DFKI-MLST at DialAM-2024 Shared Task

Arne Binder[1*], Tatiana Anikina[2*], Leonhard Hennig[1], Simon Ostermann[2]

[1] SLT department, DFKI Berlin, [2] MLT department, DFKI Saarbrücken, Germany

The DialAM-2024 Shared Task is about **dialogue argument mining**. It focuses on the **relation extraction** (RE) part using the Inference Anchoring Theory (IAT) to represent the data. DialAM subtasks require detection of:

- **argumentative (ARI) relations**: *Inference*, *Conflict*, *Rephrase*
- **illocutionary (ILO) relations**: *Asserting*, *Agreeing*, etc.

DialAM Data

Inference Anchoring Theory (IAT) uses a bipartite-like representation to model argumentative and dialogue-related information in a single framework.

Propositions (I-nodes, left side) are linked by **argumentative relations** (S-nodes, left side) and both are anchored in **dialogue** (right side) by **illocutionary relations** (YA-nodes, center).

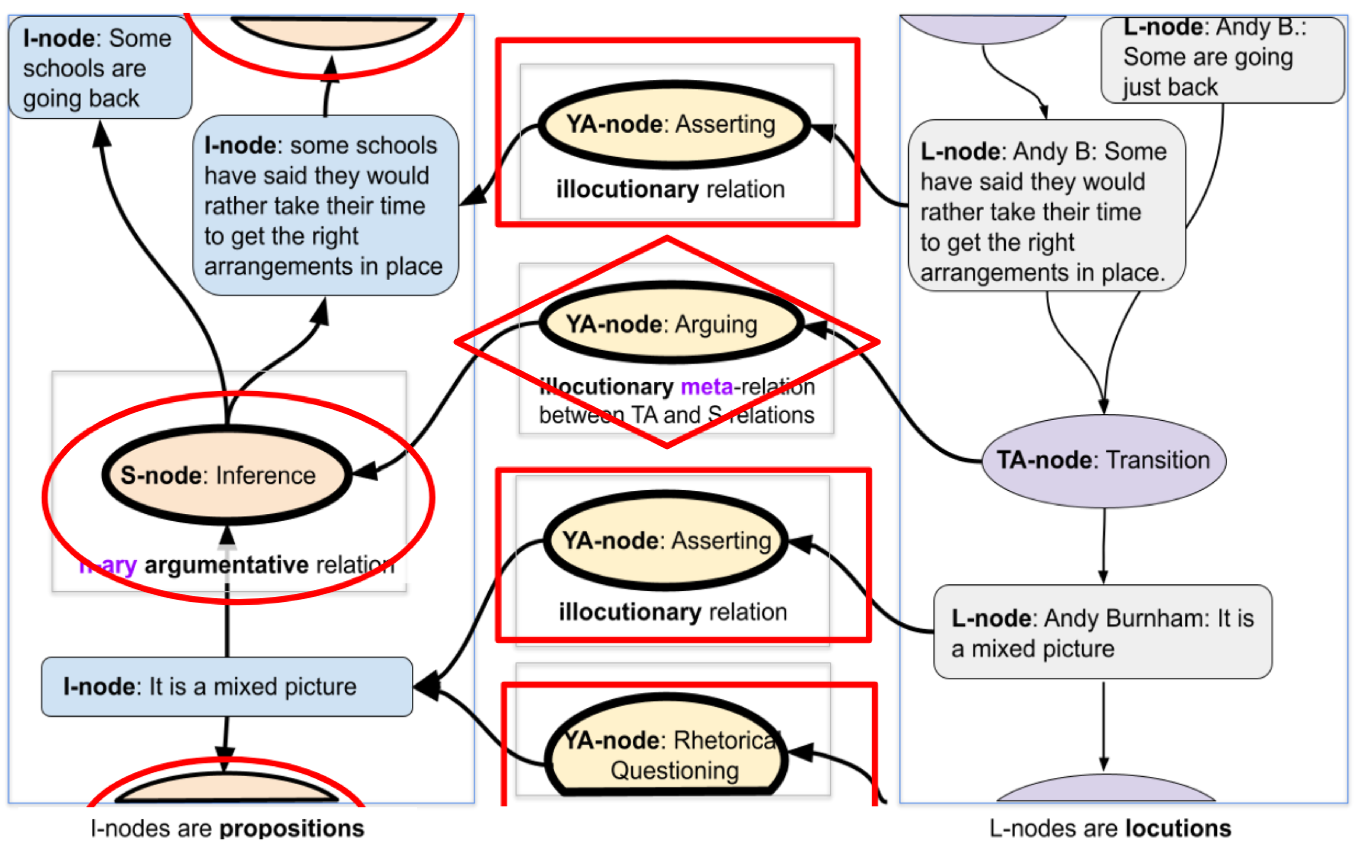


Figure 1: The DialAM 2024 shared task requires identification as well as classification of (1) argumentative S node relations (ARI; ⬭) and (2) illocutionary YA node relations (ILO; ▭ and ◇), see bold circles and arrows.

DialAM Challenges

- **Meta-relations:** some illocutionary relations (YA-nodes) **anchor argumentative relations** (S-nodes) in dialogue.
- **N-ary relation** type of the argumentative relation *Inference* (S-node) since it can have **multiple output nodes**.
- The **direction of argumentative relations** labeled as *Inference* may not follow the direction of dialogue flow.

Our Approach

We train a **single model for both subtasks** by framing them as **n-ary relation extraction on the locution nodes** (L-nodes):

1. Construct relation **candidate tuples** for all relation categories, i.e., ▭, ⬭, and ◇ (see right side).
2. Encode candidate tuples as **contiguous text**:
   1. **Serialize the locution nodes** along the dialogue flow.
   2. **Encode each candidate tuple**: Mark the respective locution node texts with special tokens (see right side).
3. Pass the annotated texts to a **BERT-like model with a linear classification head** (e.g. *DeBERTa-v3*).
4. During training: Assign **gold labels** by matching candidate tuples with gold tuples or use *NONE* if this is not possible.

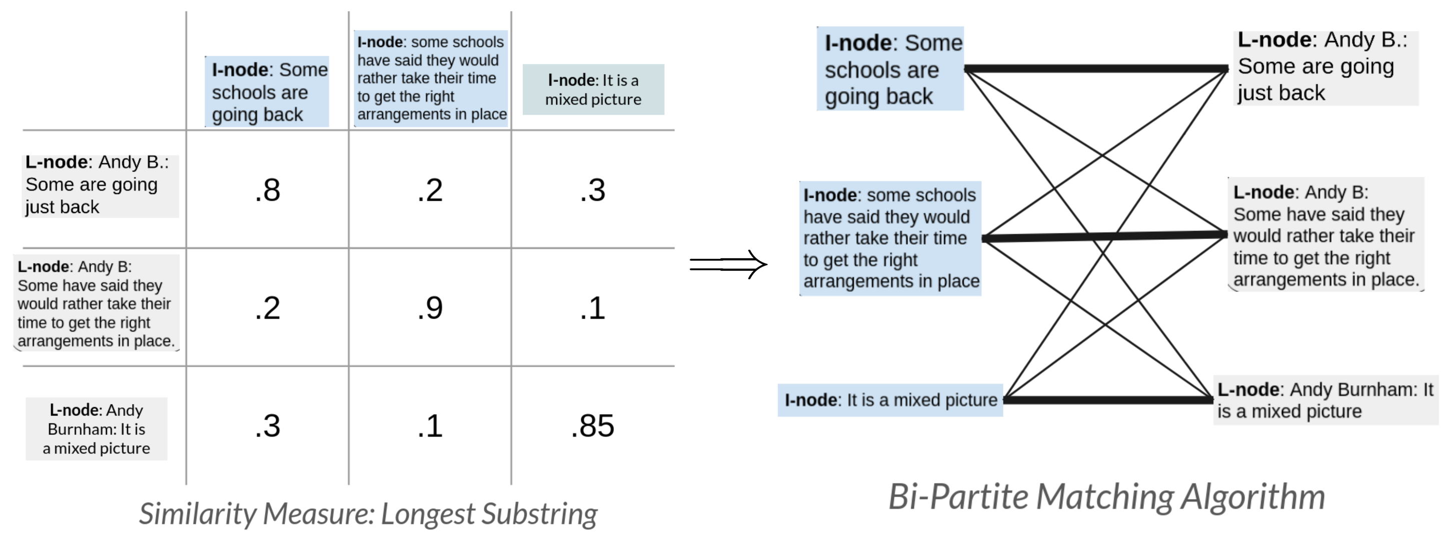
This work has been supported by the German Ministry of Education and Research (BMBF) as part of the project TRAILS.

The 11th Workshop on Argument Mining at ACL 2024, Bangkok, Thailand.

Candidate Tuple Construction

We construct relation argument **tuples consisting of nodes with roles** (arc directions, i.e., *incoming* or *outgoing*) for all relation categories.

▭ **Illocutionary YA-relations (L-node → I-node):** Construct a **bipartite matching between proposition and locution nodes** based on the longest common substring.



⬭ **Argumentative S-relations:** Use the same matching to **map the dialogue *TA relations*** along their relation arguments.

◇ **Illocutionary YA-relations (TA-node → S-node):** Connect the new **argumentative *S-relations* with the dialogue *TA relations*** from which they originate in the previous step.

Encoding Candidate Tuples

We encode the candidate tuples by marking the respective L-node (locution) texts with special tokens: **<{relation category}:{argument role}>**, e.g., **<illo–ta–s:in>**. We determine the L-nodes and roles as follows:

▭ **Illocutionary YA-relations (L-node → I-node)**: Unary relation with the anchoring L-node as the only argument.

⬭ **Argumentative S-relations**: N-ary relation with the L-nodes that anchor its arguments.

◇ **Illocutionary YA-relations (TA-node → S-node)**: N-ary relation with the arguments of the anchoring TA-node.

Some text before.</S><illo-ta-s:in>Andy B.: Some are going just back</illo-ta-s:in>
</S><illo-ta-s:in>Andy B.: Some have said [...] to get the right arrangements in place.
</illo-ta-s:in></S><illo-ta-s:out>Andy B.: It is a mixed picture.</illo-ta-s:out></S>
Some more text. relation label: *illo-ta-s:Arguing*

Figure 2: Encoding of a candidate tuple for an **illocutionary YA-relation (TA-node → S-node)** with roles ***in**coming* and ***out**going*. Role markers and relation labels are prefixed with the relation category, e.g. *illo-ta-s* for *Illocutionary YA-relations (TA-node → S-node)*, so all three cases are distinct. This results in **20 different relation labels** and **40 relation argument roles**.

Results & Conclusion

Model	ARI		ILO		GLOBAL	
	Focused	*General*	*Focused*	*General*	*Focused*	*General*
baseline	22.80	26.46	**72.09**	45.75	47.45	36.10
best-competitor	**35.89**	46.22	69.95	**81.17**	45.23	63.70
dfki-mlst (ours)	30.40	**55.33**	66.10	78.78	**48.25**	**67.05**

Table 1: Argumentative (ARI), illocutionary (ILO) and combined (GLOBAL) relation detection performance (macro F1).

- We frame the dialogue argument mining task as **n-ary relation classification** over dialogue turns (locutions).
- We apply **data simplification and normalization** to solve the task in a unified manner.
- DFKI-MLST achieves **the best scores** in the global setting of the Shared Task: 48.25 F1 Focused and 67.05 F1 General.