

# TYPIC: A Corpus of Template-Based Diagnostic Comments on Argumentation

Shoichi Naito<sup>1,2,5</sup>, Shintaro Sawada<sup>3</sup>, Chihiro Nakagawa<sup>3,2</sup>, Naoya Inoue<sup>4,2</sup>

Kenshi Yamaguchi<sup>1</sup>, Iori Shimizu<sup>3</sup>, Farjana Sultana Mim<sup>1</sup>, Keshav Singh<sup>1</sup>, Kentaro Inui<sup>1,2</sup>

<sup>1</sup>Tohoku University, <sup>2</sup>RIKEN, <sup>3</sup>Osaka Prefecture University, <sup>4</sup>Stony Brook University, <sup>5</sup>Ricoh Company, Ltd.

## Overview

### Goal

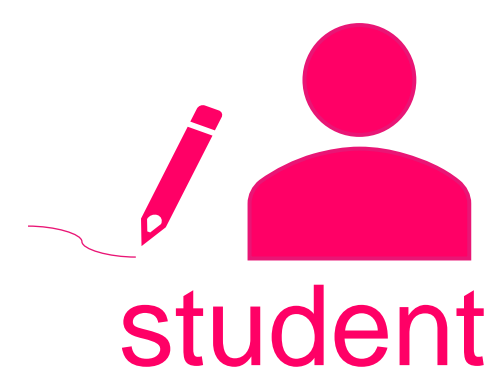
- To provide **specific** diagnostic comments on the students' counterargument for fostering critical thinking skills

### Original argument

We believe homework should be abolished. Forceful study by parents deteriorates family relationships and is one of the reasons why students dislike studying.

### Counterargument

They said homework can worsen family relationships. However, It is not true. Asking parents about homework can rather create an opportunity for communication with the family.



### Diagnostic comments

#### Previous work [Ke+'19]

Partially justified: The thesis justifies some of the author's opinions

#### This work

It lacks an explanation of why homework is a better method to realize family communication instead of other topics

### Challenge

- What task setting should be designed?
- No dataset with specific diagnostic comments annotated on an argumentation

### Contribution

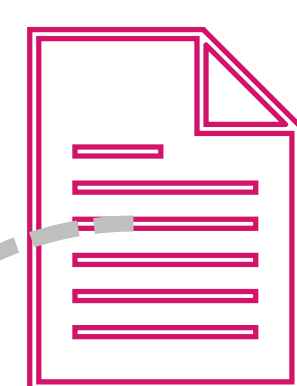
- Propose task setting for **Template-Based Diagnosis** and evaluate the quality of the induced template set
- Release TYPIC corpus, consisting of counterarguments, some of them are annotated with specific diagnostic comments in both natural language and template format

## Dataset

### Original argument



### Counterargument



### Template set (total 24 patterns)



### Statistics

|   |       |
|---|-------|
| # of <b>Original argument</b>                             | 10    |
| # of <b>Counterargument</b>                               | 1,000 |
| # of <b>Diagnostic comments</b>                           | 1,082 |
| Diagnostic comments are annotated on 197 counterarguments |       |

### Diagnostic comments

#### Natural language

Insufficient explanation of why homework is necessary as a topic of communication between parents and children, as other topics seem to exist as well.

Convert using template sets

#### Template format

It lacks an explanation of why { homework } is a better method to realize { family communication } instead of { other topics }

| Category   | Template  |
|------------|---|
| Causality  | It is unclear why { x } causes a good result of { y }                                     |
| Value      | It is unclear why { x } is good for { y }   |
| Example    | It is a limiting situation that { x }   |
| Comparison | It lacks an explanation of why { x } is a better method to realize { y } instead of { z } |
| ⋮          | ⋮   |

### Quality of induced template set

- 92.2% of comments can be represented using our template set
- 78% of comments are judged to have the same specificity compared to natural language
- The IAA of template selection is 0.517 for Cohen's Kappa

## Task Setting

### Template-Based Diagnosis

#### Step1. Template Selection

It lacks an explanation of why  is a better method to realize  instead of

#### Step2. Slot Filling

It lacks an explanation of why  is a better method to realize  instead of

### Template Set



### Motivation behind the task setting

- Reliability of automatic evaluation
  - When formulated as a generative task, reliable automatic evaluation is problematic
  - Considering the characteristics of this task, frequently occurring diagnostic comments are limited and the templates can cover many of them
- Template Selection can be reliably evaluated by metrics such as F1 and Accuracy

[Link to the dataset](#)

