Explain to Me What Is Wrong With My Arguments: A Survey about Explanations in Argumentation

Camélia Guerraoui[1] Paul Reisert[3] Keshav Singh[1] Farjana Sultana Mim[1]
Naoya Inoue[4,2] Shoichi Naito[1,2,5] Wenzhi Wang[1,2] Kentaro Inui[1,2]
[1]Tohoku University [2]RIKEN [3]Beyond Reason [4]JAIST [5]Ricoh Company, Ltd.
{guerraoui.camelia.kenza.q4, singh.keshav.t4, mim.farjana.sultana.t3, naito.shoichi.t1, wang.wenzhi.r7}@dc.tohoku.ac.jp
beyond.reason.sp@gmail.com, naoya-i@jaist.ac.jp, kentaro.inui@tohoku.ac.jp

Abstract

The use of argumentation in education has been shown to improve critical thinking skills for end-users such as students, and computational models for argumentation have been developed to assist in this process. Although these models are useful for evaluating the quality of an argument, they oftentimes cannot explain why a particular argument is considered poor or not, which makes it difficult to provide constructive feedback to users to strengthen their critical thinking skills. In this survey, we aim to explore the type of explanations provided by the current computational models for argumentation, and the possibility of enhancing the explainability of such models, ultimately helping learners improve their critical thinking skills.

1 Introduction

Argumentation is the field of elaboration and presentation of arguments to debate, persuade, and agree, where an argument is made of a conclusion (i.e., a claim) supported by reasons (i.e., premises) [1]. By analogy with computational linguistics, *computational argumentation* refers to the use of computer-based methods to analyze and create arguments and debates [2]. It is a subfield of artificial intelligence that deals with the automated representation, evaluation, and generation of arguments. This field includes important applications such as mining arguments [3], assessing an argument's quality [4], reconstructing implicit assumptions in arguments [5] or even providing constructive feedback for improving arguments [6].

In the context of education, learning argumentation (e.g. writing argumentative essays, debates, etc.) has been shown to improve students' critical thinking skills [7, 8]. A number of researchers have been working on computational argumentation to support and provide tools to assist learners in improving the quality of their arguments.

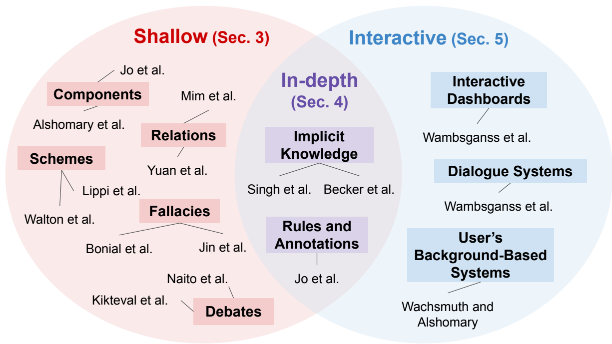

Figure 1 Overview of works focusing on the evaluation and improvement in argumentation.

Although computational models for argumentation are proven to assist students' learning and reduce teachers' workload [9, 10], such models still lack to efficiently explain how an argument can be improved; e.g., why a particular argument was labeled bad or given a low score by their automatic evaluation rubrics. In other words, the model should be not only able to provide its results but also be able to *explain the results in a comprehensive way* for the users so that users can understand, and ultimately improve their argumentation skills.

We argue that the output for current computational models for argumentation act as a type of explanation. For our survey, we categorize explanations into the following types:

- *Shallow*: The model identifies an error but doesn't explain how and why it should be corrected (i.e., *what* is wrong)
- *In-depth*: The model identifies an error and explains

Argument — *American cuisine is not healthy. Indeed hamburgers are not healthy. American cuisine is also expensive.*

Shallow Explanations (Sec. 3) — **What is wrong:** *Your argument has a hasty generalization fallacy between the first and second sentence.*

In-depth Explanations (Sec. 4) — **Why it is wrong:** *Your argument has a hasty generalization fallacy between the first and second sentence.* Without a modifier before hamburgers such as "*some*", you are implying that "*All hamburgers are not healthy*".

Interactive Explanations (Sec. 5) — **How to explain why it is wrong:**
-Model: *Are all hamburgers unhealthy?*
-Child: Uhm I don't know, maybe?
-Model: *Actually, there are some healthy burgers!*
...

Figure 2 Example of *shallow*, *in-depth* and *interactive* explanations for the same input argument

how to refine the argument in a way the end-user can understand (i.e., *why* it is wrong)

- *Interactive*: The model identifies an error at an in-depth level and adapts its explanations based on the end-user's understanding and interactions to *avoid reproduction* of the same error (i.e., *how* to explain)

In Figure 2, an argument consisting of two claims and one premise is shown. The first explanations are categorized as *shallow* as they only mention the existence of a hasty generalization fallacy. The second provides a more *in-depth* reasoning by pointing out why it is a fallacy. The last *interacts* with a child to explain more easily the fallacy.

Towards explainable computational argumentation, this paper aims to give an overview of computational argumentation on automated quality assessment. We explore work providing *shallow* (§3) and *in-depth* explanations (§4). Finally, we discuss how to develop argumentation systems that provide *interactive* explanations in a way in which learners can improve their critical thinking skills (§5).

We believe that our survey can help the community focus more on explanation in argumentation and apply it to newer models, thus making the system more explainable. [1)]

2 Related Work

2.1 Explainable AI

Explainable AI (XAI) is a research area to make AI models easily understandable for humans [11]. Based on Clinciu and Hastie [12], XAI will help both expert and non-expert users "to have a deeper understanding and the appropriate level of trust [in AI systems], which will hopefully lead to increased adoption of this technology."

To the best of our knowledge, XAI has been relatively little studied but has great potential in argumentation. Ideally, computational models should evaluate the quality of argumentation while providing efficient comments or feedback, i.e., explaining the results. However, bridging the gap between argumentation and XAI has remained unexplored.

2.2 Explainable Computational Argumentation

Several surveys have been done in the field of argumentation ([13, 14, 15, 16]) and explainability ([17, 18, 19]). In this section, we focus on the recent surveys related to explainability in argumentation.

First, Vassiliades et al. [20] highlight the potential of argumentation in explainable systems. They provide an exhaustive overview of argumentation systems for XAI by grouping them by domain, such as law, medicine, and semantic web. For each domain, papers are compared by tasks (e.g, argument classification). Despite the extensiveness of this survey, some topics important for improving explanations in argumentative systems received little attention. For example, frameworks that include arguments with commonsense knowledge have rarely been discussed, even though, they can enhance the model's explainability [21].

Čyras et al. [22] focus on the different frameworks, types, and forms of explanations. They distinguish intrinsic approaches (models using argumentative methods) from post-hoc approaches (non-argumentative models that provide complete or partial explanations). They discuss multiple forms of argumentation, such as dialogue, extensions, and sub-graphs. Their final roadmap covers the need to focus more on properties and computational aspects of argumentation-based explanations. Whereas they focus on argumentation used to explain, our work discusses how computational argumentation needs more explanation.

Moreover, our work distinguishes itself from the two previous surveys [20, 22] by focusing on evaluating and improving users' critical thinking skills.

3 Shallow Explanations

To improve students' critical thinking skills, we first need to evaluate their argumentative texts, i.e., identify

1) For more details, papers mentioned in this survey are categorized at `https://cl-tohoku.github.io/explain_arguments`.

argumentative errors. In this section, we focus on models providing *shallow* explanations, i.e., models that identify *what* should be corrected in the arguments. We discuss recent works that identify properties such as the structure of arguments helpful to assist in this process.

3.1 Argumentative Structure

As shown in Figure 1, shallow explanations consist of multiple criteria. We first discuss the **components**, **relations** followed by the **schemes**.

Components: Identifying argumentative components is one of the fundamental tasks in argumentation [23, 24, 25]. Such works primarily focus on identifying components such as *claims* and *premises*. More recently, the usefulness of identifying such components can be seen in tasks such as counter-argument generation. For example, in Alshomary et al. [26], weak premises are identified and ranked in order to generate counter-arguments.

Relations: After identifying the different components of an argumentative text, it is necessary to distinguish the multiple relations between them to assert the quality of the arguments' quality. Indeed, supporting or refuting a claim is made of complex logical moves, such as promoting, contradicting, or acknowledging a fact. Therefore it is not trivial to use correct logic. To identify the different relations patterns, Yuan et al. [27] focus on finding interactive argument pairs, whereas Mim et al. [28] enables annotating complex attack relations.

Schemes: In addition to components and relations, Walton et al. [1] proposed a set of roughly 80 logical argumentation schemes to categorize the underlying logic. Each scheme has a set of critical questions which provide a template to assess the strength of the argument depending upon the associated scheme. Since the first work on automatically detecting argumentation schemes in argumentative texts [29], the use of such schemes has been explored in tasks such as essay scoring [30].

3.2 Complex Properties

Although a good structure with a claim and premises is necessary for a good argument, it is not sufficient. Indeed an argument has other more complex properties, such as its logical, dialectical, and rhetorical aspects. In this section, we focus only on logical fallacies (i.e., the use of false or invalid inferences) and on debate patterns (i.e., interactions between arguments from different points of view).

Fallacies: Towards giving effective feedback to students and explaining the results, logical fallacies, or errors in logical reasoning, have received attention [31, 32, 33]. Given the large number of logical fallacies that exist (over 100 types), it has been increasingly difficult to identify them in argumentative texts. Habernal et al. [31] created a gamification method for capturing common fallacies through the use of crowdsourcing. Motivated by the latter, Bonial et al. [32] aimed to capture similar fallacy types for news articles, but the low distribution of fallacy types in the wild makes identification challenging.

Debates: In a case of a debate, an opponent is willing to give a counter-argument synchronously and interactively. Analyzing and evaluating a debate is a difficult task as we need to retrieve not only the argumentation structure of each opponent but also the relations between them.

Bao et al. [34] focuses on argument pair extraction (APE), which consists of finding two interactive arguments from two argumentative passages of a discussion. Although the APE task gives insights into relations between different argumentative texts, it does not indicate complex relations (i.e., how claims, supports, attacks and the intention of the speakers are interrelated). To palliate this issue, Hautli-Janisz et al. [35] identified and analyzed the dialogical argumentative structure of debates using Inference Anchoring Theory (IAT) [36]. Following the same IAT theory, Kikteva et al. [37] investigated the role of different types of questions (e.g., pure, assertive, and rhetorical questions) in dialogical argumentative setting and showed that different type of question leads to different argumentative discourse. Focused more on the opponent's side of a debate, Naito et al. [38] propose diagnostic comments for assessing the quality of counter-arguments by providing expressive, informative and unique templates. The comments are then written by template selection and slot filling.

Although the identification of such argumentative structures (components, relations, and schemes) and properties (fallacies and debates pattern) is important, it has limitations in terms of effective feedback. Identifying a missing claim or a wrong premise is not enough to properly understand how to improve the argumentation. Therefore we relate the identification of structure and properties to *shallow explanations* in the sense that end-users can still benefit from the output of the models.

4 In-Depth Explanations

Although shallow explanations help end-users to identify their mistakes, they tend to be minimalist and need more guidance. Shallow explanations can be hard to understand, specially for beginners in argumentation. To explain more effectively the errors in an argument, a model should go a step further, hence by providing *in-depth* explanations, which attempt to identify the argument's implicit components to explain *why* there is an error in an argument.

Implicit Knowledge and Reasoning in Arguments: To provide *in-depth* explanations, we need to know how to refine the argument, i.e., how to identify implicit information. Recently many works have focused their attention on this aim. The main goal of such studies is to make the structure and reasoning of arguments explicit to better explain the arguments for humans. Additionally, this focus can eventually help build Robust Argumentation Machines that can be enriched with language understanding capacity. The ExpLAIN project Becker et al. [39] and Jo et al. [40] are one such example that focuses extensively on reconstructing implicit knowledge in arguments by relying on knowledge graphs among others. Taking a step further in this direction, Singh et al. [41] proposed to utilize such implicit information to bridge the implicit reasoning gap in arguments to help students explain their arguments better.

Rules and Annotations: Another way to provide *in-depth* explanations is to understand how a model reaches its conclusion when asserting the quality of an argument. For example, in the case of evaluation of an argument's logic, Jo et al. [42] provided LogBERT, a more interpretable model based on logical and theory-informed mechanisms between two statements. LogBert relies on multiple rules that specify evidence for the support and attack relations between a claim and a statement. Although the use of rules gives a glance of explanation, LogBert remains "a black-box model with some insightful explainability." If we know how a model identifies a mistake in an argument, we can use these mechanisms to explain the diagnosis of an argument, which can help refine it.

5 Interactive Explanations

Even if *in-depth* explanations are a step towards understanding and guidance, they are static, which can be problematic depending on the end-user. Indeed beginners or professionals in argumentation do not need the same amount of feedback. A child and an adult have different levels of understanding and knowledge. Therefore it is important that a model knows *how* to explain the errors and hence adapts its output by providing *interactive* explanations.

To reach that aim, Wambsganss et al. [43] provide a visual feedback dashboard to help students see any logical error in their argumentation. Based on this learning support system, a user can easily see if a premise or a claim is missing. The dashboard provides different granularity levels of explanations, which enables the user to control the amount of needed information. Taking a step further in this direction, Wambsganss et al. [10] created an interactive educational system that uses interactive dialogues to teach students about the argumentative structure of a text. The system provides not only feedback on the user's texts but also learning session with different exercises.

Although Wambsganss et al. [10] propose different granularity levels of explanations, their study is restrained to students from their university. Having end-users from different backgrounds may imply the need for new levels of explanations. Indeed, Wachsmuth and Alshomary [44] showed that the age of the explainee changes the way an explainer explains the topic at hand. Information such as the age of the learner should be considered in future interactive argumentative feedback systems, where terminology such as *fallacy* and their existence would require different approaches of explanation for younger students (i.e., elementary) in comparison to older students.

Therefore we think models should in the future provide more *interactive* explanations (i.e., precisely adjusted by considering the background of the learner) to efficiently improve the critical thinking skills of an end-user.

6 Conclusion

In this survey, we explored several works providing explanations in argumentation, following the categories *shallow* (§3), *in-depth* (§4) and *interactive* (§5). In the future, we will extend our survey with other works, specially focusing on quality framework and scoring, which have been left behind. We will also approach the rhetorical and dialectical aspects of an argument which have received little or even no attention in this survey.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 22H00524.

References

[1] Douglas Walton, Christopher Reed, and Fabrizio Macagno. **Argumentation schemes**. Cambridge University Press, 2008.

[2] Iryna Gurevych, Eduard H. Hovy, Noam Slonim, and Benno Stein. 20 15512 – debating technologies 2 table of contents executive summary. 2016.

[3] Khalid Al-Khatib, Henning Wachsmuth, Matthias Hagen, Jonas Köhler, and Benno Stein. Cross-domain mining of argumentative text through distant supervision. In **Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies**, pp. 1395–1404, San Diego, California, June 2016. Association for Computational Linguistics.

[4] Roxanne El Baff, Henning Wachsmuth, Khalid Al-Khatib, and Benno Stein. Challenge or empower: Revisiting argumentation quality in a news editorial corpus. In **Proceedings of the 22nd Conference on Computational Natural Language Learning**, pp. 454–464, Brussels, Belgium, October 2018. Association for Computational Linguistics.

[5] Ivan Habernal, Henning Wachsmuth, Iryna Gurevych, and Benno Stein. The argument reasoning comprehension task: Identification and reconstruction of implicit warrants. In **Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)**, pp. 1930–1940, New Orleans, Louisiana, June 2018. Association for Computational Linguistics.

[6] Thiemo Wambsganss, Christina Niklaus, Matthias Söllner, Siegfried Handschuh, and Jan Marco Leimeister. A corpus for argumentative writing support in German. In **Proceedings of the 28th International Conference on Computational Linguistics**, pp. 856–869, Barcelona, Spain (Online), December 2020. International Committee on Computational Linguistics.

[7] R.T. Pithers and Rebecca Soden. Critical thinking in education: a review. **Educational Research**, Vol. 42, No. 3, pp. 237–249, 2000.

[8] Linda Behar-Horenstein and Lian Niu. Teaching critical thinking skills in higher education: A review of the literature. **Journal of College Teaching and Learning**, Vol. 8, , 02 2011.

[9] Charles Twardy. Argument maps improve critical thinking. **Teaching Philosophy**, Vol. 27, , 06 2004.

[10] Thiemo Wambsganß, Tobias Kueng, Matthias Söllner, and Jan Marco Leimeister. Arguetutor: An adaptive dialog-based learning system for argumentation skills. pp. 1–13, 05 2021.

[11] Amina Adadi and Mohammed Berrada. Peeking inside the black-box: A survey on explainable artificial intelligence (xai). **IEEE Access**, Vol. PP, pp. 1–1, 09 2018.

[12] Miruna-Adriana Clinciu and Helen Hastie. A survey of explainable AI terminology. In **Proceedings of the 1st Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence (NL4XAI 2019)**, pp. 8–13. Association for Computational Linguistics, 2019.

[13] Zixuan Ke and Vincent Ng. Automated essay scoring: A survey of the state of the art. In **Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19**, pp. 6300–6308. International Joint Conferences on Artificial Intelligence Organization, 7 2019.

[14] Beata Beigman Klebanov and Nitin Madnani. Automated evaluation of writing – 50 years and counting. In **Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics**, pp. 7796–7810, Online, July 2020. Association for Computational Linguistics.

[15] John Lawrence and Chris Reed. Argument Mining: A Survey. **Computational Linguistics**, Vol. 45, No. 4, pp. 765–818, 01 2020.

[16] Xinyu Wang, Yohan Lee, and Juneyoung Park. Automated evaluation for student argumentative writing: A survey, 2022.

[17] Marina Danilevsky, Kun Qian, Ranit Aharonov, Yannis Katsis, Ban Kawas, and Prithviraj Sen. A survey of the state of explainable ai for natural language processing, 2020.

[18] Sheikh Rabiul Islam, William Eberle, Sheikh Khaled Ghafoor, and Mohiuddin Ahmed. Explainable artificial intelligence approaches: A survey, 2021.

[19] Mareike Hartmann and Daniel Sonntag. A survey on improving nlp models with human explanations, 2022.

[20] Alexandros Vassiliades, Nick Bassiliades, and Theodore Patkos. Argumentation and explainable artificial intelligence: a survey. **The Knowledge Engineering Review**, Vol. 36, p. e5, 2021.

[21] Swarnadeep Saha, Prateek Yadav, Lisa Bauer, and Mohit Bansal. ExplaGraphs: An explanation graph generation task for structured commonsense reasoning. In **Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing**, pp. 7716–7740, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.

[22] Kristijonas Čyras, Antonio Rago, Emanuele Albini, Pietro Baroni, and Francesca Toni. Argumentative xai: A survey, 2021.

[23] Simone Teufel. Argumentative zoning: Information extraction from scientific text. 1999.

[24] Christian Stab and Iryna Gurevych. Identifying argumentative discourse structures in persuasive essays. In **Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)**, pp. 46–56, Doha, Qatar, October 2014. Association for Computational Linguistics.

[25] Yohan Jo, Jacky Visser, Chris Reed, and Eduard Hovy. Extracting implicitly asserted propositions in argumentation, 2020.

[26] Milad Alshomary, Shahbaz Syed, Arkajit Dhar, Martin Potthast, and Henning Wachsmuth. Counter-argument generation by attacking weak premises. In **Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021**, pp. 1816–1827, Online, August 2021. Association for Computational Linguistics.

[27] Jian Yuan, Zhongyu Wei, Donghua Zhao, Qi Zhang, and Changjian Jiang. Leveraging argumentation knowledge graph for interactive argument pair identification. In **Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021**, pp. 2310–2319, Online, August 2021. Association for Computational Linguistics.

[28] Farjana Sultana Mim, Naoya Inoue, Shoichi Naito, Keshav Singh, and Kentaro Inui. LPAttack: A feasible annotation scheme for capturing logic pattern of attacks in arguments. In **Proceedings of the Thirteenth Language Resources and Evaluation Conference**, pp. 2446–2459, Marseille, France, June 2022. European Language Resources Association.

[29] Vanessa Wei Feng and Graeme Hirst. Classifying arguments by scheme. In **Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies**, pp. 987–996, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.

[30] Yi Song, Michael Heilman, Beata Beigman Klebanov, and Paul Deane. Applying argumentation schemes for essay scoring. In **Proceedings of the First Workshop on Argumentation Mining**, pp. 69–78, Baltimore, Maryland, June 2014. Association for Computational Linguistics.

[31] Ivan Habernal, Raffael Hannemann, Christian Pollak, Christopher Klamm, Patrick Pauli, and Iryna Gurevych. Argotario: Computational argumentation meets serious games. In **Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations**, pp. 7–12, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.

[32] Claire Bonial, Austin Blodgett, Taylor Hudson, Stephanie M. Lukin, Jeffrey Micher, Douglas Summers-Stay, Peter Sutor, and Clare Voss. The search for agreement on logical fallacy annotation of an infodemic. In **Proceedings of the Thirteenth Language Resources and Evaluation Conference**, pp. 4430–4438, Marseille, France, June 2022. European Language Resources Association.

[33] Zhijing Jin, Abhinav Lalwani, Tejas Vaidhya, Xiaoyu Shen, Yiwen Ding, Zhiheng Lyu, Mrinmaya Sachan, Rada Mihalcea, and Bernhard Schölkopf. Logical fallacy detection, 2022.

[34] Jianzhu Bao, Jingyi Sun, Qinglin Zhu, and Ruifeng Xu. Have my arguments been replied to? argument pair extraction as machine reading comprehension. In **Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)**, pp. 29–35, Dublin, Ireland, May 2022. Association for Computational Linguistics.

[35] Annette Hautli-Janisz, Zlata Kikteva, Wassiliki Siskou, Kamila Gorska, Ray Becker, and Chris Reed. QT30: A corpus of argument and conflict in broadcast debate. In **Proceedings of the Thirteenth Language Resources and Evaluation Conference**, pp. 3291–3300, Marseille, France, June 2022. European Language Resources Association.

[36] Kasia Budsziyska, Mathilde Janier, Chris Reed, Patrick Saint-Dizier, Manfred Stede, and Olena Yakorska. A model for processing illocutionary structures and argumentation in debates. In **9th International Conference on Language Resources and Evaluation (LREC 2014)**, Vol. 14, pp. electronic–medium. European Language Resources Association (ELRA), 2014.

[37] Zlata Kikteva, Kamila Gorska, Wassiliki Siskou, Annette Hautli-Janisz, and Chris Reed. The keystone role played by questions in debate. In **Proceedings of the 3rd Workshop on Computational Approaches to Discourse**, pp. 54–63, Gyeongju, Republic of Korea and Online, October 2022. International Conference on Computational Linguistics.

[38] Shoichi Naito, Shintaro Sawada, Chihiro Nakagawa, Naoya Inoue, Kenshi Yamaguchi, Iori Shimizu, Farjana Sultana Mim, Keshav Singh, and Kentaro Inui. TYPIC: A corpus of template-based diagnostic comments on argumentation. In **Proceedings of the Thirteenth Language Resources and Evaluation Conference**, pp. 5916–5928, Marseille, France, June 2022. European Language Resources Association.

[39] Maria Becker, Siting Liang, and Anette Frank. Reconstructing implicit knowledge with language models. In **Proceedings of Deep Learning Inside Out (DeeLIO): The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures**, pp. 11–24, Online, June 2021. Association for Computational Linguistics.

[40] Yohan Jo, Haneul Yoo, JinYeong Bak, Alice Oh, Chris Reed, and Eduard Hovy. Knowledge-enhanced evidence retrieval for counterargument generation. In **Findings of the Association for Computational Linguistics: EMNLP 2021**, Online, November 2021. Association for Computational Linguistics.

[41] Keshav Singh, Naoya Inoue, Farjana Sultana Mim, Shoichi Naito, and Kentaro Inui. IRAC: A domain-specific annotated corpus of implicit reasoning in arguments. In **Proceedings of the Thirteenth Language Resources and Evaluation Conference**, pp. 4674–4683, Marseille, France, June 2022. European Language Resources Association.

[42] Yohan Jo, Seojin Bang, Chris Reed, and Eduard Hovy. Classifying argumentative relations using logical mechanisms and argumentation schemes. **Transactions of the Association for Computational Linguistics**, Vol. 9, pp. 721–739, 2021.

[43] Thiemo Wambsganß, Christina Niklaus, Matthias Cetto, Matthias Söllner, Siegfried Handschuh, and Jan Marco Leimeister. Al: An adaptive learning support system for argumentation skills. 04 2020.

[44] Henning Wachsmuth and Milad Alshomary. “mama always had a way of explaining things so I could understand”: A dialogue corpus for learning to construct explanations. In **Proceedings of the 29th International Conference on Computational Linguistics**, pp. 344–354, Gyeongju, Republic of Korea, October 2022. International Committee on Computational Linguistics.