered the phone, and She was making dinner
Combine the following sentences naturally. Inside the parentheses is the intent of each sentence. :
{utt1} (intent: {intents[0]})  + {utt2} (intent: {intents[1]})
```

제시했던
명시적인 제약조건을
따르지 않는 결과 반환

Data Construction
- Concatenation

- Complexity side

- Methodology side
  - Manual approach
  - Generative approach

- **Failure of Concatenation using ChatGPT**

failed

succeeded

Utterance 1	Utterance 2	Concatenation	Implication
weather for frankfort (`GetWeather`)	3 out of 6 for the last album (`RateBook`)	frankfort's weather gets a 3 out of 6 rating for the last album	Intent removed
what day of the week do flights from nashville to tacoma fly on (`atis_day_name`)	flight number from houston to dallas (`atis_flight_no`)	flights from nashville to tacoma fly on what day of the week and what is the flight number from houston to dallas	Intent changed
what is mci (`atis_abbreviation`)	list la (`atis_city`)	combine the sentences: "what is mci?" and "list la".	Failed to merge
play the playlist funtime activity (`AddToPlaylist`)	add adele onto funtime activity playlist (`PlayMusic`)	play the playlist, adding adele to the funtime activity playlist	Gerund phrase
how do i freeze my bank account (`account_blocked`)	do you know why my bank account is frozen (`freeze_account`)	how do i freeze my bank account, and do you know why it is frozen	Coreference

Table 1: Failed (Top) and succeeded (Bottom) results of the **Generative Approach** and their implications.

- Over 90% of the 100 randomly sampled instances were connected by 'and'
- Failed to maintain intent in at least 15 instances
- The cherry-picked example was, brilliantly, implicitly concatenated, which is what we intended

HYU 한양대학교 HANYANG UNIVERSITY

Overview of MID Dataset Construction

Single-Intent Datasets

- Banking77
- ATIS
- CLINC150
- SNIPS

Preprocessing

- Lowercase
- Remove punctuation

① For Manual Approach

- AND variants
- Various Conjunctions
- Inherent Ambiguity
- Gerund Phrase

Naïve Approach
- Approach employed by MixX

Manual Approach
- Rule-based concatenation technique

② For Generative Approach

Utterance Selection

θ

Utterance Concatenation

- AND variants
- Various Conjunctions
- Inherent Ambiguity
- Gerund Phrase
- Omission
- Coreference

ChatGPT

Generative Approach
- Concatenate utterances by using generative language model, ChatGPT

Filtering Concatenated Utterances

$$W(utt, n) = 1$$

Wrong prediction of TFMN

Filtering by Human

Concatenation

HYU 한양대학교 HANYANG UNIVERSITY
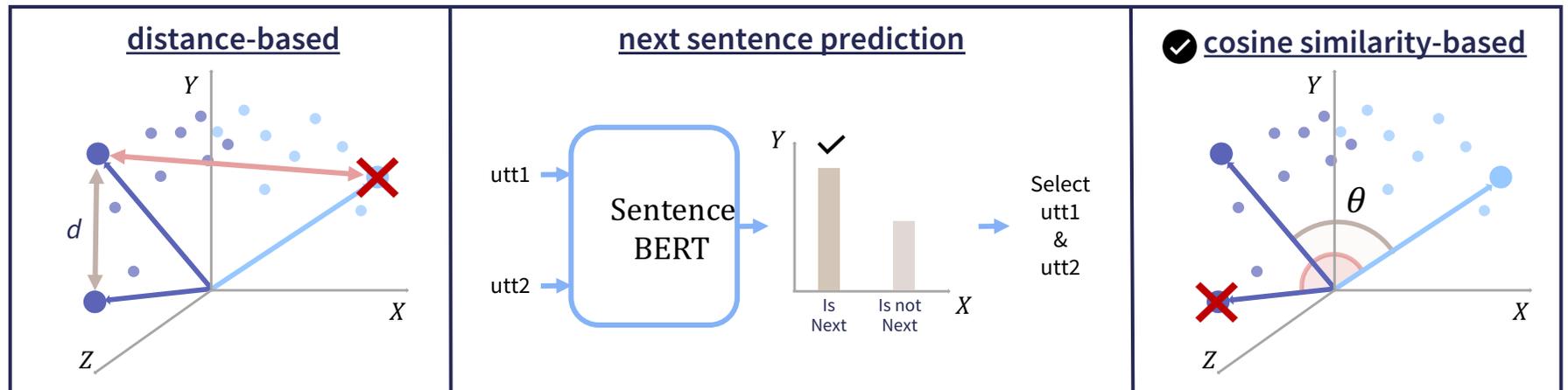
Data Construction
- Selection

- **Process**
    1. SentenceBERT를 이용해 각 single-intent utterance마다 임베딩을 생성
    2. 임베딩 간 similarity가 높다고 판단된 utterance들을 concatenate 대상으로 선정
       * 선택되는 발화들은 서로 다른 의도를 가짐.

- **Selection approach**
    - distance-based : 임베딩 간의 거리가 가까운 발화 선택 (L-2 distance)
    - next sentence prediction : 주어진 쌍의 발화가 이어지는지를 SentenceBERT가 이진분류
    - cosine similarity-based : 임베딩 간의 코사인 유사도가 큰 발화를 선택

- **3-Metric**

  - utt: concatenated utterance with 2 or more intents
  - n: concatenation에 사용한 single-intent utterance 개수

W(utt, n)
Word count

$$\mathrm{W}(utt, n) \overset{\text{def}}{=} \mathbf{1}_{\mathbb{Z}-\mathbb{N}}\left(|utt|_{word} - \sum_{i=1}^{n} |utt_i|_{word}\right).$$

- 발화 연결 전후의 <mark>단어 개수</mark> 차이가 0 이하인지

  (단어 개수가 감소하였는지 판단)

C(utt, n)
Conjunction

$$\mathrm{C}(utt, n) \overset{\text{def}}{=} \mathbf{1}_{\mathbb{Z}-\mathbb{N}}\left(|utt|_{conj} - \sum_{i=1}^{n} |utt_i|_{conj}\right).$$

- 발화 연결 전후의 <mark>접속사 개수</mark> 차이가 0 이하인지

  (접속사를 사용하지 않거나 되려 감소하였는지 판단)

- 정의된 접속사: and, or, before, after, additionally, finally, ','(쉼표), ';'(세미콜론)

P(utt, n)
Pronoun

$$\mathrm{P}(utt, n) \overset{\text{def}}{=} \mathbf{1}_{\mathbb{N}}\left(|utt|_{pron} - \sum_{i=1}^{n} |utt_i|_{pron}\right).$$

- 발화 연결 전후의 <mark>대명사 개수</mark> 차이가 1 이상인지

  (대명사를 사용했는지 판단)

- 정의된 대명사: it, them, their, theirs, this, that, those, these

> An implicitly concatenated utterance is likely to receive 1 in the metric evaluation.

HYU 한양대학교 HANYANG UNIVERSITY

- **Examples**

Utterance 1	play my 88 keys playlist (`PlayMusic`)
Utterance 2	add another song to my 88 keys playlist (`AddToPlaylist`)

Strategies	Concatenation Results	$W(utt, 2)$	$C(utt, 2)$	$P(utt, 2)$
Explicit Concatenation	play my 88 keys playlist **and** also add another song to my 88 keys playlist	0	0	0
Implicit Concatenation				
Inherent Ambiguity	play my 88 keys playlist add another song to my 88 keys playlist	1	1	0
Omissions	play my 88 keys playlist and add another song	1	0	0
Coreferences	play my 88 keys playlist and add another song to it	1	0	1
Gerund Phrase	add another song to my 88 keys playlist playing it	1	1	1

Table 3: Various concatenation classes, accompanied by their examples and respective metric values.

	concatenation	utt1	utt2	Difference	Metric
	add another song to my 88 keys playlist playing it	play my 88 keys playlist	add another song to my 88 keys playlist		
Words	10	5	8	10 − (5 + 8) = −3	W(· , 2) = 1
Conjunctions	0	0	0	0 − (0 + 0) = 0	C(· , 2) = 1
Pronouns	1	0	0	1 − (0 + 0) = 1	P(· , 2) = 1
	i need to clear my to-do list and then repeat it	clear my to do list	repeat my to do list		
Words	11	5	5	11 − (5 + 5) = 1	W(· , 2) = 0
Conjunctions	1	0	0	1 − (0 + 0) = 1	C(· , 2) = 0
Pronouns	1	0	0	1 − (0 + 0) = 1	P(· , 2) = 1

Data Evaluation - 3-Metric

- **Analysis of each concatenation approach**

Metric	SNIPS			ATIS			Banking77			CLINC150		
	Naïve	Manual	Generative	Naïve	Manual	Generative	Naïve	Manual	Generative	Naïve	Manual	Generative
$W(utt, 2)(\uparrow)$	0%	**37%**	29%	0%	**36%**	18%	0%	**46%**	37%	0%	**48%**	28%
$C(utt, 2)(\uparrow)$	0%	**56%**	10%	0%	**52%**	15%	0%	**50%**	27%	0%	**56%**	32%
$P(utt, 2)(\uparrow)$	0%	0%	**7%**	0%	0%	**8%**	0%	0%	**13%**	0%	0%	**6%**

Table 4: Comparative analysis of the three concatenation approaches: Naïve, Manual, and Generative. Notably, the Manual method demonstrates pronounced efficiency in reducing utterance length.

- Manual and Generative Approach만이 implicit concatenation이 가능하다.
  - 모든 Naïve Approach는 발화 길이 단축이나 대명사 사용이 존재하지 않고, 연결어도 반드시 사용한다. □
  - 특히 Manual Approach는 ChatGPT Concatenation보다 최대 2배까지 더 발화 길이를 줄이고 conjunction 사용 빈도도 낮춘다.
  - Coreference 발생을 추정할 수 있는 P metric은 generative approach에서만 발생된다. □

- **Analysis of each concatenation approach**

Metric	SNIPS			ATIS			Banking77			CLINC150		
	Naïve	Manual	Generative	Naïve	Manual	Generative	Naïve	Manual	Generative	Naïve	Manual	Generative
$W(utt, 2)(\uparrow)$	0%	**37%**	29%	0%	**36%**	18%	0%	**46%**	37%	0%	**48%**	28%
$C(utt, 2)(\uparrow)$	0%	**56%**	10%	0%	**52%**	15%	0%	**50%**	27%	0%	**56%**	32%
$P(utt, 2)(\uparrow)$	0%	0%	**7%**	0%	0%	**8%**	0%	0%	**13%**	0%	0%	**6%**

Table 4: Comparative analysis of the three concatenation approaches: Naïve, Manual, and Generative. Notably, the Manual method demonstrates pronounced efficiency in reducing utterance length.

- Manual and Generative Approach만이 implicit concatenation이 가능하다.
  - 모든 Naïve Approach는 발화 길이 단축이나 대명사 사용이 존재하지 않고, 연결어도 반드시 사용한다.
  - 특히 Manual Approach는 ChatGPT Concatenation보다 최대 2배까지 더 발화 길이를 줄이고 conjunction 사용 빈도도 낮춘다.
  - Coreference 발생을 추정할 수 있는 P metric은 generative approach에서만 발생된다.
- ChatGPT는 Connector 사용 빈도가 높다.
  - 최소 68%에서 최대 90%까지 'and' 사용하는 경향 확인

• **Analysis of ChatGPT concatenation with utterance selection**

ChatGPT concatenation이
실패한 비율

Metric	SNIPS		ATIS		Banking77		CLINC150	
	Random	Sim.	Random	Sim.	Random	Sim.	Random	Sim.
Cosine sim.	0.105	0.746	0.214	0.758	0.212	0.748	0.093	0.749
Error rate ($\downarrow$)	16%	**14%**	41%	**10%**	22%	**9%**	19%	**13%**
$W(utt, 2)(\uparrow)$	27.38%	**44.87%**	10.17%	**27.78%**	**34.62%**	30.77%	30.86%	**31.03%**
$C(utt, 2)(\uparrow)$	**8.33%**	1.28%	3.39%	**4.44%**	**28.21%**	15.38%	**25.93%**	3.45%
$P(utt, 2)(\uparrow)$	3.57%	**10.26%**	1.69%	**12.22%**	10.26%	**20.88%**	3.70%	**14.94%**

Table 2: Comparison of Random and Similarity-Based (Sim.) utterance selection across datasets when applied to ChatGPT. We find that Sim. leads to a reduced error rate in ChatGPT's data generation.

- Random selection의 경우보다 similarity-based selection한 경우 error rate가 줄어든다. ☐
  - 최소 2%에서 최대 31%까지 감소 확인
- Pronoun 사용은 늘고, 대부분의 경우 word count가 감소한다. ☐
- Conjunction 사용의 경우, 대부분 similarity-based selection에서 보다 증가한다. ☐
  - Implicit concatenation, 특히 omission이나 coreference 발생 빈도가 증가함에 따라, 의미적 명확성을 보장하기 위해 자연스럽게 단순 연결어 'and' variants의 사용 증가

Release Data

- **BlendX: Complex multi-intent detection with blended patterns**

Dataset	# of intents	Training	Dev	Test	Total
BlendSNIPS	7	50,625	2,613	2,615	55,853
BlendATIS	18	20,250	1,125	1,125	22,500
BlendBanking77	77	36,390	2,009	2,021	40,420
BlendCLINC150	147	54,899	2,889	2,977	60,765

Σ (total) = 179,583

Table 5: Statistics of the constituents of BlendX.

- Source Dataset: SNIPS, ATIS, Banking77, CLINC150
- Concatenation approach: Naïve, Manual, Generative
  - Ratio: (Naïve : Manual : Generative) = (1 : 3 : 0.5)
- Selection
  - Random selection for Manual Approach
  - Cosine Similarity-based selection for Generative Approach
- Train-Dev-Test Ratio: (90 : 5 : 5)

Discussion #2

Experiment & Analysis
Custom metric
Baseline evaluation
Visualization

- **Measure 3-metric**

Metric	MixX		BlendSNIPS		BlendATIS		BlendBanking77		BlendCLINC150	
$W(\cdot)$	0	(0%)	10,420	(25.51%)	4,465	(27.06%)	11,547	(38.96%)	14,865	(33.22%)
$C(\cdot)$	0	(0%)	12,649	(30.96%)	5,124	(31.05%)	9,712	(32.77%)	14,522	(32.46%)
$P(\cdot)$	0	(0%)	329	(0.81%)	202	(1.22%)	1,052	(3.55%)	1,377	(3.08%)

Table 6: Comparing the metric of 2- or 3- intent concatenated utterances in MixX and BlendX. Specify that $W(\cdot) = W(\cdot, 2) + W(\cdot, 3)$, and the same applies to $C(\cdot)$ and $P(\cdot)$.

- Original MixX에 비해 Implicit Concatenation 증가
  - 단어 개수(W)와 접속사(C)의 개수는 줄어들지만 대명사(P)의 수는 증가

- **Evaluate Baseline**
  - 3-baseline: implemented w/o slot-filling part
    - TFMN: predict # of intents k, and then top-k intents over the probability distribution
    - SLIM: threshold-based classification model using sigmoid function
    - ChatGPT: OpenAI's generative model

Model	Option		Dataset (Metric: accuracy)			
	Training	Test	SNIPS	ATIS	Banking77	CLINC150
TFMN	MixX	MixX	95.96*	76.80*	76.11	85.60
	MixX	BlendX	51.01	50.40	36.96	46.15
	BlendX	BlendX	92.96	76.00	62.69	78.06
SLIM	MixX	MixX	95.88*	91.48*	0.06	86.85
	MixX	BlendX	92.96	64.09	0.06	74.47
	BlendX	BlendX	95.72	77.33	0.10	84.44
gpt-3.5-turbo	-	MixX	77.56	33.60	23.72	45.55
	-	BlendX	73.23	29.96	22.76	40.98

Table 6: Evaluating 3 MID models based on their training and test set sources: MixX vs. BlendX. *: these numbers are from our re-implementation, ignoring the impact of joint learning with slot filling modules.

  - For supervised SOTA models, we consistently observe a performance drop on our **BlendX** datasets with explicit as well as implicit concatenations.

27

- **ChatGPT Inference prompt**

명시적인 제약조건
(최대 intent 3개,
answer format 등)을
따르지 않는 결과 반환
→ 후처리 후 성능측정

```
You are an Intent Detection Model on single utterance.

[Task Definition]
        Detect single or more intent(s) of each utterance, but you can only classify UP TO 3
        most plausible intents on 1 utterance.
[Labels] atis_airport, atis_ground_service, atis_abbreviation, atis_city, atis_aircraft,
        atis_ground_fare, atis_flight, atis_airfare, atis_meal, atis_distance, atis_cheapest,
        atis_capacity, atis_restriction, atis_quantity, atis_airline, atis_flight_no,
        atis_flight_time, atis_day_name
[Answer format]
        If more than one, concatenate with '#', such as {Label}#{Label}.
            e.g. atis_ground_fare#atis_distance

[Example 1]
        [Utterance] what is restriction ap80
        [Answer] atis_restriction
[Example 2]
        [Utterance] what does the fare code qx mean , what is the distance between Pittsburgh
                    airport and downtown pittsburgh and what is restriction ap80
        [Answer] atis_abbreviation#atis_distance#atis_restriction

Detect a single or up to 3 intent(s) on this following utterance: utt
```
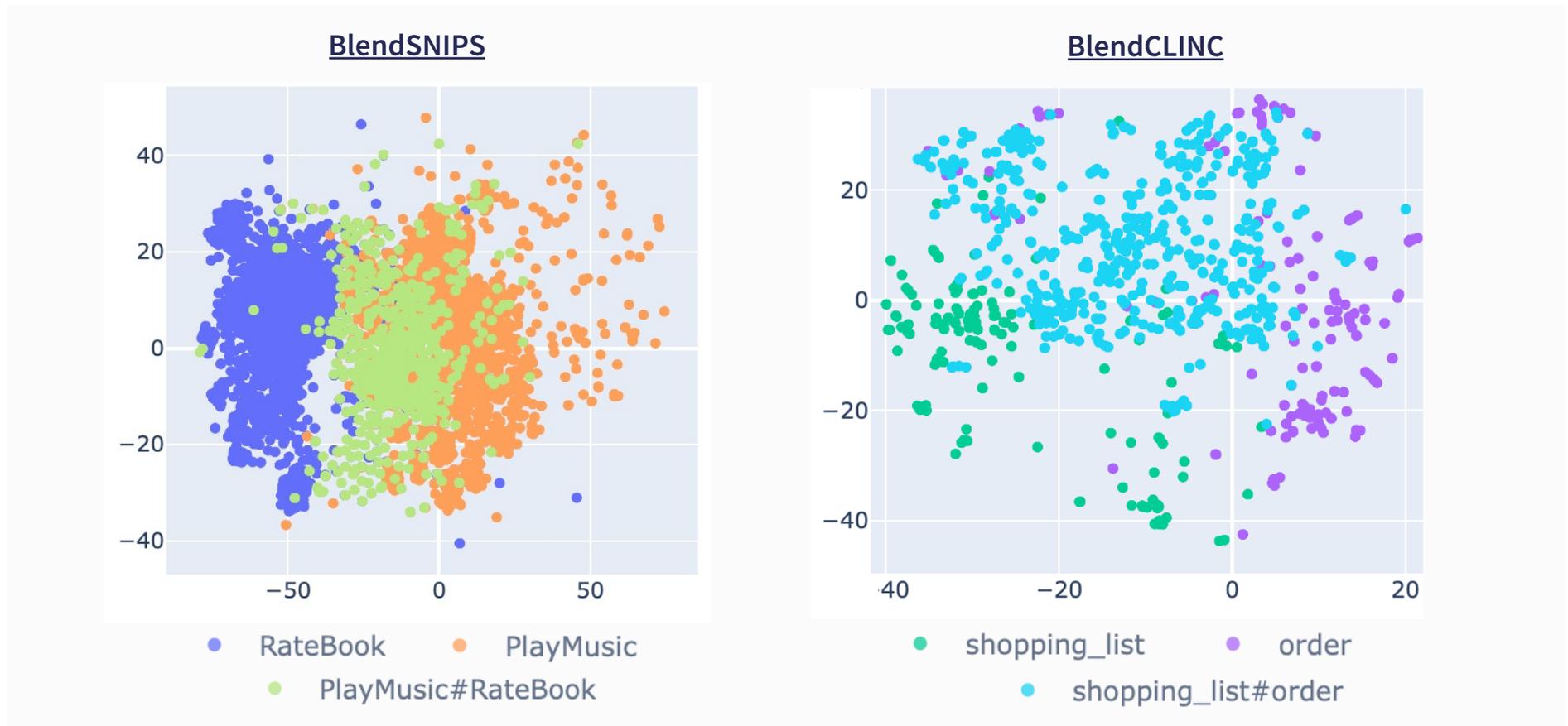
- **Visualization of X and BlendX utterances on 2-dimensional space**
  - BlendX의 concatenated utterance가 두 source utterance의 의미를 보존하면서, multi-intent를 갖는 결과 확인

BlendSNIPS

BlendCLINC

- **Ablation study for concatenation methods**

Approach	Dataset (Metric: accuracy)			
	SNIPS	ATIS	Banking77	CLINC150
Naïve-Only	95.32	73.23	62.30	80.73
Manual-Only	25.32	42.40	8.05	25.73
Generative-Only	81.58	53.93	27.95	60.17

Table 7: Experiments results for 2-intent utterances concatenated using each approach.

- Manual-Only, Generative-Only 방식이 Naïve-Only 방식보다 성능 하락함의 주요 원인
  - Manual concatenation은 explicit하고 implicit한 utterance 방식 모두를 잘 생성한다.
  - 우리가 정의한 implicit concatenated utterance(omission과 coreference의 발생)은 Generative concatenation에서만 생성될 수 있지만, 그 비율이 크지 않다.

Conclusion

Contribution
Limitation
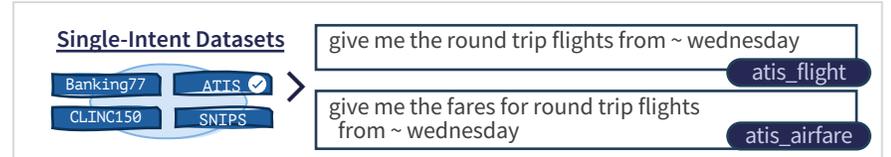Future work

Contribution

- **기존의 multi-intent dataset의 한계 확인**
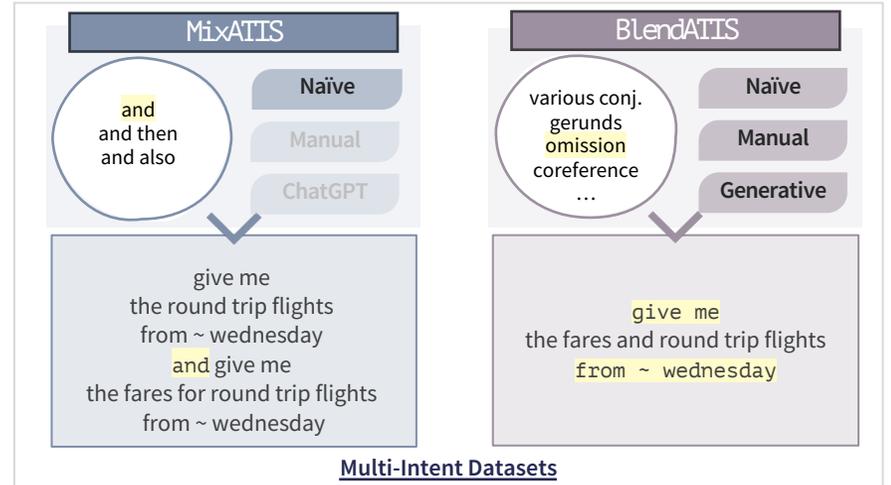 - 'and' connector 에 의존하는 explicit한 concatenation만 반영

- **BlendX**: 보다 복잡하면서 현실성을 반영한 multi-intent dataset 구축
  - 3가지의 novel concatenation approach 제안 : Naïve, Manual, Generative
  - Sentence selection 관점에서 기존의 random selection 뿐만 아니라 similarity-based strategy를 적용, generative concatenation approach에 반영
  - BlendX를 기존의 MixX와 비교/검증하기 위한 3가지의 statistical metric 고안: W, C, P
  - 데이터셋 공개 예정 : MixX의 확장(CLINC150/Banking77)과 BlendX 데이터셋 공개

#1 Selection

Single-Intent Datasets

Banking77, ATIS ✓, CLINC150, SNIPS

give me the round trip flights from ~ wednesday — atis_flight

give me the fares for round trip flights from ~ wednesday — atis_airfare

#2 Concatenation

MixATIS

and, and then, and also — Naïve, Manual, ChatGPT

give me
the round trip flights
from ~ wednesday
and give me
the fares for round trip flights
from ~ wednesday

BlendATIS

various conj., gerunds, omission, coreference … — Naïve, Manual, Generative

give me
the fares and round trip flights
from ~ wednesday

Multi-Intent Datasets

#3 Evaluation

	Mix	Blend
$W(utt)$	0	1
$C(utt)$	0	0
$P(utt)$	0	0

custom metric   baseline evaluation (ChatGPT)   visualization (Mix, Blend)

HYU 한양대학교 HANYANG UNIVERSITY

Limitation

- **Limitation of 3-Metric**

    - Semantic Complexity 고려의 필요성

        - Metric 상 유의하다면 과연 그 발화는 의미적으로도 complex한가?

            e.g.
            - i'd like to improve my credit score (improve_credit_score)
            - can you help me find my credit score (credit_score)
            → i'd like to improve my credit score can you help me find it

    - 각 metric 별 독립성

        - 결합한 발화에 대명사가 있다면, 결합 후 단어 개수가 줄어드는가?
        - 결합한 발화에 접속사가 없다면, 결합 후 단어 개수가 줄어드는가?
        - 결합 후 단어 개수가 줄어든다면, 결합한 발화에 대명사가 있거나 접속사가 없는가?

Limitation

- **Single dataset issue**

  - Label overlapping

    - CLINC150: *improve_credit_score*, *credit_score*
    - Banking77: *getting_virtual_card*, *get_disposable_virtual_card*
    → ChatGPT concatenation 시 문장을 임의로 생략하거나, intent를 유지하지 못하는 결과 다수 발생

  - 삭제된 문장 부호

    - 발화 의미 파악에 모호성을 증가시킬 가능성
    e.g. add nazad nazad kalino mome to escapada and find a world apart AddToPlaylist#SearchCreativeWork

> Cecilia Ying and Stephen Thomas. Label Errors in BANKING77. ACL 2022 Workshop on Insights from Negative Results in NLP

Limitation

- **Limitation of Manual approach**

utt1	utt2	concatenation	implication
what is the least expensive fare from boston to salt lake city (atis_fare)	what are the fares for ground transportation in denver (atis_ground_fare)	what is the least expensive fare from boston to salt lake city finally what are the fares for ground transportation in denver	두 의문문을 finally로 연결
why isn't my id being verified (unable_to_verify_identity)	my top up was denied in the app (top_up_failed)	why isn't my id being verified or my top up was denied in the app	의문문과 평서문을 or로 연결

- 모든 문법 규칙을 rule-based로 반영하는 것은 현실적으로 불가능하며 비효율적
  - BlendX에 적용되지 않은 문법 규칙: 부정문, 의문문, 조건문, 관계사절, 불규칙 동사, …
  - 부자연스러운 발화 존재

- **Limitation of Generative approach**

  - 프롬프트의 지시 사항을 따르지 않은 답변 출력 → 부적절한 답변을 수동으로 필터링

utt1	utt2	concatenation	implication
weather for frankfort (GetWeather)	3 out of 6 for the last album (RateBook)	frankfort's weather gets a 3 out of 6 rating for the last album	'RateBook' 삭제
what day of the week do flights from nashwille to tacoma fly on (atis_day_name)	flight numbers from houston to dallas (atis_flight_no)	flights from nashwille to tacoma fly on what day of the week and what is the flight number from houstom to dallas	'atis_day_name' → 'atis_flight'
what is mci(atis_abbreviation)	list la (atis_city)	combine the sentences: "what is mci" and "list la"	결합 실패

Future Work

- **산출물 정리 및 검증 예정**
  - 공개할 데이터셋 정돈
  - concatenation 과정 전체 검수

- **Multi-intent 서비스 시나리오 선행 검토**

  1. **'~하고', '~하고 나서'와 같은 접속사(연결 어미)가 문장에 포함된 경우의 처리**
     - '에어컨 켜고 18도로 설정해줘' → '에어컨 켜고' + '18도로 설정해줘'

  2. **복수 slot 추출을 통한 복합 제어**
     - '핸들이랑 시트 열선 최대로 켜줘' → <핸들> + <시트> 슬롯 추출

  3. **Complex multi-intent 발화의 한국어 적용**
     - '밖에 미세먼지 좋으면 창문 다 열어줘'

Future Work

- **산출물 정리 및 검증 예정**
  - 공개할 데이터셋 정돈
  - concatenation 과정 전체 검수

Meng et al. DialogUSR: Complex Dialogue Utterance Splitting and Reformulation for Multiple Intent Detection, EMNLP 2022 Findings.

- **Multi-intent 서비스 시나리오 선행 검토**

  1. **'~하고', '~하고 나서'와 같은 접속사(연결 어미)가 문장에 포함된 경우의 처리**
  - '에어컨 켜고 18도로 설정해줘' → '에어컨 켜고' + '18도로 설정해줘'

- Q1: Multi-intent query in DialogUSR
- Q2: Split Multi-intent query to Single-intent queries
- Q3: Delete Conjunctions   sequence generation task
- Q4: Recover Missing Info. (coreference, omission)
- Q5: Recover 1st split-query   independent!
- Q6: Recover 2nd split-query and concatenate w/Q5
- Q7: Recover 3rd split-query and concatenate w/Q6

End-to-end Generative Models: Q1 → Q4
2-stage Generative Models
 (once)   2-stage model
          : Q1 → Q2 → Q4
 (casual) 2-stage model
          : Q1 → Q2 → [Q5 → Q6 → Q7]
                          almost 100% split

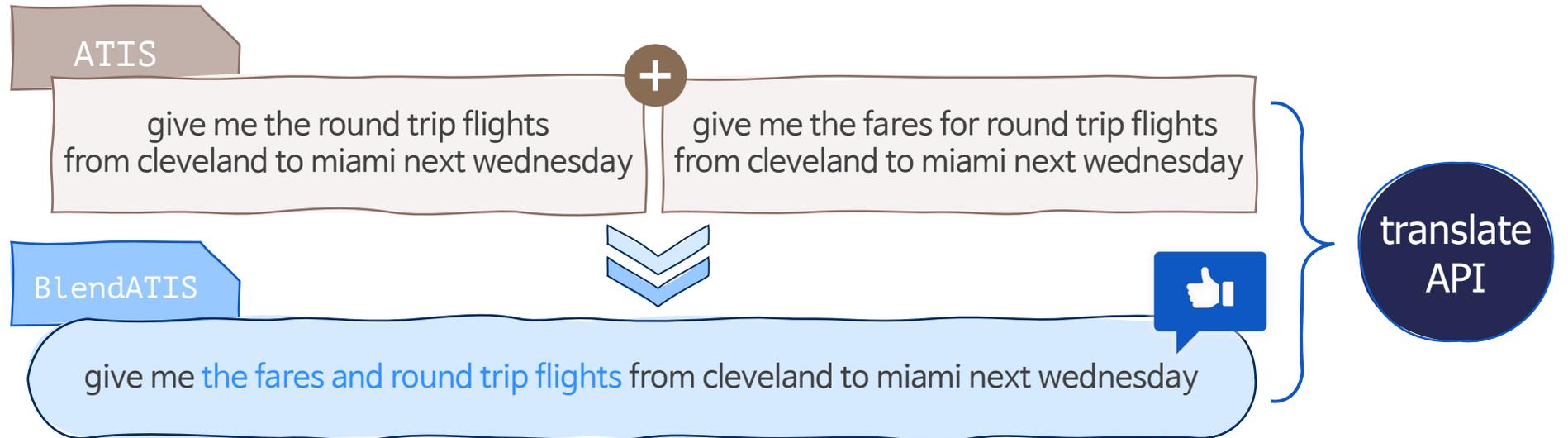| Model	MixSNIPS		MixATIS	
	BLEU	EM	BLEU	EM
T5-base	99.46	95.13	96.94	74.88
T5-large	99.60	97.64	98.52	88.77
T5-xl	**99.62**	**98.14**	**99.87**	**98.55**

Future Work

- **산출물 정리 및 검증 예정**
  - 공개할 데이터셋 정돈
  - concatenation 과정 전체 검수

- **Multi-intent 서비스 시나리오 선행 검토**

  3. **Complex multi-intent 발화의 한국어 적용**
     - '밖에 미세먼지 좋으면 창문 다 열어줘'

{ End Page }

HYU NLP Lab. (advisor: prof. 김태욱)
Hanyang University

윤예진