

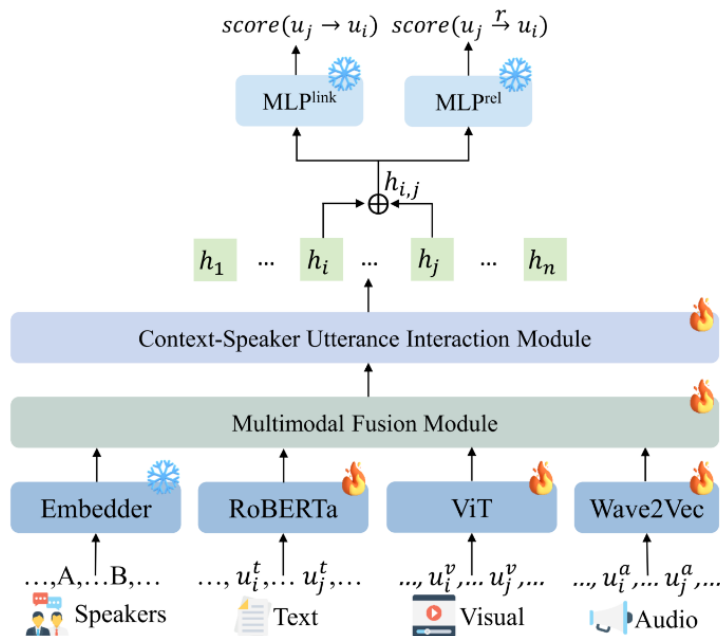
Bridging Modalities: A Unified Framework for Textual and Multimodal Dialogue Discourse Parsing

Chen GONG, Nan YU, Guohong FU

Frontiers of Computer Science, DOI: [10.1007/s11704-025-50170-0](https://doi.org/10.1007/s11704-025-50170-0)

Problems & Ideas

- Problems of existing approaches for dialogue discourse parsing:
 - develop separate dialogue discourse parsers for text-only and multimodal scenarios, largely due to the scarcity of parallel multimodal annotated datasets.
 - limits the ability to fully utilize diverse data with different modalities and poses challenges for real-world artificial intelligence applications.
- Ideas: propose a unified dialogue discourse parsing framework that bridges text-only and multimodal parsing within a single model.



The architecture of the proposed unified multimodal parser.

Main Contributions

- Contributions:
 - A unified dialogue discourse parsing framework that leverages heterogeneous multimodal data sources;
 - Bridging text-only and multimodal parsing, enabling effective dialogue discourse parsing across multiple modalities within a single unified framework;
 - Enhancing parsing capabilities in multimodal scenarios while maintaining strong performance in text-only scenarios.

Model	Dev								Test							
	STAC		Molweni		MODDP		Avg		STAC		Molweni		MODDP		Avg	
	UAS	LAS	UAS	LAS	UAS	LAS	UAS	LAS	UAS	LAS	UAS	LAS	UAS	LAS	UAS	LAS
Training with Basic Textual Dialogue Discourse Parser																
1 STAC	72.2	52.9	56.1	33.5	69.0	24.3	65.8	36.9	71.8	55.9	56.0	33.0	69.3	25.7	65.7	38.2
2 Molweni	42.2	26.4	84.0	60.4	47.7	21.6	58.0	36.2	46.5	28.3	82.7	59.6	48.1	22.6	59.1	36.9
3 MODDP	60.9	32.6	69.1	33.8	89.5	44.7	73.1	37.0	61.6	35.1	67.5	34.7	90.1	45.3	73.1	38.4
4 STAC+Molweni	70.5	49.5	84.4	60.7	75.7	29.1	76.9	46.4	71.0	53.9	82.4	59.5	76.6	30.0	76.7	47.8
5 STAC+Molweni+MODDP	54.4	32.7	81.1	58.7	79.2	48.6	71.6	46.6	56.5	36.5	78.6	56.0	89.2	36.9	74.8	43.1
Training with Unified Multimodal Dialogue Discourse Parser																
6 MODDP	58.5	30.5	69.2	38.5	89.6	43.2	72.4	37.4	60.2	34.2	67.4	38.3	90.3	44.4	72.6	39.0
7 STAC+Molweni+MODDP (1 stage)	44.7	25.2	75.5	53.1	71.3	44.2	63.8	40.8	46.3	24.0	75.5	53.1	68.3	27.0	63.4	34.7
8 STAC+Molweni+MODDP (2 stage)	70.5	49.5	84.4	60.7	88.1	39.1	81.0	49.8	71.0	53.9	82.4	59.5	88.7	40.4	80.7	51.3
Training with the Parser of Gong et al. [19]																
9 MODDP	63.5	36.1	72.4	39.2	91.2	47.6	75.7	41.0	65.1	38.9	71.2	38.9	90.5	48.4	75.6	42.1
10 STAC+Molweni+MODDP	51.5	39.0	75.5	57.1	40.2	22.2	55.7	39.5	51.9	39.2	73.7	55.9	38.4	21.3	54.6	38.8

Main results on STAC, Molweni, and MODDP datasets.