

Comprehensive Argument Analysis for Chinese Argumentative Essay

1. Task Background

Engaging in thorough analysis of arguments enables a profound comprehension of the logical framework within an argumentative essay, allowing for an accurate grasp of the author's line of reasoning and thought process. This, in turn, facilitates a more comprehensive understanding of the intrinsic connections between the ideas and arguments presented in the essay. This task not only focuses on identifying argument components (e.g., claims, evidences, and elaboration), but further extracts the argumentative relations between components (e.g., support). The dataset for this task consists of argumentative essays written by Chinese high school students, with manual annotations provided by relevant language experts. The ultimate goal of this task is to improve the depth of comprehensive argumentative analysis, as well as promoting the development of applications in the fields of education.

This shared task includes two tracks:

- Track 1: Argument Component Detection (ACD): This track aims to build the fundamental argumentative structure by identifying the argumentative components. Given an essay as input, the output is expected to be the argumentative component categories of each sentence block, which indicates the role it serves within the text.
- Track 2: Argument Relation Identification (ARI): The objective of this task is to extract and classify the possible argumentative relations between various argumentative components, thereby forming an comprehensive understanding of the author's argumentative logic.

2. Task Overview

2.1 Track 1. Argumentative Component Detection (ACD)

2.1.1 Task Description

Argument components identification for argumentative essays is important for the assessment of argumentative essays. This task takes the whole argumentative essay as input and categorizes each sentence in it into 4 categories of assertion, evidence, elaboration, and others at a coarse-level. According to the content and location of the assertion sentence, this task further divides assertion into main claim, claim and restate claim. This task divides evidence into fact, anecdote, quotation, proverb and axiom according to the type and source of evidence as well.

2.1.2 Task Definition

In this task, an essay will be inputted and the model needs to predict the fine-grained category of argumentative components to which each sentence belongs. This task defined 10 fine-grained categories as displayed in the Table below.

Coarse-grained	Fine-grained	Definition	Example
----------------	--------------	------------	---------

Coarse-grained	Fine-grained	Definition	Example
论点 Assertion	中心论点 Major Claim	The main idea or proposition of the entire document, indicating the author's point of view or claim	当此之时，我们当如何抉择这些影响，如何决定自我判断？我以为，当构筑自我之思，以实践抉择之，检验之。
	分论点 Claim	Secondary ideas around the main claim	唯有前人经验不断地积累，方有如今科技之进步，文明之发展。
	重申论点 Restate Claim	Restatement of the main claims or claims	勇于探索，永于探索，以实践构建自我之思。
论据 Evidence	史实事例 Fact	Specific facts, historical facts or phenomena in society	恰如《诗经》《楚辞》开创我国悠长诗歌文脉，西方科学技术的发展也依托于牛顿，欧几里得等先贤的智慧结晶。
	个人事例 Anecdote	Examples that happened to the author	我们最常做的，就是“向多数人学习”，别人做，我也做。
	名人名言 Quotation	Citing of the writings of famous people	马克思曾以辩证唯物论告诉我们认识当从实践中来，也当经过实践检验，掷于实践中去。
	谚语俗语 Proverb	Idioms or phrases passed down among the masses	良言一句三冬暖，恶语伤人六月寒。
	公理规律 Axiom	Recognized common sense or scientific laws	因为人类对于事物的认识总是螺旋前进的。
阐述 Elaboration	-	Explanatory notes or analytical discussions of the assertion	在认识事物时，我们常受外界因素来决定自己的判断。
其他 Others	-	Sentences that do not fall into the above categories	那么，如何抉择？

2.1.3 Evaluation Metric

The total score for this track consists of two parts: a coarse-grained categorization score as well as a fine-grained categorization score, which is computed as follows:

$$Score_{Track1} = 0.5 \times MicroF1_{Track1} + 0.5 \times MacroF1_{Track1}$$

In which $MicroF1_1$ and $MacroF1$ are both computed in the **fine grained level** in our definition.

2.2 Track 2. Argument Relation Identification (ARI)

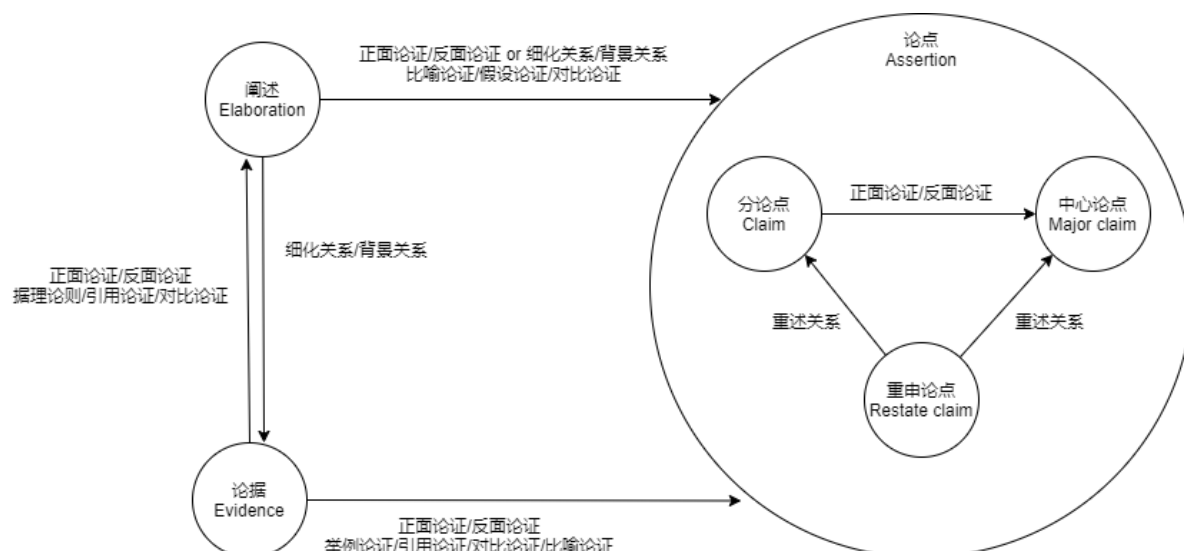
2.2.1 Task Description

Argument relation identification aims to reveal the internal argumentative connection between the components. By identifying the supportive, refuting, causal and other relationships between argument components, it is possible to understand the argumentative structure of the text in a more comprehensive way, and thus to grasp more accurately the author's point of view and the intent of the argument.

2.2.2 Task Definition

In this track, 14 types of argumentative relationships are defined at a finer level of granularity, both vertically (focusing on the relationships between different types of argumentative elements) and horizontally (focusing on the relationships between the same types of argumentative elements), which are shown in the table below. This track aims to predict the possible argument relation between **chunks of sentences**. A chunk of sentences is adjacent sentences (**with the same argument component type**) that plays similar role in the text.

Below is the definition of all possible argumentative types that appear in the dataset.

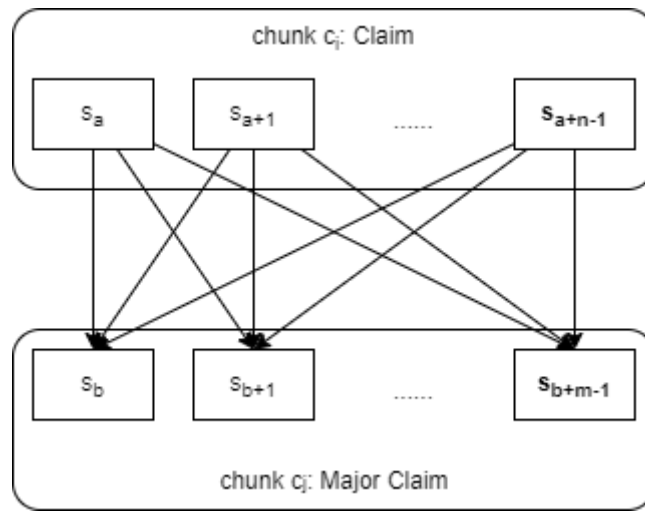


Also, we offer an example for each category, which is shown in the following table:

Argumentative Relation Type	Source Argument Unit	Target Argument Unit
正面论证 Positive	若是将认识事物比作烤蛋糕，那认识事物的方法就是模具，只听信专家权威者的话，将他人的模具套上，那烤出的蛋糕再香甜，也不会是适合自己的形状。 (Elaboration)	对于专家权威的笃信，也反映了更深层次的问题——自己并没有对事物基本认识的方法。(Claim)
反面论证 Negative	因为人失去了独立思考能力后，便不再“活着”了，正如那句“我思故我在”。 (Quotation)	综上所述，外部因素和独立思考都固然重要，于是如何将其结合便成了难处，不妨从其本质不同下手。(Claim)

Argumentative Relation Type	Source Argument Unit	Target Argument Unit
对比论证 Comparative	仪式感，它是一把双刃剑，用的好，可以帮助我们克服一切的阻碍，带着我们至达终得。但是一旦有所放纵，就宛如跳下悬崖般粉身碎骨，失去我们所拥有的一切。 (Elaboration)	只有适度的仪式感，才能让我们铭记过往，向往明天的生活。(Claim)
举例论证 Example	首先，自我思考与外部信息存在难以避免的主客观偏差，正如我与陈公孰美的闹剧，来源于其亲人对其不客观的评价所导致的信息差，对自我思考带来的影响。 (Fact)	于是我们应多去收集他人的客观信息，并结合自己的主观判断，发挥自我和他人信息的最大用处，便可更加合理的作出判断。(Claim)
引用论证 Citation	因为人失去了独立思考能力后，便不再“活着”了，正如那句“我思故我在”。 (Quotation)	综上所述我认为，外部因素和独立思考都固然重要，于是如何将其结合便成了难处，不妨从其本质不同下手。(Claim)
比喻论证 Metaphorical	若是将认识事物比作烤蛋糕，那认识事物的方法就是模具，只听信专家权威者的话，将他人的模具套上，那烤出的蛋糕再香甜，也不会是适合自己的形状。 (Elaboration)	对于专家权威的笃信，也反映了更深层次的问题——自己并没有对事物基本认识的方法。(Claim)
假设论证 Hypothetical	倘若探索陌生世界只以好奇心为唯一动力，那么好奇心最旺盛的孩童理应成为未知世界的领航者，主导一切前沿的求索，这显然不合常理。(Elaboration)	但好奇心绝不是自发性的、乐意探索的唯一诱因。(Claim)
重述关系 Restatement	可见，一个人乐意去探索陌生世界，远不止好奇心。(Restate Claim)	但好奇心绝不是自发性的、乐意探索的唯一诱因。(Claim)
细化关系 Detail	一般，变成了他人思想的木偶，也无法为巨人的肩更增一层，会使所有人安于现状，不思进取。(Elaboration)	因为人失去了独立思考能力后，便不再“活着”了，正如那句“我思故我在”。(Quotation)
背景关系 Background	上述因素，对我们现代青年亦有启发。 (Elaboration)	马克思主义不是一成不变，它随着中国国情与时俱进。 (Fact)

To make it easier, if there exists some argumentative relation $r = [r_1, r_2, \dots]$ from chunk c_i to c_j which respectively have n and m sentences, we consider **each sentence** in c_i has the same r to **each sentence** in c_j .



2.2.3 Evaluation Metric

Due to the imbalance distribution of the dataset, we utilize **micro f1** to evaluate systems. So the score for this track is computed as follows:

$$Score_{Track2} = MicroF1_{Track2}$$

3. Data Sample

```

1  {
2    "title": "时间经向下的价值认识之思",
3    "essay_ID": 0,
4    "essay_text": "当叔本华的著作问世之际，所得到的回应却是“一片死寂”，叔本华冷静的悲观主义论调似乎无法为彼时西欧的乐观主义氛围所接受。而当西欧经历诸多战争的洗礼后，人们才恍惚发觉叔本华的学术价值。有人说，经过时间的沉淀，事物的价值才能被别人所认识。\\n未经时间沉淀的事物，其价值不为我们所识，本质来自于我们人类的短视。.....",
5    "sents": [
6      {
7        "sent_text": "当叔本华的著作问世之际，所得到的回应却是“一片死寂”，叔本华冷静的悲观主义论调似乎无法为彼时西欧的乐观主义氛围所接受。",
8        "sent_type": "史实事例"
9      },
10     {
11       "sent_text": "而当西欧经历诸多战争的洗礼后，人们才恍惚发觉叔本华的学术价值。",
12       "sent_type": "史实事例"
13     },
14     {
15       "sent_text": "有人说，经过时间的沉淀，事物的价值才能被别人所认识。",
16       "sent_type": "阐述"
17     },
18     {
19       "sent_text": "未经时间沉淀的事物，其价值不为我们所识，本质来自于我们人类的短视。",
20       "sent_type": "分论点"
21     },
22     .....
23   ],
24   "relations": [
25     {
26       "from_sent": 0,
```

```

27         "to_sent": 3,
28         "relation_type": [
29             "正面论证",
30             "举例论证"
31         ]
32     },
33     {
34         "from_sent": 1,
35         "to_sent": 3,
36         "relation_type": [
37             "正面论证",
38             "举例论证"
39         ]
40     },
41     {
42         "from_sent": 2,
43         "to_sent": 3,
44         "relation_type": [
45             "背景关系"
46         ]
47     },
48     .....
49 ]
50 }

```

This shows a simplified sample from the dataset used in this task, including the basic information of the essay, i.e. `essay_ID`, `title` and document as `doc`. For track 1, you may focus on `sents`, which indicates the **argument component** category for each **sentence**. For track 2, you may focus on `relations`, which represent the annotated **argument relations** between them.

In this task, we offer some samples from our dataset for reference, which can be found in `data/samples.json`. We encourage participants to utilize other open source datasets, extra annotation and techniques to enhance their training process.

4. Testing

The submission should consist of **the source code of your model**, **a development report** that includes instructions for model usage, and **the prediction results of the testing datasets**. It is crucial to ensure that the format of the model input is consistent with the input of the testing datasets. To submit your work of this track, please write the prediction results into a JSON file with the same order with the input from testing datasets. Below is a sample input and output for your reference.

Input sample:

```

1  [
2      {
3          "title": "时间经向下的价值认识之思",
4          "essay_ID": 0,
5          "essay_text": "当叔本华的著作问世之际，所得到的回应却是“一片死寂”，叔本华冷静的悲观主义论调似乎无法为彼时西欧的乐观主义氛围所接受。而当西欧经历诸多战争的洗礼后，人们才恍惚发觉叔本华的学术价值。有人说，经过时间的沉淀，事物的价值才能被别人所认识。\\n未经时间沉淀的事物，其价值不为我们所识，本质上来来自于我们人类的短视。.....",
6          "sents": [

```

```

7         "当叔本华的著作问世之际，所得到的回应却是“一片死寂”，叔本华冷静的悲观主义论
      调似乎无法为彼时西欧的乐观主义氛围所接受。",
8         "而当西欧经历诸多战争的洗礼后，人们才恍惚发觉叔本华的学术价值。",
9         "有人说，经过时间的沉淀，事物的价值才能被别人所认识。",
10        "未经时间沉淀的事物，其价值不为我们所识，本质上来自于我们人类的短视。",
11        .....
12    ]
13 },
14 ...
15 ]

```

Output Sample:

```

1  {
2      {
3          "title": "时间经向下的价值认识之思",
4          "essay_ID": 0,
5          "sents": [
6              {
7                  "sent_text": "当叔本华的著作问世之际，所得到的回应却是“一片死寂”，
      叔本华冷静的悲观主义论调似乎无法为彼时西欧的乐观主义氛围所接受。",
8                  "sent_type": "史实事例"
9              },
10             {
11                 "sent_text": "而当西欧经历诸多战争的洗礼后，人们才恍惚发觉叔本华的学
      术价值。",
12                 "sent_type": "史实事例"
13             },
14             {
15                 "sent_text": "有人说，经过时间的沉淀，事物的价值才能被别人所认识。",
16                 "sent_type": "阐述"
17             },
18             {
19                 "sent_text": "未经时间沉淀的事物，其价值不为我们所识，本质上来自于我
      们人类的短视。",
20                 "sent_type": "分论点"
21             },
22             .....
23         ],
24         "relations": [
25             {
26                 "from_sent": 0,
27                 "to_sent": 3,
28                 "relation_type": [
29                     "正面论证",
30                     "举例论证"
31                 ]
32             },
33             {
34                 "from_sent": 1,
35                 "to_sent": 3,
36                 "relation_type": [
37                     "正面论证",
38                     "举例论证"
39                 ]

```

```

40         },
41         {
42             "from_sent": 2,
43             "to_sent": 3,
44             "relation_type": [
45                 "背景关系"
46             ]
47         },
48         .....
49     ]
50 ]
51 }

```

The overall score for the participated system is calculated as follows:

$$Score = 0.4 \times Score_{Track1} + 0.6 \times Score_{Track2}$$

Contact

Zheqin YIN(zqyin@stu.ecnu.edu.cn), East China Normal University