

Conch: Competitive Debate Analysis via Visualizing Clash Points and Hierarchical Strategies

Qianhe Chen , Yong Wang , Yixin Yu , Xiyuan Zhu , Xuerou Yu, and Ran Wang

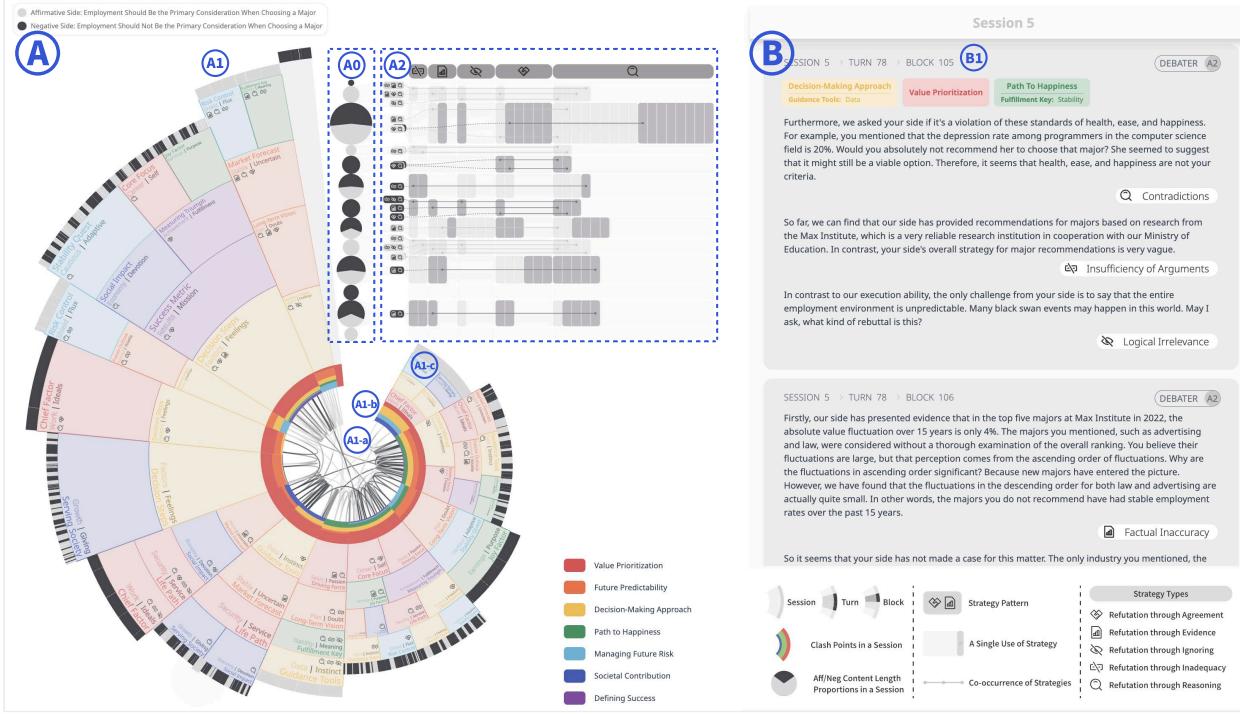


Fig. 1: The user interface of *Conch*. (A) Overview consists of two parts: the Process View(A0) and the Session View(A1) showing the evolution and interactions of debate among **blocks** based on **clash points**, disagreements, and viewpoints; the Strategy View (A2) displaying the usage and co-occurrence of debate strategies. (B) The Content View presents the specific textual content (B1), allowing users to closely examine arguments and strategies. In this figure, *Conch* visually represents the evolution and interactions of debates among **blocks**, helping users intuitively understand how disagreements and strategies develop and interact over time.

Abstract— In-depth analysis of competitive debates is essential for participants to develop argumentative skills and refine strategies, and further improve their debating performance. However, manual analysis of unstructured and unlabeled textual records of debating is time-consuming and ineffective, as it is challenging to reconstruct contextual semantics and track logical connections from raw data. To address this, we propose *Conch*, an interactive visualization system that systematically analyzes both what is debated and how it is debated. In particular, we propose a novel parallel spiral visualization that compactly traces the multidimensional evolution of clash points and participant interactions throughout debate process. In addition, we leverage large language models with well-designed prompts to automatically identify critical debate elements such as clash points, disagreements, viewpoints, and strategies, enabling participants to understand the debate context comprehensively. Finally, through two case studies on real-world debates and a carefully-designed user study, we demonstrate *Conch*'s effectiveness and usability for competitive debate analysis.

Index Terms—Competitive debate, debate analysis, clash point, visual analytics

1 INTRODUCTION

- Qianhe Chen, Y. Yu, X. Zhu, X. Yu and R. Wang are with School of Journalism and Information Communication, Huazhong University of Science and Technology (HUST). E-mail: qianhechen01@gmail.com.
- Yong Wang is with the College of Computing and Data Science, Nanyang Technological University. E-mail: yong-wang@ntu.edu.sg.
- Ran Wang is the corresponding author, also with School of Future Technology, HUST. E-mail: rex_wang@hust.edu.cn.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

The competitive debate is a structured and competitive form of communication that challenges participants' comprehensive abilities, including logical thinking, expression skills, rapid analysis, argument construction, and rebuttal techniques [64]. By analyzing previous debates, they can learn effective strategies [4], identify common mistakes [49], and understand how successful arguments are built [25]. However, this process currently relies on manually reviewing long transcripts or videos [54], which is time consuming and makes it hard to track how arguments develop or connect across different parts of a debate. For example, a team might establish a strong argument early but fail to defend it later, and this common tactic detail often goes unnoticed during conventional manual analysis. Therefore, automated methods are essential to help debaters and coaches uncover hidden patterns and

interaction dynamics in historical debates effectively and efficiently to improve performance.

Previous work on debate analysis (e.g., online [8], formal [54], political debates [4]) has contributed to identifying argument components [19, 40], labeling claims [34], and applying predictive models to assess persuasiveness [48]. These approaches offer valuable insights into what is being said in a debate, especially at the sentence level [1]. However, they remain insufficient to address two key challenges faced by debaters and coaches when analyzing competitive debates. On one hand, most existing methods focus on extracting claims, arguments, or keywords from individual sentences [1, 8, 63]. While this is useful for basic content analysis, it overlooks key elements that are essential to competitive debating, such as *clash points*, disagreements, viewpoints, and strategies [5, 7]. These elements usually connect multiple speaking turns and require examining the full debate context to properly identify and interpret them.

On the other hand, current approaches pay little attention to how debates unfold over time. Competitive debates involve multiple sessions, speaker roles, and strategy interactions across different stages [6, 64]. Debaters often respond to earlier disagreements, adapt their strategies, and establish dominance at critical *clash points* over the debate's progression [7]. However, most existing methods for debate analysis typically process debates as disconnected claims or arguments [34], overlooking how opposing views develop over time and interact across different debate stages to form meaningful argument structures [8, 33]. As a result, existing approaches struggle to adequately address two fundamental aspects of competitive debates for experts: what to debate and how to debate.

To address these challenges, we first introduce *clash points* and refutation strategies as essential elements for analyzing competitive debates. ***Clash points***, representing the core disagreements between two sides, provide a focused understanding of critical conflict areas [22]. Meanwhile, refutation strategies, such as evidence-based or reasoning-based refutation, illustrate how debaters systematically dismantle opposing arguments [22, 61]. By combining these two elements, we can track the progression of debates more effectively, highlighting both the key points of contention and the strategy approaches used by debaters and coaches. This approach allows for a more detailed and dynamic analysis of debate interactions, capturing both the structure and tactics central to competitive debates.

We propose *Conch*, an interactive visualization system that enhances the analysis of debate competitions by summarizing debate dynamics and visualizing key logical interactions. The system consists of two primary views: an overview and a detail view. The overview is divided into two parts: (1) A process view, which illustrates the evolution of *clash points*, their progression, and the interactions between debaters over time; (2) A strategy view, which depicts the distribution and co-occurrence of various refutation strategies used throughout the debate. These visualizations provide users with a comprehensive and in-depth understanding of both the argumentative structure and strategy choices. Additionally, the detail view complements the overview by displaying the specific textual content of the debate, allowing users to closely examine the arguments and strategies presented. The spiral-shaped design of *Conch*, inspired by the natural growth patterns of conch shells, uses circular timelines to show how competitive debates develop and interact over time. The effectiveness of *Conch* was demonstrated through two case studies and a carefully-designed user interview. *Conch* can be publicly accessed¹. In summary, our main contributions are:

1. An interactive system *Conch*, for the first time, is proposed to facilitate the analysis of *clash points* and strategy interactions hierarchically in competitive debates for debaters and coaches.
2. A compact parallel spiral visualization is designed to represent the temporal and structural evolution of competitive debates centered on *clash points*. Inspired by *Archimedean spirals*, the layout is arranged to optimize space while enabling block-wise exploration of debate dynamics across sessions.

3. An augmented stacked bar chart design that enables detailed comparison of strategies both within and across different debate sessions, providing a comprehensive view of strategy patterns.
4. Two case studies and a user study are conducted on two distinct datasets with experts on competitive debates to show the usefulness and effectiveness of our system.

2 RELATED WORK

Argument Analysis. Research in argument visualization has focused on representing the logical structure of arguments. These works use diagrams and graphs to map out the elements of an argument, such as premises and conclusions, and the relationships between them [17, 65]. The primary goal is to help users understand complex reasoning within static texts [30]. However, these methods are designed for stable content where arguments are already complete and presented in a sequential order. They are not well-suited for live debate competitions, which are dynamic and involve real-time, back-and-forth exchanges [61]. When we look specifically at the analysis of debate competitions, existing studies tend to focus on the debate process. They examine aspects like how debaters construct their arguments, how they interact, and the rhetorical strategies they employ [1, 2, 28]. While this research provides valuable insights into the components of a debate, it often analyzes arguments or strategies in isolation. It gives less direct attention to *clash points* — the specific moments where opposing arguments directly confront each other — and the detailed refutation strategies debaters use in response. Therefore, our work addresses this gap by introducing a framework for systematically identifying and categorizing *clash points* and the refutation strategies. This allows us to provide a deeper understanding of the strategic core of competitive debating.

Visual Analytics of Presentation. While some research focuses on visualizing single-presenter performance [58, 62], our work is concerned with visualizing the interactions among multiple participants, which is essential for understanding debates. Existing approaches in this area typically visualize two main aspects: discussion content and participant interactions. For the discussion content, researchers often extract high-level topics using methods like Latent Dirichlet allocation (LDA) or word embeddings to show what is being discussed [16, 52]. For interactions, they commonly use network graphs or timelines to show connections between speakers or the sequence of turns [15, 50]. However, these established methods have two key limitations in the context of a debate. First, topic-level views are often too general; they show what is being discussed but lose the specific arguments and the logical connections between them [18]. Second, common interaction visualizations show that participants are interacting but not how. For example, a simple line connecting two speakers cannot distinguish a basic reply from a direct, strategic refutation [52]. Our work is designed to overcome these limitations. To solve the problem of over-generalization, our system visualizes the debate using a layered structure of *clash points*, disagreements, and viewpoints, which preserves the essential argumentative context. To reveal the nature of interactions, our novel visual design explicitly shows the temporal flow and strategic purpose of each exchange. This allows users to see not just that a clash occurred, but precisely how an argument was constructed, challenged, and refuted over time.

Text Visualization. Text visualization can be broadly classified into node-link diagrams, matrix-based layouts, spatial region encodings, and glyph-augmented representations [3, 31]. Node-link diagrams are dominant for visualizing hierarchical and relational data [26, 45, 47]. Their applications include tree and graph layouts that preserve structural links [29, 37], as well as enhanced versions like extended Chord diagrams that reveal keyword co-occurrence patterns in social media [27]. Matrix-based techniques facilitate pattern detection in sequential discourse by representing similarity across utterance pairs [11, 32, 53]. Spatial layouts present a radial diagram for script-based stories [59], using concentric rings to encode scene settings, character presence, and emotional arcs, enabling structural and comparative analyses via difference overlays. Glyph-based forms integrate micro-visualizations like sparklines into tag clouds [35] or node-link diagrams [13] to simultaneously convey temporal trends and categorical frequency. These

¹<https://debate.datavizu.app/>

diverse design strategies demonstrate how structural properties of text — such as hierarchy, temporal sequence, or faceted annotation — can be mapped to distinct visual encodings, with each encoding supporting specific tasks such as exploration, comparison, or trend analysis. However, such layouts struggle with the unique structure of competitive debates: long-range argument threads spanning sessions create visual clutter in linear or tree representations. Furthermore, they typically visualize general topics or sentiment. None of them can capture the deeper, domain-specific semantic hierarchy of **clash points**, disagreements, and refutation strategies, which is critical for debate analysis and is the focus of this work.

3 REQUIREMENTS ANALYSIS

To better understand the major challenges and design requirements of debate competition visualization, we have worked closely with six debaters (**D1-6**), four debate coaches (**C1-4**), and three experts (**E1-3**) for over 6 months. We used snowball sampling [24] to find debaters and coaches with experiences of debate competitions across different languages (Chinese and English) and multiple formats with a wide range of experience (1 - 6 years). We conducted a series of semi-structured interviews with them; each interview lasted for more than one hour. Details regarding each participant and interview questions are available in the Supplementary Materials. Following their feedback, we summarized major design requirements into **R1-R4** as follows:

R1 Explore the overall debate evolution across sessions. Post-

competition analysis is a crucial practice for debaters and coaches, as demonstrated in our interviews. They review debates to capture the structure and evolution of the debate. **D1**, **D4**, and **D5** usually summarize each session separately, while **D3** and **C3** also highlighted that different sessions have specific functions. However, they find it is time-intensive and requires additional effort to trace the overall evolution. Therefore, it is crucial to present debate's chronological evolution across sessions, enabling them to quickly understand the debate's development.

R2 Identify and analyze clash points, disagreements, viewpoints, and strategies. Compared to other types of debates, competitive debates often involve more direct and explicit conflicts between debaters. Therefore, clearly identifying and analyzing these specific debate elements is particularly important. However, debaters and coaches find it difficult to accurately capture these elements from recordings and recall them later. Specifically, **D2** mentioned that understanding **clash points** and disagreements helps identify the main conflicts and key arguments. Both **D5** and **C3** emphasized that identifying debaters' viewpoints on disagreements helps understand the core differences between two sides. Moreover, all coaches agreed that identifying the debater strategies, such as refutation and defense approaches, is essential for analyzing performance and improving training. Therefore, it is necessary to clearly visualize these debate elements and allow users to explore their interrelationships.

R3 Analyze the interaction within each disagreement. Analyzing how debaters interact within each disagreement is crucial for effective debate analysis, as emphasized by debaters and coaches. In particular, **D4** and **C2** mentioned that examining how debaters refute one another reveals the logic and reasoning of their arguments. Moreover, tracking the interactions between opposing sides helps understand how each disagreement evolves. Such analysis helps debaters and coaches to assess the strengths and weaknesses of each side's argumentation. Therefore, a clear and intuitive visualization of these interactions would be highly valuable for debaters and coaches.

R4 Examine the detail of original text content. Beyond general overviews, debaters emphasized the need to analyze and learn from the concrete language used in debates. **C1** noted that the text is useful for learning strategy applications directly, and **C2** highlighted it as a resource for in-depth study. Therefore, it is necessary to present the detailed text, including specific strategies and the corresponding language.

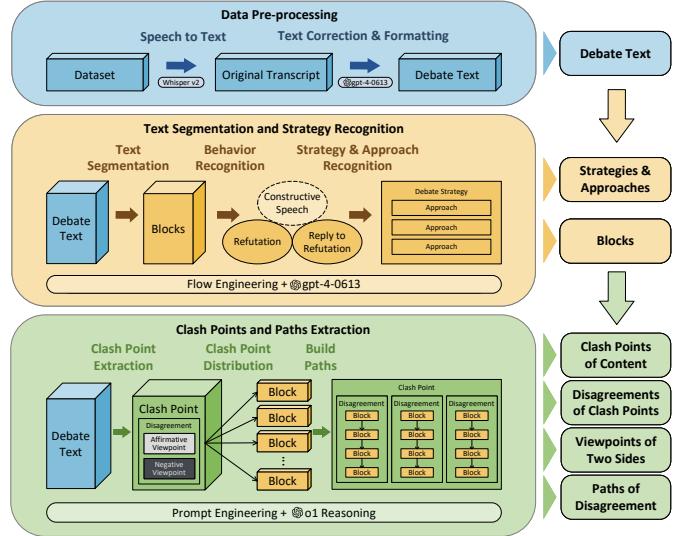


Fig. 2: The technical framework for segmenting **blocks**, extracting **clash points**, disagreements, and viewpoints, and constructing interaction paths.

4 DATA ABSTRACTION AND PROCESSING

As shown in Tab. 1, we collected a dataset from three representative competitive debates: the top-tier Chinese debate competition - International Chinese Debate Invitational² (ICDI), the largest academic competition - National Speech and Debate Tournament³ (NSDT), and the most prestigious debating society - the Oxford Union⁴ (TOU). This collection includes debates conducted in two languages (Chinese and English) and three formats, offering a cross-cultural perspective on debate structures and strategies. Notably, interactions are more frequent and direct in ICDI compared to the other competitions.

Table 1: Overview of the debate competition dataset

Competition	Language	Debater	Session	Turn	Duration	Word
ICDI	Chinese	8	13	181	1.1 h	~12 K
NSDT	English	6	8	8	1 h	~10 K
TOU	English	8	8	8	1.5 h	~12 K

Debate text extraction and structuring. The debate text was first extracted from debate videos and then structurally divided. A debate has a clear and well-defined structure. It consists of **sessions**, **turns**, and **blocks**. **Sessions** are parts of the debate defined by the debate rules. **Turns** are uninterrupted speeches given by a single debater within a session. **Blocks** are the smallest units of debate content, defined as brief arguments that include reasoning and evidence [22]. A session is consisted of one to several turns, and a turn is consisted of one to several **blocks**. Firstly, we transcribed the audio content using Whisper [44], and used the GPT-4 [41] to automatically correct errors in the initial transcription⁵. Based on the transcripts and associated competition information, we annotated the debate sides, debaters, sessions and turns of the textual content. To further evaluate the debate structure, we used large language model (LLM) techniques and designed prompts to guide GPT-4⁶ in segmenting debate texts into more integrated **blocks**. The segmentation results underwent manual review by three debate experts, achieving an average precision of 93.2% with a Fleiss' Kappa score [23] of 0.89, thereby validating the robustness of this approach.

²<https://bilibili.com/video/BV1Ca4y1y72t/>

³<https://youtu.be/ytA7V3wXux0/>

⁴<https://youtu.be/ByrhYvbwJoA/>

⁵<https://platform.openai.com/docs/guides/speech-to-text/>

⁶All mentions of GPT-4 in this paper refer to gpt-4-0613.

Construction of hierarchical debate analysis framework. We constructed a hierarchical debate analysis framework by combining the strategy framework and the content framework. To analyze debate strategies and approaches, a framework for organizing refutations in debate competitions is firstly established by integrating instructional books on debates [22, 43, 61], and its details are provided in the Supplemental Material. This framework provides a comprehensive coverage of refutation strategies in debates. Strategies and approaches of **blocks** were identified by GPT-4. The performance of GPT-4 was enhanced through a series of detailed prompts utilizing LangGPT [56], few-shot learning [39], and flow engineering [46]. Prompts included at least three standardized examples for each strategy facilitated few-shot learning. All identification results were manually validated by three debate experts and achieved an average precision of 95.2% with a Fleiss' Kappa score of 0.95, validating the accuracy and reliability of our method.

Additionally, we introduced a content framework to analyze debate content by first identifying **clash points**, then examining the **disagreements**, and finally extracting the **viewpoints**. **Clash point** is a high-level concept, which was defined as the fundamental disagreement between the two sides and all the other disagreements simply derive from that first disagreement. Disagreements derive from **clash points**, and viewpoints refer to the different opinions each side holds in response to these disagreements. The development process of each disagreement can be summarized as a **path**, which is a sequence of connected **blocks**.

The identification of **clash points**, disagreements, viewpoints and paths requires understanding of long context and complex logic, which is challenging to accomplish in a single step. As a result, we designed a three-steps approach to leverage the strongest capabilities of reasoning LLMs at each stage, ultimately ensuring accurate results. First, we used the OpenAI o1⁷ [42] to precisely extract all **clash points** and related disagreements with affirmative and negative viewpoints from the entire debate text, using the **clash point** and disagreement definition as an instruction. Each **clash point** is a phrase consisting of 2 to 4 words, and each disagreement is a phrase with 2 or 3 words, while viewpoints are concluded into single words. Then, **clash points** and disagreements were distributed to **blocks** where they were referenced via prompts and gpt-o1. **Blocks** with less than 20 words are considered too short to contain **clash points** and disagreements. Each **clash point** and disagreement may be distributed across multiple **blocks** and each **block** may be associated to multiple **clash points** and disagreements. The distribution of **clash points** and disagreements was manually validated by three debate experts using the same method, achieving an average precision of 92.7% with a Fleiss' Kappa score of 0.95. To thoroughly analyze the relationship between these **blocks**, specially designed prompts and o1 were used to identify **blocks** with logical relevance. Mutually related **blocks** can be connected into relational paths, and each disagreement is represented by a path consisting of multiple **blocks**. All paths have been verified by three experts and are considered to meet their professional analysis needs, with essentially no omissions in terms of quantity.

Expert Validation Methodology. To validate our LLM-based method, we recruited three experts, each with over five years of debate experience, to create a ground truth by annotating our dataset. The annotation process followed a structured protocol to ensure consistency [9]. First, we conducted a training session to familiarize experts with the annotation tasks. Next, they performed a pilot annotation on a small data sample and discussed their results to establish a unified annotation standard. After reaching an agreement on the standard, they proceeded with the formal annotation. During this process, any disagreements were resolved by majority rule. In the rare case that all three experts had different annotations, they would discuss the case together to determine the final label. Based on these annotations, we calculated the inter-annotator agreement using Fleiss' Kappa [23]. The resulting high level of agreement confirmed the reliability of our ground truth. The specific technical workflow and all prompts are available in the Supplemental Material.

⁷All mentions of OpenAI o1 in this paper refer to o1-2024-12-17 with high reasoning effort.

5 VISUAL DESIGN

The visual interface of *Conch* supports the exploration of debate competition evolution both in **clash points** and in strategies with intuitive visualization designs. Fig. 1 shows the snapshot of the interactive system annotated with an overview (including the Session View (A0), the Process View (A1), and the Strategy View (A2)), and the Content View (B). The Process View (Fig. 1A1) provides an integrated visualization of the temporal evolution of the debate across sessions, the main **clash points** and the interactions among **blocks**, and the Strategy View (Fig. 1A2) reveals the co-occurring strategies and their usages in each session. The Session View (Fig. 1A0) shows all sessions in the competition, and connects the Process View and the Strategy View, while the Content View (Fig. 1B) offers detailed insights into the specific content. *Conch* also facilitates interactive and collaborative exploration across the three views. A unified color encoding scheme is used throughout the interface to distinguish between the affirmative (white) and negative (black) sides and differentiate **clash points** through distinct colors.

5.1 Process View

The Process View (Fig. 1A1) is designed to demonstrate the evolution of debate (**R1**), highlighting **clash point** and disagreements (**R2**), and their interactions across sessions (**R3**). It consists of a novel designed conch-like diagram outside (Fig. 1A1-c), a chordal graph inside (Fig. 1A1-a) and a ring of color-filled blocks in the middle (Fig. 1A1-b). These elements collaboratively present an overview of the debate's evolution and interactions among **blocks** based on **clash points**.

The conch-like diagram (Fig. 1A1-a) is composed of a series of parallel **Archimedean spirals** segments. As shown in Fig. 3, each **Archimedean spiral** segment corresponds to a debate session, arranged clockwise in chronological order, with its arc length representing the session's content length. Each white and black segment (Fig. 3F) within the **Archimedean spiral** represents a **block**, with its arc length indicating the content length and color encoding two sides. With all **Archimedean spirals** constructed in a vertically oriented polar coordinate system, their start points are positioned on the axis and serve as centers for a column of tangent circles (corresponding number 1-8 in Fig. 3) in the Session View. These circles connect the Process View and the Strategy View, which are designed to represent the session's content length and two sides' content proportions.

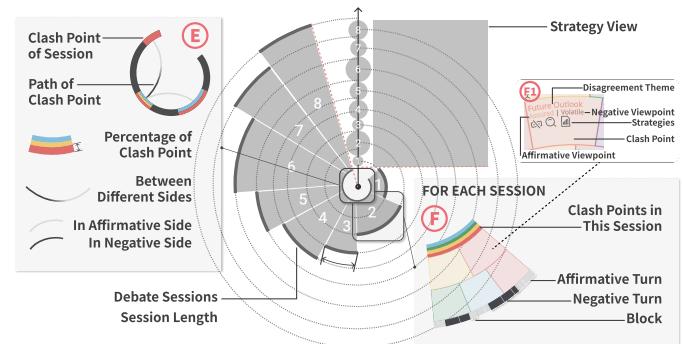


Fig. 3: The glyph design of the Process View. (E) illustrates the chord diagram and its surrounding elements; (F) illustrates the structure of the sector-shaped areas; (F1) illustrates the disagreement block content.

To analyze the **clash-points**-based interactions among **blocks** (**R3**), a chordal diagram is integrated inside the **Archimedean spirals** (Fig. 1A1-a), where each filled chord illustrates an interaction over a **clash point** between two **blocks** by connecting their respective sessions. As shown in Fig. 3E, the start and end points of the chord on each session's arc indicate the **block**'s position within that session. The color of the chord indicates whether the two connected parties are on the same side or different sides regarding the **clash point**. If the parties are from different sides for this **clash point**, the line will display two distinct colors. Each color occupies half of the line and corresponds to the color of one party's side. Conversely, if both parties belong to the same

side concerning this *clash point*, the entire line will be a single color, representing their side's color. Users are allowed to filter chords over a certain *clash point* or *block*. The periphery of the chordal graph is an *Archimedean spiral* (Fig. 1A1-b and Fig. 1A1-c). As shown in Fig. 3F, the area near the pole (Fig. 1A1-b) is radially divided into multiple color-filled sections. Each section represents a *clash points* in this session, and its width indicates the percentage of its corresponding *clash point*. The remaining central spiral and its adjacent sector-like areas (Fig. 1A1-c) display detailed information about sessions (R2). Each sector-like area (Fig. 3F), representing a single session, is further divided into multiple blocks by the disagreements within the session. The area of each block is proportional to the number of *blocks* in the corresponding disagreement, while their order is not meaningful. Each block (Fig. 3F1) is then filled with a specific color to indicate the *clash point* under discussion. In each block, the theme of the disagreement is highlighted in a larger font, while the viewpoints of the opposing sides are juxtaposed in a smaller font. Additionally, the debate strategies used in the disagreement are depicted as icons in the innermost layer of the area. All content is organized into blocks and arranged along a curve. The font size in each block is maximized based on the available space and the block's proportional size. While this approach makes efficient use of space, it can cause some reading difficulty [51]. To address this, we provide a popup that displays the full text content when the user hovers over a block, ensuring better readability. Notably, a debater might mention a *clash point* in a session, but it does not become part of an interaction with another debater. In these situations, the color for that *clash point* will appear in the middle ring (Fig. 1A1-b) to show the topic was present, but there will be no corresponding disagreement block in the outer diagram (Fig. 1A1-c). This reflects the reality of debates, where not every point raised leads to a direct conflict.

Design Alternatives and Advantages. A key design requirement is to show the entire multi-session debate on a single screen. This allows easy analysis of chronological evolution across sessions (R1) and the interactions among disagreements (R3). At the same time, users must be able to see specific details like disagreements, viewpoints, and debate strategies (R2). These goals created two constraints for our Process View. First, the view needs enough internal space to present the debate's step-by-step progression alongside detailed information. Second, it must work well with external views, such as the Strategy View. These needs naturally pointed to a timeline-based visualization.

To select a suitable layout for debate visualization, we compared three common timeline designs: linear, radial, and spiral [10, 20]. Linear layouts are intuitive for showing chronological order [12, 21], but struggle with long-range interactions common in debates (e.g., when a speaker in the final session rebuts a point from the first). Visualizing these connections creates long, overlapping lines, which leads to visual clutter and difficulties in tracing argument flow in a single screen. Radial layouts can display an entire debate on a single screen [10, 57]. However, their closed structure makes it difficult to integrate with other information views in a coordinated display. Therefore, we chose a spiral layout. Its open and extensible nature is ideal for connecting the timeline with other views. Building on prior work confirming the spiral's space efficiency [14, 60], we developed a multi-spiral design. This approach provides more sufficient and adjustable space for visualizing the temporal progression and related details in a compact way.

Our multiple *Archimedean spirals* design employs a multi-spiral layout to represent debate sessions, but its primary challenge is achieving visual balance. In our approach, concentric spirals expand outwards from the center, with each spiral segment corresponding to a single session. This arrangement naturally shows the debate's temporal progression. However, since session's content have variable lengths, a simple layout would make some segments too large while others become illegibly small. To address these challenges, we propose two key solutions: an adaptive layout algorithm and a separate view for interaction links. First, our novel algorithm balances the layout. It maps each session's content to a specific arc length, then optimizes the radius and central angle for all segments simultaneously. This process minimizes space and prevents extreme size differences. Second, to reduce visual clutter, we move all interaction links to a central chord

diagram. Links connecting the diagram to the spiral are shown only on-demand when a user selects the related *clash point*, keeping the main timeline clean. This integrated visualization is compact, balanced, and easy to interpret. Its adaptability makes the design generalizable for different debates, fulfilling the criteria [10] for a viable design: being purposeful, interpretable, and generalizable. A detailed explanation of our algorithm is available in the Supplemental Material.

5.2 Strategy View

Strategy View (Fig. 1A2) is designed to simultaneously show the usage of debate strategies and any strategies that appear together (R2). An augmented stacked bar diagram is used to help compare how often strategies are used by opposing sides. To support easy exploration, co-occurring strategies are positioned between the Session View (Fig. 1A0) and the relevant augmented stacked bar for each session.

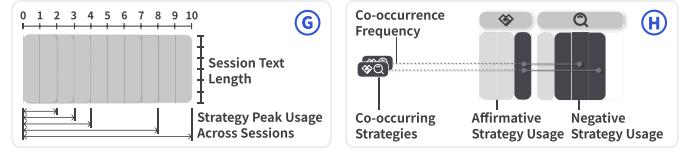


Fig. 4: The glyph design of the Strategy View. (G) augmented stacked bar design, where column height encodes session length and column width encodes the peak usage frequency of strategies within each session; (H) lines and boxes design, illustrating strategy co-occurrence frequency.

The augmented stacked bar diagram is the core component of the Strategy View (Fig. 1A2), elaborately designed to display detailed strategy usage. This diagram is positioned alongside the Session View (Fig. 1A0) with each of its rows directly corresponding to a circle in the Session View (Fig. 1A0). The height of each row matches the circle's diameter, visually representing the length of that session's content. The diagram's columns represent different debate strategy types, and the width of each column is significant: it indicates the highest frequency (peak usage) of that particular strategy across all sessions (Fig. 4G). These columns are arranged from left to right, starting with the strategy that has the lowest peak usage (narrowest column) and progressing to the one with the highest (widest column). Within this structure, each small rectangle, signifies a single instance of a strategy being used. The color of the unit clearly indicates which debate side used the strategy, while its position in a column tells us which specific strategy it was; an icon at the top of each column further identifies the strategy type. To visualize co-occurrence, if a *block* involves multiple strategies, the units representing these strategies are connected by solid lines across their respective columns, as shown in Fig. 4H. The leftmost of these connected units then has a dashed line extending to a dedicated area on the diagram's left. This area displays icons representing all strategies that co-occurred in that specific block, with their colors also indicating the debate side. Furthermore, if the same set of co-occurring strategies appears multiple times within a session, their icon-based representations on the left may overlap. Each instance in the main bar chart will still link to this representation. Consequently, the number of dashed lines pointing to a particular box or overlapping area of co-occurring strategy icons directly reflects how often that specific combination of strategies was used together in the session.

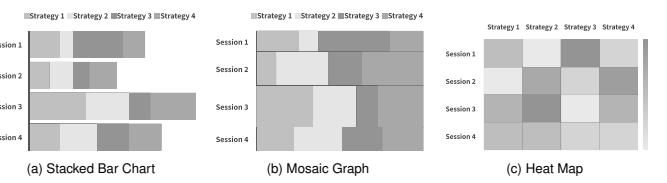


Fig. 5: Alternative designs for the Strategy View.

Design alternatives and Advantages. Common visualization methods for this type of data, such as stacked bar charts, mosaic graphs,

and heat maps, each present certain limitations when trying to clearly display all necessary information. Traditional stacked bar charts (Fig. 5 a) make it difficult to accurately compare the sizes of different segments because they lack a common baseline, which also obscures the true differences between these segments. While mosaic graphs (Fig. 5 b) are more effective for showing proportions and the lengths of individual components, they still pose challenges for comparing components across different groups and tend to downplay the actual values involved. Similarly, heat maps (Fig. 5 c), though useful for illustrating the frequency of strategy use through color sequences, do not account for variations in debate duration, potentially leading to incorrect interpretations of the debate's overall structure. These limitations highlighted several key challenges for our design: accurately showing strategy usage across different sessions, enabling easy comparison of strategy use both within a single session and between multiple sessions, and clearly representing differences in content length from one session to another. To address these challenges, we developed an augmented stacked bar diagram. This new design builds upon the familiar concept of a stacked bar chart but introduces two significant enhancements: a variable-width for the bars and non-full-format filling. The primary advantage of our approach is its ability to integrate multiple dimensions of information — including the content length of each session, the specific strategies employed, and the distribution of these strategy usages across all sessions — into a single, comprehensive view.

5.3 Content View

To help users examine the original text content (R4), the Content View (Fig. 1B) presents detailed content in a dedicated view. In this view, each **block** is shown as a card (Fig. 1B1). Each card is color-coded by its side to indicate its origin. Users can select specific **blocks** to read their details together with the surrounding context.

On each card, **block** identifiers are in the upper-left corner, while debater identifiers (e.g., “DEBATER A1” means the affirmative side’s first speaker) are placed in the upper-right corner. Below these identifiers, the card provides a collection of the **block**’s main **clash point** and key viewpoints. This allows users to quickly grasp the main points before reading full texts. To better show debate strategies within **block** content (R2), text in each **block** is segmented at sentences where strategies are applied. Strategy icons and descriptive labels are displayed beneath each text segment, indicating the specific approaches used.

5.4 Interactions

The three views in *Conch* are interconnected with globally linked interactions. Selecting **blocks** highlights the corresponding spiral segments in the Process View, units in the Strategy View, and cards in the Content View, with the linked card auto-scrolling to the top. *Conch* supports both session-based and **clash-point**-based analysis alongside hovering interactions for condensed information display.

Session-based Exploration. Users are allowed to select a session by clicking a spiral segment or a circle in the Session View, activating the Content View to display all **blocks** in the session and highlighting the corresponding row in the Strategy View. Further selection of a turn via shorter spiral segments, or direct **block** selection via units in the Strategy View or cards in the Content View reveals contextual content and auto-scrolls the selected card.

Clash-point-based Exploration. Clicking a **clash point** in any view (legend, chordal graph, Process View, or **block** card) highlights relevant **blocks** and filters the display to its interaction chords. These chords are color-matched to the **clash point**, showing the development of that specific disagreement (R3).

Expandable Information Interaction. Clicking on a condensed element, such as a disagreement block or a phrase in the **clash point** legend, expands that legend to show more content. A further click on an individual disagreement in the legend reveals its full explanation. Hovering over various elements — including markers, icons, and phrases across the **clash point**, **block**, and Strategy View views — displays tooltips with their definitions.

6 CASE STUDY

This section presents two case studies conducted on representative debate competitions in China and the United States to demonstrate the effectiveness of *Conch*, by describing how two debate experts (E1 and E2) used *Conch* to explore and analyze debates.

6.1 Case 1: Analyzing Debate Evolution and Interaction

E1 is the captain of a Chinese debate competition team, who has five years of debate experience. She was asked to analyze a Chinese debate competition from ICDI by *Conch*. The debate topic was “Whether Employment Prospects Should Be the Primary Consideration When Choosing University Majors After the Gaokao⁸”. Firstly, she explored the Overview Part, as shown in Fig. 1A. From the colorful filled blocks on the border of the chordal graph (Fig. 6(a)), she noticed that “Value Prioritization” was the main **clash point** of the debate and was discussed in all sessions, as indicated by the continuous appearance and large proportion of the red blocks. In addition, she found “Future Predictability” and “Decision-Making Approach” were also discussed continuously, with orange and yellow blocks appearing in multiple session areas. However, “Path to Happiness” was mainly discussed in the first half of the debate, as green blocks mainly appeared in the earlier area. Moreover, blocks in other colors were found appearing scattered across sessions, and therefore considered the corresponding **clash point** as non-core disagreements.

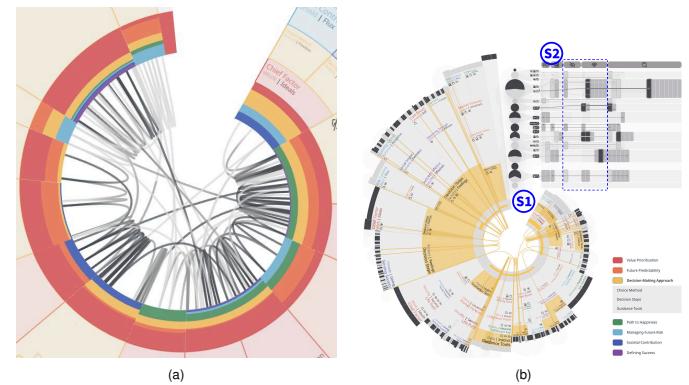


Fig. 6: (a) Chordal diagram from Case Study 1. (b) S1: The disagreement and interaction path of the **clash point** (“Decision-Making Approach”); S2: Strategies View from Case Study 1.

Explore the **clash points and the strategies.** E1 noticed that yellow blocks appeared only in sessions involving the affirmative side (S1 in Fig. 6(b)). Therefore, she started focusing on the corresponding **clash point** — “Decision-Making Approach”. Based on this observation, she concluded that the affirmative side had an advantage on this **clash point**, as the negative side rarely participated in related discussions. Then, she used the strategy view (S2 in Fig. 6(b)) to explore the strategies used in relevant sessions. She noticed that black **blocks**, representing the negative side, appeared in the column marked with a handshake icon (“Refutation through Agreement”), but never appeared in the column marked with an eye icon (“Refutation through Ignoring”). This observation further confirmed her judgment, as the negative’s strategy is mainly to agree with the affirmative side. Consequently, she further analyzed the disagreements and related interactions of this **clash point**. She chose the three disagreements of this **clash point** in turn and observed their development path in the chordal graph. As shown in Fig. 7a, Fig. 7b and Fig. 7c, E1 found the sequential clustering of yellow blocks, and these three disagreements were all initiated by the affirmative side, clearly evolved one after another over time. Specifically, the disagreement “Choice Method” appeared first, starting from session 1 and lasting for four sessions, which formed the earliest cluster of yellow blocks (Fig. 7a). Immediately afterward, the disagreement

⁸National College Entrance Examination (NCEE)

“Guidance Tool” became the main topic and was actively discussed until session 5, forming the second clear cluster (Fig. 7b). Finally, in the second half of the debate, the disagreement “Decision Steps” emerged as the primary focus, as it was heavily discussed during the final six sessions and formed the last cluster of yellow blocks (Fig. 7c).

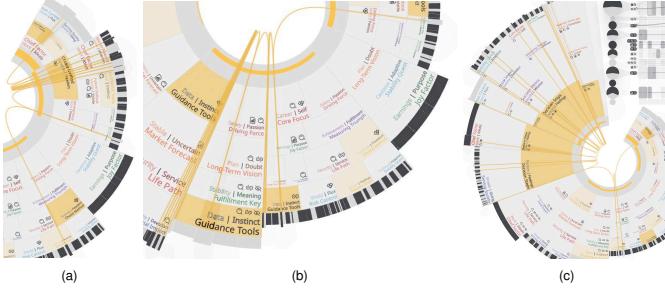


Fig. 7: Interaction paths within the **clash point** (“Decision-Making Approach”): (a) “Choice Method”. (b) “Guidance Tool”. (c) “Decision Steps”.

Delve into the interaction within a disagreement. By reviewing the three progressively developed disagreements and related viewpoints, **E1** found the negative side hold the view that “people should first reflect on themselves to learn what they want to learn before choosing college majors”. Since “Decision Steps” was the last disagreement of this **clash point**, whoever had an advantage on this disagreement would determine who won the entire **clash point**. E1 then followed the path under “Decision Steps” (Fig. 7c) to examine the detailed content. She noticed that both the beginning and end of the path corresponded to white blocks, and a long black block existed in the middle of the path. She focused particularly on understanding these three corresponding blocks. Through careful examination, she found that the affirmative side presented this disagreement using detailed and solid data, while the negative side only briefly argued their position and explained their position by “ideal”. In the final block, the affirmative side pointed out the negative side’s flaw of inconsistent reasoning, criticizing them for shifting from “reflecting on oneself” to “idealism” above all. Ultimately, she concluded that the affirmative side won this **clash point**. Interestingly, she noted that the negative side’s initial suggestion of “reflecting on oneself” did not even serve as a valid rebuttal to the affirmative side. If a person is indeed materialistic, then “reflecting on oneself” would naturally lead them to choose their major based on employment opportunities, which clearly aligns with the affirmative side’s position. This suggests that the negative side itself has a flaw in its logic.

6.2 Case 2: Understanding Overall Debate tactics

E2 is the captain of an English debate club, who primarily coaches beginners. She was asked to analyze an English competition from NSDT by *Conch*, with the topic “This House Believes That Housing Is a Guaranteed Right”. **E2** explained to us that, compared to being a debater who deeply explores every detail during a debate, she — as a coach — pays more attention to the high-level, overall tactics of the debate.

E2 first noticed that black and white chords appeared separately in the chord diagram, with very few chords mixing black and white (Fig. 8U0). This indicated that the affirmative and negative sides mainly developed arguments internally within their own sides, rather than between the two sides, which means direct interactions and clashes between the two sides are limited. She found this phenomenon unusual. Therefore, she further looked into the **clash point** labeled “Interpretation of the motion”, represented by the red color sections that frequently appeared and occupied significant space (Fig. 8U1). She discovered that the two sides differed fundamentally in their understanding and definitions of the debate topic itself. This difference explained why there were relatively few direct confrontations between them.

Next, **E2** examined two other **clash points**, “Pragmatism vs. Principle” and “Role of government policies”, represented by orange and

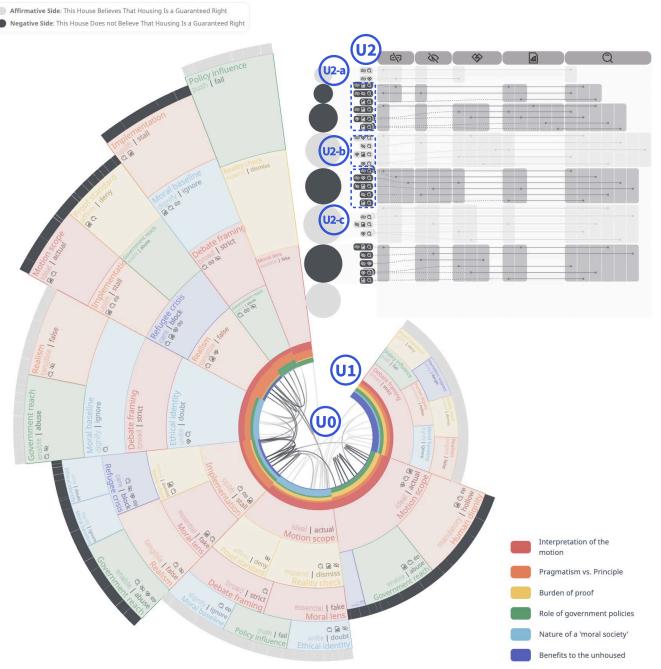


Fig. 8: The chord diagram (U0) shows few interactions between the two sides, indicated by black-and-white lines. The **clash point** “Interpretation of the motion” (U1, red) reveals fundamental differences in how each side defined the debate topic. Additional **clash points**, “Pragmatism vs. Principle” (orange) and “Role of government policies” (yellow), illustrate independent argumentation from each side. The Strategy View (U2) highlights clear differences in strategy patterns between the two sides.

yellow colors respectively, as these also occupied large spaces. She found that under these two **clash points**, the two sides had different definitions of “should”, and both sides continued independently arguing their own interpretations of the motion: the affirmative side argued that “Housing ought to be a guaranteed right”, while the negative side argued that “Housing is already a guaranteed right”.

Given these observations, **E2** shifted her attention to the strategy view (Fig. 8U2) to better understand the strategic framework of the entire debate. She noticed that black blocks appeared extensively across all columns, whereas white blocks were more concentrated in the columns labeled “Refutation through Agreement” and “Refutation through Reasoning”. Considering both sides’ differing definitions of the motion, she identified distinctly opposite patterns between the two teams: the affirmative side primarily used reasoning to indirectly strengthen their own definition, rarely directly refuting the negative side. In contrast, the negative side heavily relied on evidence to directly refute the affirmative side.

E2 further examined co-occurring strategies and discovered that for the black side (negative), “Refutation through Reasoning” frequently appeared together with “Refutation through Evidence” (Fig. 8U2-a, Fig. 8U2-c). On the other hand, for the white side (affirmative), “Refutation through Reasoning” was more often combined with “Refutation through Agreement” (Fig. 8U2-b). She concluded that these insights further highlighted the systematic differences between the affirmative side and negative side throughout the debate, not only on basic interpretations, but also on strategic framework.

7 USER STUDY

We conducted a user study by comparing the effects of debate analysis using *Conch* and other common analyzing methods: watching debate videos and reading debate transcripts. Manual debate analysis is time-consuming and cognitively demanding [54] because it is burdensome for human to identify and piece together the logic of **clash points** and refutation strategies from the extensive context of the entire debate.

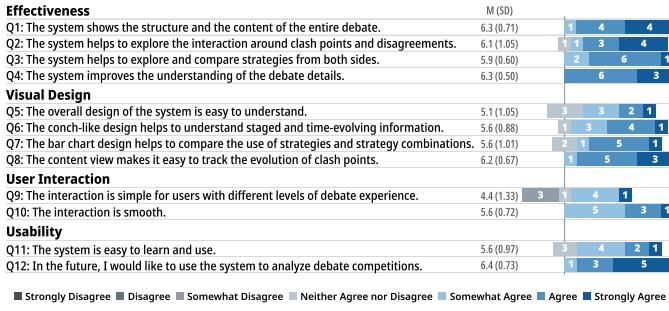


Fig. 9: User interview questionnaire results, including the effectiveness, visual design, interaction, and usability, rated on a 7-point Likert scale.

Therefore, when analyzing a debate, human debaters or coaches will inevitably take longer time and confront growing cognitive load. Taking above into consideration, we measured the time participants spent to evaluate efficiency, their cognitive load during analysis to assess the effectiveness [36], and their feedback and ratings for system to understand user satisfaction and perceived usability.

7.1 Study Design

Group design. In this study, we aimed to evaluate whether the use of *Conch* improves the debate analysis effects. We conducted a controlled experiment consisting of one experimental group and two control groups. Participants from all three groups were asked to analyze the same debate competition through a series of tasks. Specifically, participants in the experimental group were asked to use *Conch* to analyze the debate, while participants in the video control group and the text control group analyzed the debate using video and text respectively.

Data and tasks. The video and text data from the ICDI were used as the experimental material, as detailed in Tab. 1. This competition was chosen because it contains richer **clash points** and disagreements. Moreover, since the data is in Chinese, it reduces participants' cross-language comprehension burden. We designed seven debate analysis tasks (Tab. 2) to measure the completion rate of debate analysis from three aspects: what to debate (T1-T3), how to debate (T4-T5) and comprehensive analysis (T6-T7). Following these tasks, we assessed participants' cognitive load through a ten-item questionnaire for the measurement of **IL** (Intrinsic Load), **EL** (Extraneous Load), and **GL** (Germane Load) [36]. The detailed content of this questionnaire can be found in the supplementary materials. Additionally, a Likert scale (1 = strongly disagree, 7 = strongly agree) was designed for participants in the experimental group to evaluate the effectiveness, visual design, interaction and usability of *Conch* (Q1 - Q12 in Fig. 9).

Table 2: The tasks for participants to perform in our user interviews.

A. What to Debate
T1 Summarize how the content of both sides evolved across different sessions.
T2 List the clash points in this debate.
T3 Review how debaters discussed around clash points and list the most impressive interactions.
B. How to Debate
T4 List strategies used by debaters when making rebuttals.
T5 List strategies that were commonly combined and used together.
C. Comprehensive Analysis
T6 Summarize the performance of both sides.
T7 Summarize what you have learned.

Participants. We recruited 27 debaters and debate coaches to participate in the user study (7 female, 20 male; age: Mean(M) = 21.0, Standard Deviation(SD) = 2.4). All of them have 1-9 years of related experience in debate and debate analysis. Specifically, 19 of them have 1-3 years' experience, 4 have 3-5 years' experience and 4 have more than 5 years' experience. Among these participants, two (E3, E4) are world champions in Chinese debate competitions. Based on their duration of debate experience, we randomly divided them into three groups through stratified sampling, ensuring similar experience levels

within each group. Additionally, another expert (E5), who is also a world champion and the “Best Debater” in Chinese debate competitions, provided feedback on our system.

Procedure. The study was conducted both online and offline. Participants could choose their types of attendance based on their location and convenience. The *Conch* system, the debate video and the debate text were all accessible online. We first obtained their consent, and introduced the research background and experiment process for approximately 5 minutes. For the experimental group, we additionally provided a 15-minute introduction to the functions of *Conch*. Participants in this group were then allowed to freely explore the system until they were familiar with how to use it. After the introduction, participants in three groups should complete seven tasks by using *Conch*, watching debate video and reading debate text, respectively. Then, participants completed the ten-item cognitive load questionnaire, and participants in the experimental group also filled a system evaluation questionnaire. Task completion times spanned between 30-135 minutes ($M = 73.9$, $SD = 26.9$), with a debate duration of 66 minutes.

7.2 Results

Upon completing the debate analysis task, the experimental group using *Conch* showed shorter analyzing time than other two control groups and a better cognitive load profile for learning.

Task Completion Rate. All participants successfully completed the entire debate analysis task. We found that the experimental group (using *Conch*) and the two control groups all achieved a 100% completion rate. This indicates that *Conch* is as effective as the baseline methods in enabling users to finish the required analysis, ensuring that the tool does not introduce barriers to task completion.

Completion time. As shown in Fig. 11, participants in the experimental group (A) completed the tasks faster than those in the video control group (B) and the text control group (C). Specifically, participants using *Conch* in the experimental group ($M_A = 51.1$, $SD_A = 20.2$) spent 32.8% minutes less than participants watching debate video ($M_B = 76.1$, $SD_B = 17.8$) and 45.9% less than those reading debate text ($M_C = 94.4$, $SD_C = 23.8$), which demonstrated statistically significant differences ($p_B < 0.01$, $p_C < 0.01$) suggested by the Mann-Whitney U test [38]. Since debate analysis requires a global consideration of the entire content, single task is difficult to be completed and timed independently.

Cognitive Load. Cognitive load consists of intrinsic load (IL), extraneous load (EL), and germane load (GL). IL relates to task complexity and participants' prior knowledge, EL arises from unhelpful instructional features, and GL results from beneficial instructional features [36]. Thus, lower EL and higher GL indicate whether a system provides better support for users in performing tasks effectively.

As shown in Fig. 10, we found that IL scores were similar across all three groups (experimental: 3.67 ± 2.17 , control-video: 3.26 ± 1.73 , control-text: 3.96 ± 2.02), indicating comparable task difficulty and participant knowledge. Using the Mann-Whitney U test, the experimental group had significantly lower EL than the text control group (1.15 ± 1.12 vs. 3.15 ± 2.33 , $U = 17.5$, $p = 0.046$). Compared to the video control group, the experimental group's EL was lower on average, but the difference was not statistically significant (1.15 ± 1.12 vs. 1.59 ± 2.34 , $U = 44.5$, $p = 0.755$). Additionally, the experimental group's GL was significantly higher compared to both the video-based group (7.03 ± 1.85 vs. 4.22 ± 2.58 , $U = 66.5$, $p = 0.024$) and the text-based group (7.03 ± 1.85 vs. 3.94 ± 2.58 , $U = 64.0$, $p = 0.042$).

Questionnaire. Fig. 9 reports the results of the user interview questionnaire across four dimensions: effectiveness, visual design, user interaction, and usability. Overall, participants in the experimental group gave high ratings to *Conch*, with most of the closed-ended questions receiving positive feedback. Specifically, participants expressed satisfaction with the effectiveness and usability of the system, as they rated questions in these dimensions around 6 out of 7. They also provided generally positive ratings for visual design and user interaction, although scores in these two dimensions were slightly less consistent. However, scores for Q5 (“easy to understand”) and Q9 (“interaction is easy for users with any level of debate experience”) were relatively lower. For Q5, participants indicated that the visual design was some-

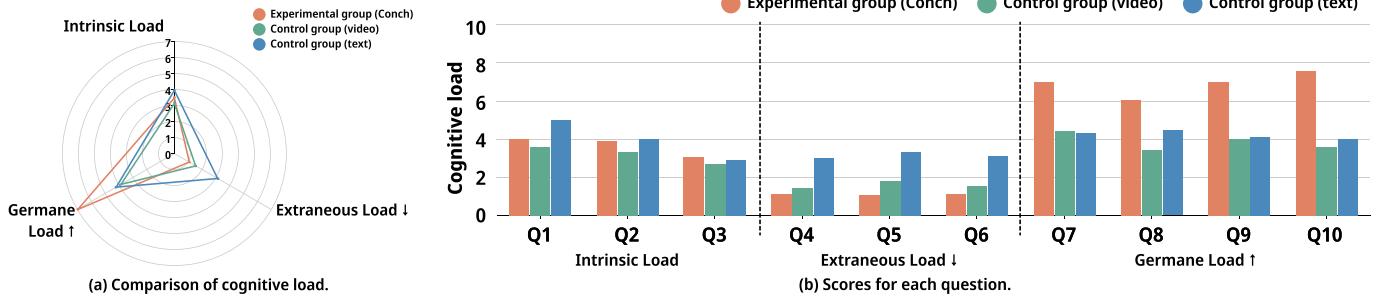


Fig. 10: Cognitive load and performance comparison among the experimental and control groups. (a) Comparison of intrinsic, extraneous, and germane cognitive load among the experimental and control groups. (b) Scores for each question across the experimental and control groups.

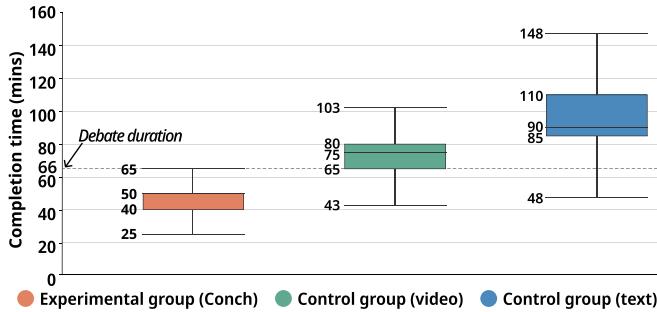


Fig. 11: Total task completion time for the experimental group and two control groups.

what challenging to understand due to the introduction of multiple visual components and novel visualization methods. Regarding Q9, participants indicated that the system offered different benefits based on user experience. For beginners, *Conch* served as a guide to understand the core elements of a debate, such as *clash points* and strategies; for experts, it functioned as a tool to make their analysis more efficient.

In the open-ended feedback, participants suggested that *Conch* was broadly beneficial. For example, experts E3 and E4 noted that the Process View was effective for tracking the evolution of arguments, showing which points were introduced, dropped, or modified over time. E5 added that the Process View (Fig. 1A1) clearly highlighted the main conflicts. The Strategy View (Fig. 1A2) was also widely praised for summarizing common tactics and identifying strategic strengths or weaknesses. Finally, E5 found the Content View (Fig. 1B) useful for filtering information to focus on the most meaningful content. Overall, most of users reported that *Conch* effectively helped them quickly learn main arguments, different strategies, and the clear evolution of debates.

Limitations and improvements. First, *Conch* doesn't show the performance of each single debater. This results from our system's main focus: the presentation of logical and structural debate content, rather than individual debater's performance. In future work, we can develop a debater's view to support speaker-level analysis, including the duration of participation, the number of contributions made, and the efficiency of argumentative output. This would enable more fine-grained assessments of participant performance. Second, current LLMs are still constrained by context window size and reasoning capabilities, which can lead to mistakes in the structured output, causing confusion in user interpretation. In future work, it is worth to fine-tune a LLM with a long context window and strong reasoning abilities to generate all structured analysis results in a single pass.

8 DISCUSSION

In this section, we summarize the lessons we learned during the development and evaluation of *Conch*.

Transferability of structured text analysis. We used the competitive debate to characterize the problem domain. The workflow also

can be applied to legal text interpretation [55], as shown in Fig. 2. This is because structured texts are characterized by logical frameworks, argumentation, and organized reasoning. We also captured the strategies and patterns used by speakers during argumentation. This enriched our understanding of content interaction and further supports the transferability of our approach.

Enhancing sequential textual analysis through hierarchical segmentation. Sequential textual data often lacks clear structure because content is presented incrementally over time, making it challenging to effectively capture relationships and track evolving contexts. While overly general summaries tend to omit essential details critical for user judgment, excessively detailed representations of key points can overwhelm users and obscure content relationships. We propose a hierarchical segmentation approach to balance macro-level summaries and detailed key points. Specifically, our method integrates high-level key points (e.g., clash points) with intermediate-level structural anchors (e.g., disagreements). This hierarchical segmentation framework should be used to enhance the structural organization of sequential textual data. It could facilitate clearer content analysis and improves user comprehension in scenarios such as online discussions, educational forums, and collaborative reviews.

9 CONCLUSION

In this paper, we presented *Conch*, an interactive visualization system designed to assist debaters and coaches in analyzing competitive debates. In our system, LLM-based natural language processing techniques with carefully designed prompts are utilized to structure debate content and identify multi-level semantic information, including clash points, disagreements, viewpoints, and argumentative paths, enabling comprehensive content analysis. The novel visualization designs, including concentric spirals and augmented stacked bar charts, effectively illustrate the evolution of debate content and the strategies adopted by debaters. Through two case studies and a carefully-designed user study, we demonstrated *Conch*'s effectiveness in helping users gain deep insights into various aspects of debate competitions, including content evolution, strategy usage, and interactions.

In future work, we will extend *Conch* to analyze everyday communication, which lack the clear logical rules of structured arguments. We will also incorporate emotion analysis to investigate how emotional factors in these texts influence a reader's acceptance and agreement.

ACKNOWLEDGMENTS

This work was supported in part by the Taihu Lake Innovation Fund for Future Technology, Huazhong University of Science and Technology (HUST), under Grant 2023-B-8; in part by the Fundamental Research Funds for the Central Universities, HUST, under Grant 82400049. The computation is completed in the HPC platform of Huazhong University of Science and Technology. Ran Wang is the corresponding author (rex_wang@hust.edu.cn).

REFERENCES

- [1] R. A. Adejare. Sentence pattern and usage in nigeria's 2015 presidential debate. *Advances in Language and Literary Studies*, 9(4):117–137, 2018. doi: [10.7575/aiac.all.v.9n.4p.117](https://doi.org/10.7575/aiac.all.v.9n.4p.117) 2
- [2] P. D. Agustini Putri. Politeness strategies used in the first trump-clinton presidential debate. *International Journal of Language and Literature*, 6(1):35–43, Jul. 2023. doi: [10.23887/ijll.v6i1.46491](https://doi.org/10.23887/ijll.v6i1.46491) 2
- [3] M. Alharbi and R. S. Laramee. Sos textvis: An extended survey of surveys on text visualization. *Computers*, 8(1), 2019. doi: [10.3390/computers8010017](https://doi.org/10.3390/computers8010017) 2
- [4] M. L. Aryanti, S. Yuliawati, D. Ekawati, N. Darmayanti, et al. The use of turn taking in the 2020 us presidential debate: A conversation analysis study. *Diglosia: Jurnal Kajian Bahasa, Sastra, dan Pengajarannya*, 7(3):511–520, 2024. 1, 2
- [5] C. Athanasopoulos and C. M. Sugino. Clash of the uncivilized: An alternative approach to policy debate. *Contemporary Argumentation and Debate*, 38(1):7, 2023. 2
- [6] M. Będkowski and K. Rogowska. Persuasive strategies in competitive debates: A corpus rhetoric approach. *Res Rhetorica*, 11(4):175–193, 2024. 2
- [7] M. Będkowski and K. J. Rogowska. Systemic means of persuasion and argument evaluation: Insights from the corpus of competitive debates. *Informal Logic*, 44(2):166–207, 2024. 2
- [8] A. Bikakis, G. Flouris, T. Patkos, and D. Plexousakis. Sketching the vision of the web of debates. *Frontiers in Artificial Intelligence*, 6, 2023. doi: [10.3389/frai.2023.1124045](https://doi.org/10.3389/frai.2023.1124045) 2
- [9] D. Braun. I beg to differ: how disagreement is handled in the annotation of legal machine learning data sets. *Artificial Intelligence and Law*, 32(3):839–862, 2024. doi: [10.1007/s10506-023-09369-4](https://doi.org/10.1007/s10506-023-09369-4) 4
- [10] M. Brehmer, B. Lee, B. Bach, N. H. Riche, and T. Munzner. Timelines revisited: A design space and considerations for expressive storytelling. *IEEE Transactions on Visualization and Computer Graphics*, 23(9):2151–2164, 2017. doi: [10.1109/TVCG.2016.2614803](https://doi.org/10.1109/TVCG.2016.2614803) 5
- [11] S. A. Brian, L. P. Salamanca, J. Whitehill, J. S. Reilly, M. S. Bartlett, D. Angus, and J. Wiles. Using recurrence plots to visualize the temporal dynamics of tutor/student interactions. In *2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, pp. 1–2, 2012. doi: [10.1109/DeLrn.2012.6400853](https://doi.org/10.1109/DeLrn.2012.6400853) 2
- [12] G. D. Cantareira, Y. Xing, N. Cole, R. Borgo, and A. Abdul-Rahman. Interactive hierarchical timeline for collaborative text negotiation in historical records. *IEEE Transactions on Visualization and Computer Graphics*, p. 1–12, 2024. doi: [10.1109/tvcg.2024.3376406](https://doi.org/10.1109/tvcg.2024.3376406) 5
- [13] N. Cao, C. Shi, S. Lin, J. Lu, Y.-R. Lin, and C.-Y. Lin. Targetvue: Visual analysis of anomalous user behaviors in online communication systems. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):280–289, 2016. doi: [10.1109/TVCG.2015.2467196](https://doi.org/10.1109/TVCG.2015.2467196) 2
- [14] J. V. Carlis and J. A. Konstan. Interactive visualization of serial periodic data. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology*, UIST '98, 10 pages, p. 29–38. Association for Computing Machinery, New York, NY, USA, 1998. doi: [10.1145/288392.288399](https://doi.org/10.1145/288392.288399) 5
- [15] M. Celepkolu, J. B. Wiggins, A. C. Galdo, and K. E. Boyer. Designing a visualization tool for children to reflect on their collaborative dialogue. *International journal of child-computer interaction*, 27:100232, 2021. 2
- [16] W.-F. Chen, F.-Y. Lin, and L.-W. Ku. WordForce: Visualizing controversial words in debates. In H. Watanabe, ed., *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*, pp. 273–277. The COLING 2016 Organizing Committee, Osaka, Japan, Dec. 2016. 2
- [17] S. Cullen, J. Fan, E. van der Brugge, and A. Elga. Improving analytical reasoning and argument understanding: a quasi-experimental field study of argument visualization. *npj Science of Learning*, 3(1), Dec. 2018. doi: [10.1038/s41539-018-0038-5](https://doi.org/10.1038/s41539-018-0038-5) 2
- [18] T. Dang and V. T. N. 0001. Commodele: Topic modeling using community detection. In *EuroVA@ EuroVis*, pp. 1–5, 2018. 2
- [19] B. Delhomme, F. Taillandier, I. Abi-Zeid, R. Thomopoulos, C. Baudrit, and L. Mora. An interface between natural language and abstract argumentation frameworks for real-time debate analysis. *Decision Support Systems*, 154:113694, 2022. doi: [10.1016/j.dss.2021.113694](https://doi.org/10.1016/j.dss.2021.113694) 2
- [20] S. Di Bartolomeo, A. Pandey, A. Leventidis, D. Saffo, U. H. Syeda, E. Carstensdottir, M. Seif El-Nasr, M. A. Borkin, and C. Dunne. Evaluating the effect of timeline shape on visualization task performance. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20. ACM, Apr. 2020. doi: [10.1145/3313831.3376237](https://doi.org/10.1145/3313831.3376237) 5
- [21] M. El-Assady, R. Sevastjanova, D. Keim, and C. Collins. Threadreconstructor: Modeling reply-chains to untangle conversational text through visual analytics. *Computer Graphics Forum*, 37(3):351–365, June 2018. doi: [10.1111/cgf.13425](https://doi.org/10.1111/cgf.13425) 5
- [22] J. Ericson, J. Murphy, and R. Zeuschner. *The Debater's Guide, Fourth Edition*. Southern Illinois University Press, Carbondale, 2011. 2, 3, 4
- [23] J. L. Fleiss. Measuring nominal scale agreement among many raters., 1971. doi: [10.1037/h0031619](https://doi.org/10.1037/h0031619) 3, 4
- [24] L. A. Goodman. Snowball Sampling. *The Annals of Mathematical Statistics*, 32(1):148 – 170, 1961. doi: [10.1214/aoms/1177705148](https://doi.org/10.1214/aoms/1177705148) 3
- [25] I. Hadiningrum. Speech functions realized by the first speakers of an english debate competition. *J-Lalite: Journal of English Studies*, 1(1):53–67, 2020. doi: [10.20884/1.jes.2020.1.1.2747](https://doi.org/10.20884/1.jes.2020.1.1.2747) 1
- [26] E. Hoque and G. Carenini. Convis: A visual text analytic system for exploring blog conversations. *Computer Graphics Forum*, 33(3):221–230, 2014. doi: [10.1111/cgf.12378](https://doi.org/10.1111/cgf.12378) 2
- [27] S. R. Humayoun, S. Ardalan, R. AlTarawneh, and A. Ebert. TExVis: An Interactive Visual Tool to Explore Twitter Data. In B. Kozlikova, T. Schreck, and T. Wischgoll, eds., *EuroVis 2017 - Short Papers*. The Eurographics Association, 2017. doi: [10.2312/eurovisshort.20171149](https://doi.org/10.2312/eurovisshort.20171149) 2
- [28] F. J. Jennings, B. R. Warner, M. S. McKinney, C. C. Kearney, M. E. Funk, and J. C. Bramlett. Learning from presidential debates: Who learns the most and why? *Communication Studies*, 71(5):896–910, 2020. doi: [10.1080/10510974.2020.1807377](https://doi.org/10.1080/10510974.2020.1807377) 2
- [29] B. Karer, D. Fernández-Prieto, and H. Hagen. The Situation Universe: Visualizing the Semantics of Integrated Data Structures. In B. Kozlikova, T. Schreck, and T. Wischgoll, eds., *EuroVis 2017 - Short Papers*. The Eurographics Association, 2017. doi: [10.2312/eurovisshort.20171153](https://doi.org/10.2312/eurovisshort.20171153) 2
- [30] D. Khatabil, C. Collins, S. Wells, B. Bach, and J. Kennedy. Design and evaluation of visualization techniques to facilitate argument exploration. *Computer Graphics Forum*, 40(6):447–465, 2021. doi: [10.1111/cgf.14389](https://doi.org/10.1111/cgf.14389) 2
- [31] K. Kucher and A. Kerren. Text visualization techniques: Taxonomy, visual survey, and community insights. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 117–121, 2015. doi: [10.1109/PACIFICVIS.2015.7156366](https://doi.org/10.1109/PACIFICVIS.2015.7156366) 2
- [32] K. Kucher, C. Paradis, M. Sahlgren, and A. Kerren. Active learning and visual analytics for stance classification with alva. *ACM Trans. Interact. Intell. Syst.*, 7(3), article no. 14, 31 pages, Oct. 2017. doi: [10.1145/3132169](https://doi.org/10.1145/3132169) 2
- [33] G. Lapesa, A. Blessing, N. Blokker, E. Dayanik, S. Haunss, J. Kuhn, and S. Padó. Analysis of political debates through newspaper reports: Methods and outcomes. *Datenbank Spektrum*, 20(2):143–153, June 2020. doi: [10.1007/s13222-020-00344-w](https://doi.org/10.1007/s13222-020-00344-w) 2
- [34] J. Lawrence and C. Reed. Argument mining: A survey. *Computational Linguistics*, 45(4):765–818, 2020. 2
- [35] B. Lee, N. H. Riche, A. K. Karlson, and S. Carpendale. Sparkclouds: Visualizing trends in tag clouds. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1182–1189, 2010. doi: [10.1109/TVCG.2010.194](https://doi.org/10.1109/TVCG.2010.194) 2
- [36] J. Leppink, F. Paas, C. P. M. Van der Vleuten, T. Van Gog, and J. J. G. Van Merriënboer. Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(4):1058–1072, Apr. 2013. doi: [10.3758/s13428-013-0334-1](https://doi.org/10.3758/s13428-013-0334-1) 8
- [37] S. Liu, Y. Chen, H. Wei, J. Yang, K. Zhou, and S. M. Drucker. Exploring topical lead-lag across corpora. *IEEE Transactions on Knowledge and Data Engineering*, 27(1):115–129, 2015. doi: [10.1109/TKDE.2014.2324581](https://doi.org/10.1109/TKDE.2014.2324581) 2
- [38] H. B. Mann and D. R. Whitney. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1):50 – 60, 1947. doi: [10.1214/aoms/1177730491](https://doi.org/10.1214/aoms/1177730491) 8
- [39] S. Min, X. Lyu, A. Holtzman, M. Artetxe, M. Lewis, H. Hajishirzi, and L. Zettlemoyer. Rethinking the role of demonstrations: What makes in-context learning work? In Y. Goldberg, Z. Kozareva, and Y. Zhang, eds., *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pp. 11048–11064. Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, Dec. 2022. doi: [10.18653/v1/2022.emnlp-main.759](https://doi.org/10.18653/v1/2022.emnlp-main.759) 4
- [40] P. D. I. M. Obied and L. T. W. Al-Khazaali. Identity construction in some selected online debates: A conversational analysis. *International*

- Journal of Applied and Structural Mechanics(IJASM) ISSN: 2799-127X, 4(03):1–11, Apr. 2024.* doi: 10.55529/ijasm.43.1.11 2
- [41] OpenAI. Gpt-4 technical report, 2024. 3
- [42] OpenAI. Openai o1 system card, 2024. 4
- [43] S. Quinn. *Debating in the World Schools Style: A Guide*. International Debate Education Association, 2009. 4
- [44] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever. Robust speech recognition via large-scale weak supervision, 2022. 3
- [45] L. E. Resck, J. R. Ponciano, L. G. Nonato, and J. Poco. Legalvis: Exploring and inferring precedent citations in legal documents. *IEEE Transactions on Visualization and Computer Graphics*, 29(6):3105–3120, 2023. doi: 10.1109/TVCG.2022.3152450 2
- [46] T. Ridnik, D. Kredo, and I. Friedman. Code generation with alphacodium: From prompt engineering to flow engineering, 2024. 4
- [47] J. C. Roberts, P. Butcher, A. Sherlock, and S. Nason. Explanatory journeys: Visualising to understand and explain administrative justice paths of redress. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):518–528, 2022. doi: 10.1109/TVCG.2021.3114818 2
- [48] R. Ruiz-Dolz, S. Heras, and A. Garcia. Automatic debate evaluation with argumentation semantics and natural language argument graph networks. In H. Bouamor, J. Pino, and K. Bali, eds., *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 6030–6040. Association for Computational Linguistics, Singapore, Dec. 2023. doi: 10.18653/v1/2023.emnlp-main.368 2
- [49] K. M. Sari and Y. Fatmasari. Error analysis on english diphthongs pronounced by indonesian students in english debate competition posted in youtube: Array. *Jurnal Dialetika Program Studi Pendidikan Bahasa Inggris*, 7(1):38–55, 2019. doi: 10.58436/jdpbi.v7i1.442 1
- [50] Y. Seki. Visualization for analyzing usage status from dialogue systems. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 1701–1706. IEEE, 2020. 2
- [51] M. Serrano, A. Roudaut, and P. Irani. Investigating text legibility on non-rectangular displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI ’16, 11 pages, p. 498–508. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2858036.2858057 5
- [52] L. South, M. Schwab, N. Beauchamp, L. Wang, J. Wihbey, and M. A. Borkin. Debatevis: Visualizing political debates for non-expert users. In *2020 IEEE Visualization Conference (VIS)*, pp. 241–245. IEEE, Salt Lake City, UT, USA, Oct. 2020. doi: 10.1109/VIS47514.2020.00055 2
- [53] F. Stoffel, W. Jentner, M. Behrisch, J. Fuchs, and D. Keim. Interactive ambiguity resolution of named entities in fictional literature. *Computer Graphics Forum*, 36(3):189–200, 2017. doi: 10.1111/cgf.13179 2
- [54] M. Supardi and F. Sayogie. The logical pattern of argument: A case study of national university debating championship. *Elsya : Journal of English Language Studies*, 4(1):65–76, Apr. 2022. doi: 10.31849/elsya.v4i1.8466 1, 2, 7
- [55] The President and Fellows of Harvard University. Caselaw access project, 2024. Last accessed March 30, 2025. 9
- [56] M. Wang, Y. Liu, X. Zhang, S. Li, Y. Huang, C. Zhang, D. Wang, S. Feng, and J. Li. Langgpt: Rethinking structured reusable prompt design framework for llms from the programming language, 2024. 4
- [57] R. Wang, K. Du, Q. Chen, Y. Zhao, M. Tang, H. Tao, S. Wang, Y. Li, and Y. Wang. Rumorlens: Interactive analysis and validation of suspected rumors on social media. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pp. 1–7, 2022. 5
- [58] X. Wang, H. Zeng, Y. Wang, A. Wu, Z. Sun, X. Ma, and H. Qu. Voicecoach: Interactive evidence-based training for voice modulation skills in public speaking. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, vol. 26 of *CHI ’20*, p. 1–12. Association for Computing Machinery, New York, NY, USA, Apr. 2020. doi: 10.1145/3313831.3376726 2
- [59] K. Watson, S. S. Sohn, S. Schriber, M. Gross, C. M. Muniz, and M. Kapadia. Storyprint: an interactive visualization of stories. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, IUI ’19. ACM, Mar. 2019. doi: 10.1145/3301275.3302302 2
- [60] M. Weber, M. Alexa, and W. Muller. Visualizing time-series on spirals. In *IEEE Symposium on Information Visualization, 2001. INFOVIS 2001.*, pp. 7–13, 2001. doi: 10.1109/INFVIS.2001.963273 5
- [61] Z. You. *DEBATING principles and practice of academic debate*. Yeh Yeh Book Gallery, 2004. 2, 4
- [62] H. Zeng, X. Wang, A. Wu, Y. Wang, Q. Li, A. Endert, and H. Qu. Emoco: Visual analysis of emotion coherence in presentation videos. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):927–937, 2020. doi: 10.1109/TVCG.2019.2934656 2
- [63] J. Zhang, R. Kumar, S. Ravi, and C. Danescu-Niculescu-Mizil. Conversational flow in oxford-style debates. *arXiv preprint arXiv:1604.03114*, 2016. 2
- [64] Y. Zhang, X. Yang, S. Feng, D. Wang, Y. Zhang, and K. Song. Can llms beat humans in debating? a dynamic multi-agent framework for competitive debate. *arXiv preprint arXiv:2408.04472*, 2024. 1, 2
- [65] H. Zhou, N. Song, W. Chang, and X. Wang. Linking the thoughts within scientific papers: Construction and visualization of argumentation graph. *Proceedings of the Association for Information Science and Technology*, 56(1):757–759, Jan. 2019. doi: 10.1002/pra2.205 2