BlendX : Complex Multi-Intent Detection with Blended Patterns

Yejin Yoon [1], Jungyeon Lee [2], Kangsan Kim [2], Chanhee Park [3], Taeuk Kim [1,2]

1. Department of AI Application, Hanyang University, Seoul, Republic of Korea
2. Department of Artificial Intelligence, Hanyang University, Seoul, Republic of Korea
3. Hyundai Motor Company, Seoul, Republic of Korea

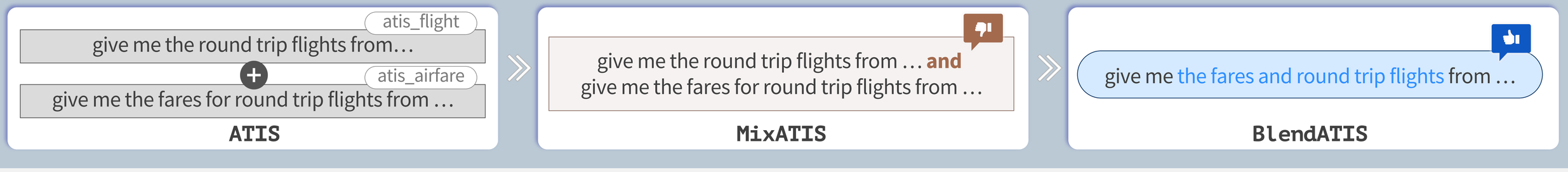
HANYANG UNIVERSITY HYUNDAI

arXiv 2403.18277

GitHub HYU-NLP/BlendX

Introduction

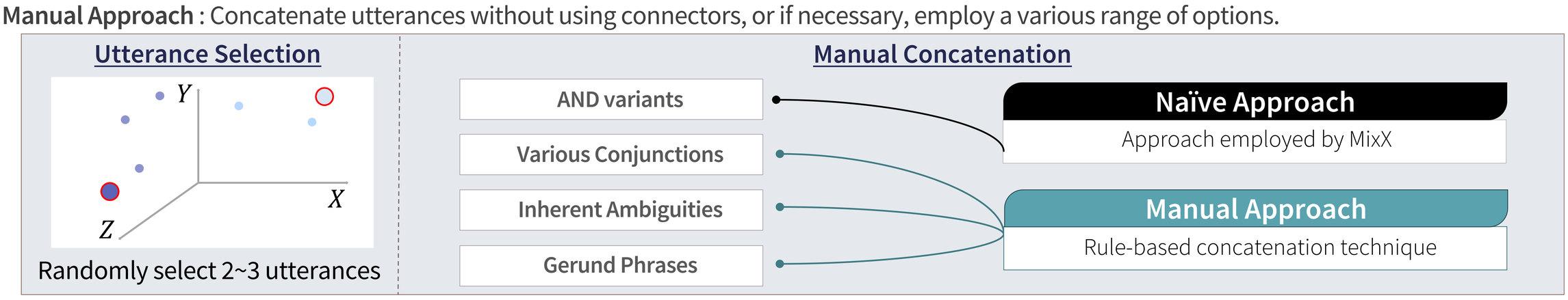
MixATIS and MixSNIPS datasets rely on only a few specific connectors (‘*and*’, ‘*and then*’, ‘*and also*’, ‘**,**(comma)’) when concatenating 2 or more single-intent utterances. Real-world conversations often involve more varied and complex ways of combining intents.



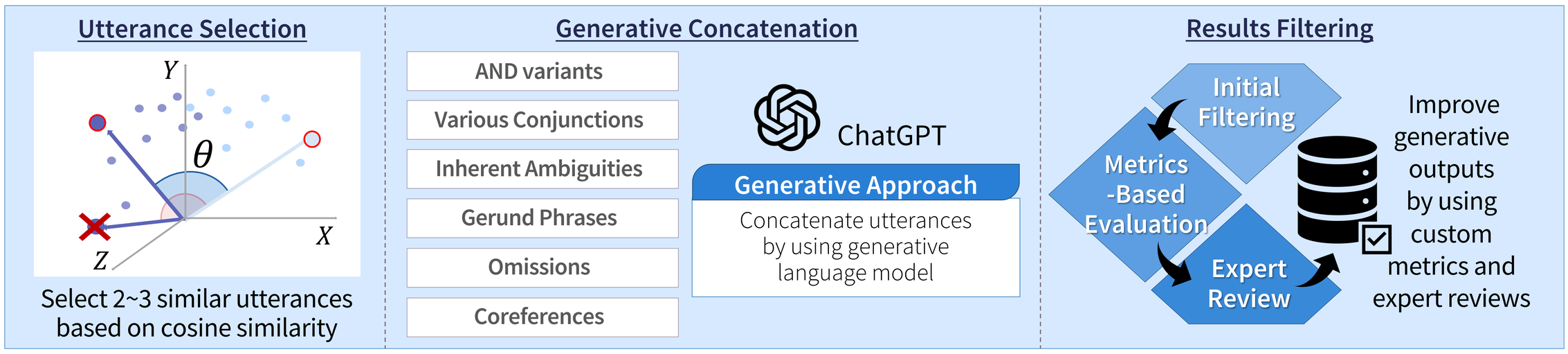
Dataset Construction

Without generating brand-new multi-intent utterances and ensuring they fit within the existing intent space, we propose 2 approaches:

- **Manual Approach** : Concatenate utterances without using connectors, or if necessary, employ a various range of options.



- **Generative Approach** : Extend ChatGPT's capabilities for producing coherent multi-intent utterances by concatenating single-intent utterances.



We’ve developed **BlendX**, a dataset that mirrors the complexity and diversity of natural dialogue. In total, it contains about 180,000 utterances, covering both manual and generative approach concatenations.

Comparative Evaluation – BlendX vs. MixX

(1) 3 CUSTOM METRICS (Differences before and after an utterance concatenation)
→ Our approach yields more realistic explicit and implicit concatenations

- $\mathrm{W}(utt, n) \stackrel{\text{def}}{=} \mathbf{1}_{\mathbb{Z}-\mathbb{N}}\left(|utt|_{word} - \sum_{i=1}^{n} |utt_i|_{word}\right)$. : Check if the word count difference
- $\mathrm{C}(utt, n) \stackrel{\text{def}}{=} \mathbf{1}_{\mathbb{Z}-\mathbb{N}}\left(|utt|_{conj} - \sum_{i=1}^{n} |utt_i|_{conj}\right)$. : Verify if the number of conjunctions
- $\mathrm{P}(utt, n) \stackrel{\text{def}}{=} \mathbf{1}_{\mathbb{N}}\left(|utt|_{pron} - \sum_{i=1}^{n} |utt_i|_{pron}\right)$. : Assess if the difference in pronoun count

Utterance 1	play my 88 keys playlist (`PlayMusic`)			
Utterance 2	add another song to my 88 keys playlist (`AddToPlaylist`)			
Strategies	**Concatenation Results**	$\mathrm{W}(utt, 2)$	$\mathrm{C}(utt, 2)$	$\mathrm{P}(utt, 2)$
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

We've used 3 metrics to validate the naturalness of our approach. An implicitly concatenated utterance is likely to receive 1 in the metrics evaluation.

(2) EVALUATE STATE-OF-THE-ART BASELINES → Significant drops

Model	Option		Dataset (Metric: accuracy)			
	Training	Test	SNIPS	ATIS	Banking77	CLINC150
TFMN	MixX	MixX	95.68* ±0.57	77.98* ±0.57	76.61 ±1.17	85.88 ±1.03
	MixX	BlendX	52.51 ±1.86	42.51 ±1.48	37.31 ±0.81	42.45 ±2.40
	BlendX	BlendX	94.93 ±0.85	76.50 ±0.83	63.99 ±0.81	77.96 ±0.82
SLIM	MixX	MixX	95.97* ±0.23	77.10* ±0.28	83.71 ±0.88	88.67 ±0.56
	MixX	BlendX	93.51 ±0.18	72.80 ±1.48	69.89 ±0.46	73.39 ±2.46
	BlendX	BlendX	95.73 ±0.86	76.92 ±0.84	75.30 ±0.71	85.62 ±0.51
gpt-3.5-turbo	-	MixX	81.68	40.30	30.90	49.22
	-	BlendX	76.18	38.84	22.67	37.55

The significant performance drop indicates BlendX’s complexity and the need for more advanced models to handle such intricate tasks.

(3) VISUALIZATION → Nearly identical distributions