From Alignment to Reason: Multi-Agent Debate for Tactical Badminton Video Retrieval

Supplementary Material

A1. Structured Game Log Specification

This appendix presents a concrete example of the structured game log used in our system. The sample in Figure A1 illustrates how rally events, spatial information, and pose cues are organized in practice. The example reflects the actual format consumed by the debate agents, and all tactical claims generated by the system must be verifiable against entries in this log.

A2. MADR Prompt Configuration

This appendix provides the detailed configuration used to instantiate the MADR framework, including the agent personas and the prompt structures that govern each stage of the debate process. We describe the configuration details below. The prompts associated with each step are shown in Figures A2-A7.

- **Analyst Personas ($\mathcal{P}$):** We embed specific 'coach_focus' definitions into the prompts to enforce adversarial viewpoints:
  - *Offense Analyst:* Specializes in judging offensive strategies, timing of attacks, and creation of scoring opportunities in badminton matches.
  - *Defense Analyst:* Specializes in judging defensive strategies, timing of defense, and analysis of shot-type data in badminton matches.

  Both agents are mandated to support their arguments with logical articulation grounded in the game log evidence.
- **Tone Decay Function (ϕ):** To simulate the evolution of a professional debate, the function $\phi(i)$ modulates the `tone_description` in the prompts via a contentiousness parameter c:
  - *Critical Phase ($c \geq 0.8$):* Adopt a critical, confrontational attitude. Focus on identifying fundamental flaws, logical contradictions, or overlooked evidence in the opponent's argument. The goal is to refute the opponent.
  - *Balanced Phase ($0.4 \leq c < 0.8$):* Adopt a balanced, exploratory approach. While questioning the opponent, also acknowledge the valid points in their perspective and try to offer a more comprehensive alternative explanation. The goal is to deepen the discussion.
  - *Cooperative Phase ($c < 0.4$):* Adopt a cooperative, consensus-seeking approach. The focus is not on refutation but on finding possibilities to integrate both viewpoints to build a more complete conclusion together. The goal is to integrate and co-construct.

Algorithm 1 provides a formalized version of the workflow. While the main paper describes the process conceptually, the algorithm here presents the same reasoning pipeline in explicit procedural form.

A3. Retrieval and Re-ranking Algorithm

This appendix details the implementation of the online retrieval inference pipeline (Algorithm 2). We implement a **Two-Stage Retrieval Architecture**. The inference process is bifurcated into two distinct phases:

- **Stage 1: Coarse Retrieval (ANN Search):** An exhaustive vector similarity search is performed over the entire index ($\mathcal{I}$) using approximate nearest neighbor (ANN) search to generate a high-recall candidate ranking, $\mathcal{L}_{\text{global}}$.
- **Stage 2: Semantic Re-ranking:** To optimize query latency while maintaining high precision, the computationally intensive cross-encoder re-ranker ($\mathcal{M}$) is applied only to the top-K subset ($\mathcal{L}_{\text{head}}$). The remaining candidates ($\mathcal{L}_{\text{tail}}$) are retained to preserve recall but are not re-scored.

The final ranking is determined by concatenating the re-scored head with the preserved tail. This design ensures that top-tier results benefit from fine-grained semantic matching without discarding the broader retrieval context or incurring prohibitive latency.

A4. Dataset Statistics and Query Taxonomy

To rigorously evaluate the retrieval capabilities of MADR, we constructed a comprehensive benchmark dataset derived from 40 professional badminton matches in the ShuttleSet dataset [42], comprising a total of 2,057 individual rally videos. Through the generation and filtering process described in Section 4.1 of the main paper, we curated a final set of 230 distinct natural-language queries. These queries naturally fall into three categories based on their semantic complexity:

- **Category 1: Factual Queries.** These queries focus on explicit attributes and low-level metadata, such as shot counts, specific stroke types, or win/loss outcomes. They require accurate detection but minimal reasoning.
- **Category 2: Relational Queries.** These queries require understanding the temporal sequence or causal relationships between consecutive shots, such as *"Sequence A leads to Outcome B"*.
- **Category 3: Strategic Reasoning Queries.** These queries involve abstract tactical concepts like "pressure," "exploitation," or "recovery." They demand high-level interpretation to infer the *implicit* intent behind player actions.

Table A1 provides representative examples for each category.

Game Log Source Data Example (Rally 1)

```
### 1. Rally 1 Context and Result ###
- Point Winner: Bottom-court Player
- Total Shots in Rally: 3
- Reason for Win/Loss: Bottom-court Player Winning shot.
### 2. Court and Player Definitions ###
- Coordinate System: Top-down view coordinates after Homography transformation.
- Singles Court Boundary Range:
    - x-axis (left-right): [47.62, 302.37]
    - y-axis (top-bottom): [149.99, 810]
- Net Y-coordinate: 480.0
- Player Information:
    - Bottom-court Player (Hits odd-numbered shots: 1, 3, 5, ...)
    - Top-court Player (Hits even-numbered shots: 2, 4, 6, ...)
- Current Score: Top-court Player 0-0 Bottom-court Player
### 3. Shot Type Description ###
- short service: Service that goes over the top of the net...
- net shot: Soft shot that places the shuttle close to the net.
- lob: Defensive shot typically performed from the front...
### 4. Detailed Shot-by-Shot Record ###
<shot index="1">
   * Video Time: 00:27:23
   * Hitter: Bottom-court Player
   * Hitting Position: Bottom-court Player's Center Mid court (182, 631)
   * Shot Action: short service [Features: Backhand, Hit Height: Below the net]
   * Hitting Pose Description: The person is in a squatting position with both knees
bent and feet shoulder-width apart.
   * Shuttle Trajectory: From Hitting Position [(182, 631)] to Landing Point [Top-court
Player's Center Mid court (130, 365)]
   * Opponent at that time: Top-court Player (located at [(143, 280)])
</shot>
<shot index="2">
   * Video Time: 00:27:23 (Interval: 0.68s)
   * Hitter: Top-court Player
   * Hitting Position: Top-court Player's Center Front court (114, 386)
   * Shot Action: net shot [Features: Forehand, Hit Height: Below the net]
   * Hitting Pose Description: The person is in a lunge position with the left leg
forward, bent at a right angle...
   * Shuttle Trajectory: From Hitting Point [(130, 365)] to Landing Point [Bottom-court
Player's Left Front court (258, 510)]
   * Opponent at that time: Bottom-court Player (located at [(177, 622)])
</shot>
<shot index="3">
   * Video Time: 00:27:24 (Interval: 0.64s)
   * Hitter: Bottom-court Player
   * Hitting Position: Bottom-court Player's Center Front court (224, 560)
   * Shot Action: lob [Features: Forehand, Hit Height: Below the net]
   * Hitting Pose Description: The person is leaning back with the right leg raised and
bent, and the left leg extended...
   * Shuttle Trajectory: From Hitting Point [(258, 510)] to Landing Point [Top-court
Player's Right Rear court (76, 169)]
   * Opponent at that time: Top-court Player (located at [(161, 333)])
   * Opponent's pose attempting to return: The person is leaning forward with their
right arm raised...
</shot>
```

Figure A1. **An example of the structured game log.**

Step 1: Independent Analysis Prompt

```
### Persona Definition: Specialist Tactician ###
You are an expert-level sports tactician, identified as {coach_name}. Your analytical
focus is strictly on {coach_focus}.

### Analytical Ontology: Tri-Layer Causal Model ###
Your entire analysis must be structured through this shared causal model:
1. Spatial Dynamics: Encompasses court positioning, tempo of movement, and overall
   spatio-temporal control.
2. Tactical Sequencing: Pertains to the selection, combination, and strategic intent
   behind shot sequences.
3. Technical Execution: Refers to the biomechanical quality and precision of a player's
   shot execution.

### Core Mandate: Propose and Defend a Definitive Opening Thesis ###
Your primary function is to establish the opening argument for the tactical debate.
Based on the provided Raw Match Data, you must construct a definitive opening thesis.
This thesis must be anchored in your specific area of expertise ({coach_focus}),
providing a primary causal explanation for the rally's outcome.

### Governing Axioms ###
1. Axiom of Thesis Mandate: Your opening_thesis must be a bold, specific, and
   verifiable claim that synthesizes factors from the three analytical layers to
   pinpoint the rally's primary causal driver.
2. Axiom of Causal Primacy: You must not only identify a primary cause but also
   construct a compelling argument in your causal_reasoning that explains why this
   factor is more influential than others. This is the core of your defense.
3. Axiom of Strict Evidentiary Grounding: Every assertion within your causal_reasoning
   must be explicitly linked to specific data points from your evidence list.
4. Axiom of Professional Lexicon: Your analysis must be articulated using precise and
   professional badminton terminology.

### Output Schema ###
Please output your analysis in the following JSON format. Your entire output must be a
single, structurally complete JSON object.
{
   "opening_thesis": "(Your bold and specific thesis statement, clearly stating the
primary cause of the rally's outcome from your expert perspective.)",
   "evidence": ["(A curated list of specific data points from the Raw Match Data that
directly support your thesis.)", "..."],
   "causal_reasoning": "(A compelling narrative that logically defends your thesis.
Methodically use the Tri-Layer Ontology to connect your evidence to your thesis,
building an irrefutable argument for the causal primacy of your chosen factors.)"
}
```

Figure A2. **Prompt for Step 1: Initial Independent Analysis.** This prompt instantiates two separate agents (Offense/Defense) to form initial hypotheses based on the Tri-Layer Causal Model.

Step 2: Debate and Refinement Prompt

Persona Definition: Specialist Tactician
You are an expert-level sports tactician, identified as {coach_name}. Your analytical focus is strictly on {coach_focus}.

Analytical Ontology: Tri-Layer Causal Model
Your entire analysis must be structured through this shared causal model:
1. **Spatial Dynamics**: Encompasses court positioning, tempo of movement, and overall spatio-temporal control.
2. **Tactical Sequencing**: Pertains to the selection, combination, and strategic intent behind shot sequences.
3. **Technical Execution**: Refers to the biomechanical quality and precision of a player's shot execution.

Core Mandate: Advance the Debate Through Expert Reframing
Your primary objective is to advance the tactical debate by critically evaluating your opponent's argument through the specific lens of your expertise ({coach_focus}). You must formulate a new, more insightful causal argument based on all available information. **Golden Rule**: Never repeat prior arguments; always add new value.

Input Context
- Debate History: {full_history}
- Opponent Argument: {opponent_summary}

Governing Axioms
1. **Axiom of Causal Integrity**: All arguments must be structured as coherent causal models that adhere to the Analytical Ontology.
2. **Axiom of Causal Reframing**: Your goal is to reframe the rally's causal story. Propose a more precise explanation by challenging causal primacy or revealing a deeper cause.
3. **Axiom of Attributional Integrity**: You MUST internally verify player-shot attribution using the odd/even sequence.
4. **Axiom of Stylistic Adaptation**: You must adapt your tone based on the directive: {tone_description}.

Output Schema
Strictly adhere to the following JSON format. Your entire output must be a single, structurally complete JSON object.

```
{
    "rejoinders": [
      {
        "counter_argument": "(Your structured counter-argument...)",
        "supporting_evidence": ["Evidence 1: ...", "Evidence 2: ..."]
      }
    ],
    "new_insights": [
      {
        "insight": "(Optional but encouraged: Present a novel insight...)",
        "evidence": ["(Specific data evidence...)"]
      }
    ]
}
```

Figure A3. **Prompt for Step 2: Dialectic Debate.** Agents use this prompt to critique the opponent's reasoning. The 'tone_description' parameter modulates the contentiousness of the debate (Critical → Balanced → Cooperative).

Step 3: Narrative Synthesis Prompt

Persona Definition: Principal Analyst
You are a Principal Analyst for an elite badminton coaching consortium. You are tasked with producing the definitive post-rally analysis. Your function is not merely to summarize but to synthesize disparate data streams (qualitative expert debate and quantitative match data) into a singular, coherent causal explanation. Your ultimate objective is to elucidate the definitive causal chain that dictated the rally's outcome.

Analytical Ontology: Tri-Layer Causal Model
All analysis must be structured through the following three-layer causal model. These layers are non-exclusive and inter-dependent.
1. **Spatial Dynamics**: Encompasses court positioning, movement tempo, and the establishment or degradation of spatio-temporal control.
2. **Tactical Sequencing**: Pertains to the selection, combination, and strategic intent behind shot sequences.
3. **Technical Execution**: Refers to the biomechanical quality of shot execution, which dictates precision, consistency, and efficacy.

Input Context: Coaching Team Debate History
{full_history}

Governing Axioms & Methodological Constraints
1. **Axiom of Factual Supremacy & Data-Driven Attribution**: This is your highest priority. Your report's factual basis MUST come directly from the Pre-computed Rally Analysis.
2. **Axiom of Epistemological Synthesis**: Critically evaluate the Coaching Team Debate History to extract causal reasoning and tactical insights, but NOT for factual claims. Assign greater evidentiary weight to arguments from the debate that align with the quantitative Pre-computed Rally Analysis.
3. **Axiom of Agent-Centric Framing**: Player-specific analysis must be strictly framed from the perspective of the player being analyzed, adhering to the validated shot list from the Key Action Index.
4. **Axiom of Causal-Narrative Coherence**: The report must construct a compelling and logically sound narrative arc (Setup, Development, Turning Point, Resolution), not a chronological list of events.
5. **Axiom of Explanatory Depth & Causal Distillation**: Identify the "causal fulcrum"|the critical 2-3 shot sequence that irreversibly determined the outcome. The win_loss_attribution must deliver a single, distilled, and authoritative verdict on the primary cause.
6. **Axiom of Formal & Structural Integrity**: Employ precise, professional badminton terminology. The final JSON output must strictly conform to the specified schema with no extraneous elements.

Output Schema
Generate a single, structurally complete JSON object for players {player_a_name} and {player_b_name}. Adhere strictly to the key names and value types specified below.

```
{
   "rally_narrative": "(A fluent, chronological narration of the rally's key events,
framed by the Causal-Narrative Coherence axiom.)",
   "{player_a_key}": "(An insightful analysis of {player_a_name}'s tactical decisions and
executed shots, adhering to the Agent-Centric Framing axiom.)",
   "{player_b_key}": "(An insightful analysis of {player_b_name}'s tactical decisions and
executed shots, adhering to the Agent-Centric Framing axiom.)",
   "turning_point": "(Identify and explain the 'causal fulcrum'|the critical 2-3 shot
sequence that irreversibly determined the rally's outcome.)",
   "win_loss_attribution": "(A clear, decisive, and singular verdict on the core causal
factor for the rally's result, as distilled from your synthesis.)",
}
```

Figure A4. **Prompt for Step 3: Rally Narrative Generation.** This agent acts as the Tactic Summarizer, resolving conflicts from the debate to produce the initial report $\mathbf{R}_{v1}$.

Step 4: Analytical Peer Review Prompt

Persona Definition: Elite Tactical Coach
You are {coach_name}, an elite-level badminton tactician and a senior member of a national coaching team. Your specialty is {coach_focus}. You are reviewing a draft report that has already passed a rigorous, automated fact-checking process.

Input Context: Review Dossier
- Fact-Checked Tactical Report for Review: {final_report_json_str}

Mandated Review Protocol: Strategic Insight Elevation
Primary Mission: The "So What?" Test
- Your input, the draft report, is considered factually accurate. Your sole mission is to assess and elevate its strategic depth from your specialist perspective of {coach_focus}.
- A report that only states 'what' happened is insufficient; it must explain 'why' it happened and 'so what' it means for future strategy.
- Evaluate the report's causal reasoning using the Tri-Layer Causal Model (Spatial Dynamics, Tactical Sequencing, Technical Execution).
- Your objective is to challenge superficial analysis. Identify where the report fails to connect the dots, misses underlying tactical patterns, or fails to explain the true cause of the outcome.
- Formulate your critique as a series of constructive, actionable enhancement_directives. If the report's analysis already meets the highest standards of strategic insight, you must return an empty array [].

Output Schema: Coach's Enhancement Directives
Your entire output must be a single, structurally perfect JSON object. It should contain only directives for analytical enhancement.

```
{
   "approval_status": "(A string: 'Needs Revision' if any enhancement_directives
exist, otherwise 'Approved'.)",
   "justification": {
      "enhancement_directives": [
         {
            "type": "ANALYTICAL_ENHANCEMENT",
            "location": "(The specific JSON field requiring deeper analysis.)",
            "instruction": "(A constructive rewrite guide. Example: 'In the
win_loss_attribution, the current text is too generic. Rewrite it to incorporate this
specific causal chain: ...')"
         }
      ]
   }
}
```

Figure A5. **Prompt for Step 4: Peer Review.** Used by the two domain-specialized agents to evaluate $\mathbf{R}_{v1}$. They provide structured "Enhancement Directives" if the analysis lacks depth.

Step 5: Integration and Revision Prompt

```
### Persona Definition: Chief Analyst & Final Synthesizer ###
You are the Lead Editor and Chief Analyst. Your mandate is to produce the definitive,
final version of the analytical report by precisely executing and intelligently
synthesizing directives from two specialist coaches. Your goal is to create a final
report that is both factually infallible and strategically profound, using all available
information to achieve this.

### Input Context: Synthesis Dossier ###
    ### Draft Report (Version 1.0) ###
    {original_report_json_str}
    ### Feedback from {coach_a_name} ({coach_a_focus}) ###
    {coach_a_review_json_str}
    ### Feedback from {coach_b_name} ({coach_b_focus}) ###
    {coach_b_review_json_str}

### Mandated Cognitive Workflow ###
1. Phase 1: Factual Correction (Highest Priority): First, meticulously apply all
   correction_directives. Use the instruction as your primary command, but you MUST
   cross-reference it with the Raw Match Data available in your context to ensure your
   correction is perfectly aligned with the ground truth. This provides a layer of
   redundancy and guarantees factual accuracy.
2. Phase 2: Strategic Synthesis with Context: With the fact-checked report as your new
   baseline, integrate the enhancement_directives. You MUST consult the Raw Match Data
   and Pre-computed Rally Analysis in your context to fully understand the reasoning
   behind each enhancement directive. This will allow you to skillfully weave the
   provided logic into the existing text, creating a richer, more coherent narrative.
3. Phase 3: Coherence Polish: After integrating all directives, perform a final
   read-through to ensure the report is fluent, consistent, and sounds like a single,
   authoritative author.

### Governing Axioms ###
1. Axiom of Ground-Truth Alignment: Your highest priority is that the final report must
   be in perfect alignment with the Raw Match Data provided in your context. Use it as
   your ultimate arbiter when applying any and all directives.
2. Axiom of Intelligent Integration: While the coaches' directives are your primary
   guide, you are expected to use the contextual Raw Match Data to make your synthesis
   more intelligent and context-aware. Your role is to execute directives, but with the
   full context of the original evidence.
3. Axiom of Surgicality: Your final JSON output must be precise. It must only contain
   the keys for the fields that have been modified from the base report. Unchanged
   fields must be omitted.

### Output Schema ###
Directly output a JSON object containing only the modified fields. Your entire output
must be a single, structurally complete JSON object.
```

Figure A6. **Prompt for Step 5: Revision.** This agent incorporates the peer reviews into a polished report $\mathbf{R}_{v2}$, ensuring new insights are integrated without violating ground-truth data.

Step 6: Factual Verification Prompt

```
### Persona Definition: Factual Auditor & Copy Editor ###
You are a factual verification agent that never alters correct information. Your
reasoning process has two stages:
1. Build an explicit immutable mapping (truth_map) from the ground truth.
2. Compare each shot in the text to verify factual correctness, modifying only when it
   strictly conflicts with the truth_map.

### Verification Protocol ###
Stage 1 --- Build the Truth Map
From the ground truth data, extract every mapping in this format:
[ "shot_number": number, "player": "Top-court Player" | "Bottom-court Player" ]
and construct a JSON object called truth_map. This truth_map is final, immutable, and
authoritative.

Stage 2 --- Sentence-by-Sentence Verification
For every sentence in <text_to_verify>:
1. Detect any mention of ``Shot N''.
2. If ``Shot N'' appears:
   • Look up truth_map[N].
   • If the same player name already appears → leave it unchanged.
   • If a different player name appears → replace only that player name, keep the rest
     identical.
   • If Shot N not in truth_map → skip (do not change).
3. If no mismatches are found → output "is_factually_correct": true.
4. If only 1--2 replacements are made → output Format A with "corrected_text".
5. If 3+ or complex changes are required → output Format B with "corrections_proposed".

Stage 3 --- Sanity Check Before Output
Before returning any correction:
• For each corrected shot number, double-check that the new player name == truth_map[N].
• If yes → keep correction.
• If already matched before → revert correction (do not ``fix'' what is already true).

### Output Schema: Intelligent Correction Report ###
Choose only one format according to the Axiom of Efficiency.
Format A -- Direct Correction
{
    "is_factually_correct": false,
    "corrected_text": "(Full corrected text with only erroneous parts fixed.)"
}
Format B -- Diff Correction
{
    "is_factually_correct": false,
    "corrections_proposed": [
      {
        "original_segment": "(The minimal text fragment containing the factual
error.)",
        "corrected_segment": "(The corrected version of that fragment.)"
      }
    ]
}
If all player attributions are correct
{
    "is_factually_correct": true
}
```

Figure A7. **Prompt for Step 6: Final Verification.** This strict algorithmic auditor ensures $\mathbf{R}_{final}$ matches the GameLog events, correcting any hallucinations generated during the reasoning process.

Algorithm 1 Multi-Agent Dialectic Reasoning (MADR)

Input:
- **Resources:** Rally Log D_r; Pre-trained LLM $\mathcal{L}$.
- **System Config** $\mathbb{C} = \{\mathcal{P}, \phi, N\}$: Personas $\mathcal{P}$, Tone Decay ϕ, and Max Rounds N.

Output: Final Analysis Report (R_{final})

function GENERATENARRATIVE($D_r, \mathbb{C}, \mathcal{L}$)
  $G \leftarrow \text{Preprocess}(D_r)$ ▷ Transform rally data into game log
  $H \leftarrow \emptyset$ ▷ Initialize Debate History

Step 1: Independent Analyses Based on the Game Log
  for all $k \in \mathbb{C}.\mathcal{P}$ **do**
    $H \leftarrow H \cup \{\mathcal{L}_{\text{Analyze}}(G, k)\}$
  end for

Step 2: Divergent Debate and Analytical Refinement
  for $i \leftarrow 1$ to $\mathbb{C}.N$ **do**
    $c \leftarrow \mathbb{C}.\phi(i)$ ▷ Update tone_description based on round i
    for all $k \in \mathbb{C}.\mathcal{P}$ **do**
      $H \leftarrow H \cup \{\mathcal{L}_{\text{Debate}}(H, G, k, c)\}$
    end for
  end for

Step 3: Rally Narrative Generation
  $R_{v1} \leftarrow \mathcal{L}_{\text{Synthesize}}(H, G)$ ▷ Generate coherent report $\mathbf{R}_{v1}$

Step 4: Analytical Peer Review
  $R_{review} \leftarrow \{\mathcal{L}_{\text{Review}}(R_{v1}, k) \mid k \in \mathbb{C}.\mathcal{P}\}$ ▷ Reviewers provide feedback $\mathbf{R}_{review}$

Step 5: Integration and Revision
  if $R_{review} \neq \emptyset$ **then**
    $R_{v2} \leftarrow \mathcal{L}_{\text{Revise}}(R_{v1}, R_{review})$ ▷ Generate refined report $\mathbf{R}_{v2}$
  else
    $R_{v2} \leftarrow R_{v1}$
  end if

Step 6: Rally Narrative Verification
  $R_{final} \leftarrow \mathcal{L}_{\text{VerifyAndCorrect}}(R_{v2}, G)$ ▷ Yield verified version $\mathbf{R}_{final}$
  return R_{final}
end function

Algorithm 2 Two-Stage Retrieval Pipeline

Input: Query (Q), Index ($\mathcal{I}$), Embedder ($\mathcal{E}$), Reranker ($\mathcal{M}$), Re-rank Depth (K)
Output: Final Ranked List (L_{final})

function RETRIEVEANDRERANK($Q, \mathcal{I}, \mathcal{E}, \mathcal{M}, K$)
Stage 1: Coarse Retrieval (ANN Search)
    $v_Q \leftarrow \mathcal{E}(Q)$
    $\mathcal{L}_{\text{global}} \leftarrow \text{VectorSearch}(\mathcal{I}, v_Q, \text{limit} = \infty)$
Candidate Partitioning
    $\mathcal{L}_{\text{head}} \leftarrow \mathcal{L}_{\text{global}}[1 \dots K]$
    $\mathcal{L}_{\text{tail}} \leftarrow \mathcal{L}_{\text{global}}[K+1 \dots |\mathcal{L}_{\text{global}}|]$
Stage 2: Semantic Re-ranking (Head Only)
    $L_{\text{refined}} \leftarrow \emptyset$
    for all id in $\mathcal{L}_{\text{head}}$ **do**
        $T_{\text{doc}} \leftarrow \text{FetchNarrative}(id)$
        $s \leftarrow \mathcal{M}(Q, T_{\text{doc}})$
        Append (id, s) **to** L_{refined}
    end for
    $L'_{\text{head}} \leftarrow \text{Sort}(L_{\text{refined}}, \text{key} = s, \text{order} = \text{descending})$
Rank Aggregation
    $L_{\text{final}} \leftarrow L'_{\text{head}} \oplus \mathcal{L}_{\text{tail}}$
    return L_{final}
end function

Table A1. **Benchmark Query Taxonomy.** The dataset moves beyond simple keyword matching, requiring the system to understand complex temporal relations and strategic intents.

Query Category	Example Generated Query
Category 1: Factual Queries	
Rally Attribute	Show rallies with fewer than 5 shots.
Simple Action	Show rallies that include a drive shot.
Specific Event	Show rallies where a drop shot was the final stroke and resulted in a lost point.
Category 2: Relational Queries	
Temporal Sequence	Find rallies featuring a backhand push to the center rear court, followed by a clear to the deep backcourt, leading to a wrist smash to the mid-court.
Causal Relation	How does a deep smash from the center rear court force a rushed net return with hit height below the net, leading to a net fault?
Category 3: Strategic Reasoning Queries	
Abstract Intent	Find strategic principles where a player exploits an opponent's technical weakness through targeted pressure to induce critical errors.
Performance Analysis	Instances of winning a point by using a lob to push the opponent deep, followed by a well-timed drop shot to exploit the opponent's compromised recovery position.

A5. Formal Definitions of Evaluation Metrics

This section provides the formal mathematical definitions for the retrieval metrics reported in Section 4.3. We begin by introducing the notation used throughout the definitions. Let Q be the set of test queries. For a given query $q_i \in Q$, let $\mathcal{G}_i$ denote the set of ground-truth relevant videos, and let $L_i = [r_{i,1}, r_{i,2}, \ldots, r_{i,N}]$ be the ranked list of retrieved items sorted by similarity score.

A5.1. Hit Rate (H@K) and Recall (R@K)

These metrics evaluate the system's ability to retrieve relevant items within the top-K positions.

Hit Rate at K (H@K) measures the proportion of queries for which *at least one* relevant item appears in the top-K results. It acts as a binary success indicator for the user:

$$\text{H@K} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \mathbb{I}\left(\exists j \leq K : r_{i,j} \in \mathcal{G}_i\right)$$

where $\mathbb{I}(\cdot)$ is the indicator function, evaluating to 1 if the condition is true and 0 otherwise.

Recall at K (R@K) measures the proportion of total relevant items found within the top-K results, averaged across all queries:

$$\text{R@K} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{|\{r_{i,j} \in \mathcal{G}_i \mid j \leq K\}|}{|\mathcal{G}_i|}$$

A5.2. Mean Average Precision (MAP)

MAP provides a comprehensive measure of retrieval quality by considering the rank position of all relevant items. It rewards models that place relevant items higher in the list. MAP is the mean of the Average Precision (AP) scores over all queries:

$$\text{MAP} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \text{AP}_i$$

where the Average Precision for the i-th query is computed as:

$$\text{AP}_i = \frac{1}{|\mathcal{G}_i|} \sum_{k=1}^{N} P_i(k) \times \mathbb{I}(r_{i,k} \in \mathcal{G}_i)$$

Here, $P_i(k)$ is the precision at cut-off k for query i:

$$P_i(k) = \frac{|\{r_{i,j} \in \mathcal{G}_i \mid j \leq k\}|}{k}$$

A5.3. Mean Rank (MnR) and Median Rank (MdR)

These metrics quantify the position of the *first* relevant result, reflecting the user effort required to find a correct match. Let rank_i be the rank position of the first relevant item for query q_i:

$$\text{rank}_i = \min\{j \mid r_{i,j} \in \mathcal{G}_i\}$$

To ensure robustness against failed queries where no relevant video is retrieved within the candidate limit, we compute these metrics over the subset of successful queries, denoted as Q_{hit}:

Mean Rank (MnR) is the arithmetic mean of these first-hit ranks:

$$\text{MnR} = \frac{1}{|Q_{\text{hit}}|} \sum_{i \in Q_{\text{hit}}} \text{rank}_i$$

Median Rank (MdR) is the median value of the first-hit ranks. MdR is particularly valuable as it offers resilience against outlier queries that may have exceptionally poor ranking performance:

$$\text{MdR} = \text{Median}(\{\text{rank}_i \mid i \in Q_{\text{hit}}\})$$

Lower values indicate better performance for both MnR and MdR.

A6. Robustness to Perception Noise

To address the inherent imperfections of game logs extracted by real-world computer vision models, we conducted a robustness study using noisy inputs. We simulated perception noise through two mechanisms: (1) 15% of shot types were randomly replaced with either visually similar types or "unknown" tokens based on a domain-aware confusion map, and (2) player and shuttle coordinates were perturbed with Gaussian noise ($\sigma = 0.25$ meters) to reflect physical tracking inaccuracies. As shown in Table A2, even when operating on these noisy logs, our MADR (Flash) maintains a competitive H@1 of 29.57%, significantly outperforming the best VLM baseline (RzenEmbed-v2 at 5.65%). Furthermore, by reporting the mean and standard deviation across three independent runs for the Qwen3-8B configurations, we demonstrate that our multi-agent consensus mechanism provides highly stable and consistent results despite input perturbations.

Table A2. Performance comparison demonstrating robustness to perception noise and multi-run stability. All videos/narratives are retrieved using the Qwen3-Embedding-4B model with re-ranking. Results for Qwen3-8B are averaged over 3 runs.

Methods	**H@1 (%)** ↑	**H@5 (%)** ↑	**H@10 (%)** ↑
Multimodal Embeddings			
VLM2Vec (7B)	4.35	10.00	17.39
RzenEmbed-v2 (7B)	5.65	12.17	18.26
Narratives generated using Qwen3-8B (Average of 3 runs)			
Single Agent (Clean Log)	23.33±1.10	46.23±1.76	56.09±1.15
MADR (Clean Log)	28.26±1.15	48.69±3.28	57.10±5.04
MADR (Noisy Log)	21.45±0.25	39.85±1.52	50.43±1.56
Narratives generated using Gemini-2.5 Models			
Single Agent (Flash, Clean Log)	29.13	52.17	64.78
Single Agent (Pro, Clean Log)	28.70	57.39	66.09
Single Agent (Pro, Video)	9.13	32.17	42.61
MADR (Flash, Clean Log)	55.65	77.39	83.04
MADR (Flash, Noisy Log)	29.57	55.22	65.22
Narratives generated using GPT-5-mini			
MADR (Clean Log)	31.74	53.04	62.17
MADR (Noisy Log)	23.48	46.96	56.09

A7. Efficiency Analysis

While MADR inherently increases token consumption due to its multi-round dialectic debate, we optimized the pipeline's cost-efficiency by utilizing the lightweight Gemini-2.5-Flash model and leveraging context caching. Specifically, caching the shared structured game log significantly reduces redundant processing costs across various debate rounds. As detailed in Table A3, despite processing a high volume of input tokens (75.9k), the total inference cost of MADR (˜$0.0356/rally) remains highly comparable to a single-agent pass using the heavier Gemini-2.5-Pro

model (˜$0.0320/rally). Notably, at this equivalent price point, MADR substantially improves the H@1 accuracy from 28.70% to 55.65%. This demonstrates that intelligently allocating compute to a structured, multi-agent reasoning framework — paired with modern caching mechanisms — is significantly more cost-effective for specialized tactical domains than simply scaling up a single dense model.

Table A3. Cost and latency statistics per rally during the online LLM narrative generation phase.

Method Configuration	H@1 (%)	In Token	Out Token	Cost ($/rally)	Latency (s/rally)
Single (Gemini-2.5-Flash)	29.13	4.1k	0.7k	$0.0029	8.2s
Single (Gemini-2.5-Pro)	28.70	4.1k	2.7k	$0.0320	29.5s
MADR (Gemini-2.5-Flash)	**55.65**	75.9k	9.2k	$0.0356	107.8s

A8. Domain-Specific Fine-Tuning

To explore whether the retrieval models could further adapt to the tactical badminton domain, we fine-tuned both the Qwen3 embedding and re-ranking models. We automatically generated 1,000 disjoint query-narrative training pairs using the MADR pipeline, employing Gemini-2.5-Pro as a strict relevance judge to filter high-quality positive matches (score $\geq 8/10$) and hard negative samples.

As shown in Table A4, domain-specific fine-tuning yielded substantial improvements, particularly when retrieving the high-quality narratives generated by the Gemini-2.5-Flash team. For example, the base Qwen3-Embedding-4B model saw an absolute H@1 gain of +11.30%. Furthermore, scaling to the 8B embedding model, coupled with the 8B re-ranker, achieved the highest overall performance, reaching an H@1 of 68.26% and an H@10 of 90.43%.

Interestingly, applying the fine-tuned Qwen3-Reranker-4B to the Qwen3-8B narratives resulted in slight performance degradation (marked in red). This likely occurs because the fine-tuned model learned to rely on explicit causal markers and a deep tactical vocabulary present in the training data. When these specific features are absent in the simpler Qwen3-8B texts, a feature mismatch occurs, leading the fine-tuned model to perform slightly worse than its zero-shot counterpart.

A9. The MADR Process: A Comprehensive Case Study

This appendix provides a full transcript of the MADR process for Rally 1. The debate is grounded in the structured game log presented in **Figure A1** (Appendix A1). It illustrates how the system evolves from conflicting initial interpretations to a converged, nuanced tactical narrative through four rounds of debate and a final peer-review phase.

Round 1: Initial Thesis Formation

In this round, agents analyze the game log independently to form their initial hypotheses.

Table A4. Retrieval Results Comparison after Domain-Specific Fine-Tuning. Dark Green indicates absolute performance improvement, while Red indicates performance decline.

Method	Hit@K (%) ↑		
	H@1	H@5	H@10
Ours (Debate Team: Qwen3-8B w/o Thinking)			
Qwen3-Embedding (0.6B)	23.04 (+4.34)	49.13 (+14.78)	59.57 (+14.79)
Qwen3-Embedding (0.6B) + Rerank	30.00 (+12.61)	50.43 (+12.60)	64.78 (+17.82)
Qwen3-Embedding (4B)	24.78 (+0.87)	53.48 (+11.74)	62.17 (+11.74)
Qwen3-Embedding (4B) + Rerank	28.26 (+0.87)	41.30 (-4.35)	48.70 (-2.60)
Qwen3-Embedding (8B)	20.87 (+3.04)	46.09 (+6.96)	56.96 (+6.96)
Qwen3-Embedding (8B) + Rerank	39.57 (+12.18)	54.35 (+6.52)	58.26 (+0.87)
Ours (Debate Team: Gemini-2.5-Flash w/o Thinking)			
Qwen3-Embedding (0.6B)	56.52 (+16.95)	79.13 (+15.65)	84.78 (+12.61)
Qwen3-Embedding (0.6B) + Rerank	56.96 (+20.87)	80.43 (+11.30)	88.70 (+12.18)
Qwen3-Embedding (4B)	52.17 (+11.30)	77.39 (+10.87)	83.04 (+8.69)
Qwen3-Embedding (4B) + Rerank	63.04 (+7.39)	81.74 (+4.35)	84.35 (+1.31)
Qwen3-Embedding (8B)	54.35 (+20.44)	77.83 (+16.53)	86.09 (+16.96)
Qwen3-Embedding (8B) + Rerank	**68.26** (+13.04)	**86.52** (+11.74)	**90.43** (+11.30)

Offense Analyst

Opening Thesis: The Bottom-court Player secured the rally victory through a well-timed offensive lob, effectively exploiting the Top-court Player's forward court positioning after a net shot, thereby creating a scoring opportunity that culminated in a winning shot.

Causal Reasoning: The Bottom-court Player's victory was primarily driven by a decisive offensive strategy, specifically the execution of a well-placed lob. From a *Spatial Dynamics* perspective, the Top-court Player's commitment to the net after their shot 2 (net shot from (114, 386)) left a significant portion of their rear court exposed. This forward positioning, evident by their location at (161, 333) during shot 3, created a critical vulnerability. The *Tactical Sequencing* employed by the Bottom-court Player was exemplary; after receiving the net shot, they recognized the open court and immediately transitioned from a defensive return to an offensive lob. This shot, executed from their own Center Front court (224, 560), was not merely a clear but a strategic attack. The *Technical Execution* of this lob was precise, sending the shuttle to the Top-court Player's Right Rear court (76, 169). This deep and wide placement, combined with the Top-court Player's compromised court coverage, rendered the return impossible, directly leading to the winning shot and highlighting the Bottom-court Player's superior offensive timing.

[Supporting Evidence]

- Point Winner: Bottom-court Player. Reason for Win/Loss: Bottom-court Player Winning shot.
- Shot 2: Hitter: Top-court Player, Hitting Position: Top-court Player's Center Front court (114, 386), Shot Action: net shot.
- Shot 3: Hitter: Bottom-court Player, Hitting Position: Bottom-court Player's Center Front court (224, 560), Shot Action: lob.
- Shot 3: Shuttle Trajectory: From Hitting Point [Bottom-court Player's Left Front court (258, 510)] to Landing Point [Top-court Player's Right Rear court (76, 169)].

Defense Analyst

Opening Thesis: The Bottom-court Player's effective defensive lob in Shot 3, executed from a challenging front-court position, was the primary causal factor for winning the rally, forcing the Top-court Player into an unrecoverable defensive posture.

Causal Reasoning: The rally's outcome was decisively shaped by the Bottom-court Player's defensive strategy, specifically the execution of the lob in Shot 3. *Spatial Dynamics:* Following the Top-court Player's net shot (Shot 2), the shuttle landed in the Bottom-court Player's Left Front court (258, 510). This forced the Bottom-court Player to move forward to a challenging Hitting Position in their Center Front court (224, 560). From this forward position, the Bottom-court Player executed a lob that sent the shuttle deep into the Top-court Player's Right Rear court (76, 169). This deep placement significantly altered the spatial dynamics, forcing the Top-court Player to cover a substantial distance to the back of their court. *Tactical Sequencing:* The Bottom-court Player's decision to play a lob from a front-court position was a crucial defensive maneuver. Instead of attempting a risky offensive shot from a compromised position, the lob served to reset the rally, gain time, and push the opponent deep. This tactical choice effectively neutralized the Top-court Player's previous net shot. *Technical Execution:* Despite hitting from a 'Below the net' height, the Bottom-court Player's lob (Shot 3) demonstrated effective technical execution by achieving a deep landing in the opponent's rear court. The Top-court Player's subsequent 'leaning forward' pose suggests they were caught off balance and out of position, struggling to reach the shuttle effectively. This suggests the lob's depth and placement were highly effective, making it an unrecoverable shot for the Top-court Player.

[Supporting Evidence]

- Point Winner: Bottom-court Player. Reason for Win/Loss: Bottom-court Player Winning shot.
- Shot 2: Hitter: Top-court Player, Shot Action: net shot, Shuttle Trajectory: From Hitting Point [Top-court Player's Center Mid court (130, 365)] to Landing Point [Bottom-court Player's Left Front court (258, 510)].
- Shot 3: Hitter: Bottom-court Player, Hitting Position: Bottom-court Player's Center Front court (224, 560).
- Shot 3: Shot Action: lob [Features: Forehand, Hit Height: Below the net].
- Shot 3: Shuttle Trajectory: From Hitting Point [Bottom-court Player's Left Front court (258, 510)] to Landing Point [Top-court Player's Right Rear court (76, 169)].
- Opponent's pose attempting to return the winning shot: The person is leaning forward with their right arm raised and left arm extended.

Round 2: Adversarial Critique

Offense Analyst

Rejoinder: The assertion that the Bottom-court Player's lob was 'defensive' fundamentally misinterprets the strategic intent and outcome. This was not a mere defensive reset; it was a calculated offensive maneuver that capitalized on the Top-court Player's aggressive net play, creating a scoring opportunity. The opponent's argument fails to recognize the offensive timing and the creation of a scoring opportunity, instead framing a winning shot as a defensive recovery.

[Supporting Evidence]

- The Bottom-court Player's lob (Shot 3) was executed from their Center Front court (224, 560), a position that, while forward, allowed for an immediate offensive response to the Top-court Player's net shot.
- The shuttle trajectory of Shot 3 landed in the Top-court Player's Right Rear court (76, 169), a deep and wide placement designed to exploit the Top-court Player's forward court positioning (161, 333) at the time of the shot.
- The rally result explicitly states 'Bottom-court Player Winning shot,' directly contradicting the notion of a purely defensive action.
- The Top-court Player's 'leaning forward' pose attempting to return the winning shot indicates they were caught off balance and out of position.

Defense Analyst

Rejoinder: The opponent's claim that the Bottom-court Player's lob was a 'calculated offensive maneuver' fundamentally misrepresents the defensive necessity and strategic timing of the shot. The lob, executed from a challenging front-court position, was a critical defensive reset, not an offensive creation of opportunity. The Top-court Player's aggressive net play forced the Bottom-court Player into a compromised position, making a defensive lob the most strategically sound option to regain court control and time, rather than an immediate offensive strike. The subsequent 'winning shot' was a consequence of the Top-court Player's inability to recover from this defensive pressure, not a direct offensive setup.

[Supporting Evidence]

- Shot 2: Top-court Player's net shot landed in Bottom-court Player's Left Front court (258, 510), forcing the Bottom-court Player to move forward.
- Shot 3: Bottom-court Player's Hitting Position was Center Front court (224, 560), indicating a forward and potentially compromised position for an immediate offensive strike.
- Shot 3: The shot action was a 'lob', which, when executed from a front-court position, typically serves to gain time and push the opponent deep, aligning with a defensive strategy.
- Final Result: Opponent's pose attempting to return the winning shot indicates a struggle to reach the shuttle, suggesting the lob's depth created an unrecoverable defensive situation for the Top-court Player, rather than a direct offensive setup by the Bottom-court Player.

Round 3: Deepening and New Insights

Offense Analyst

Rejoinder: While acknowledging the defensive pressure exerted by the Top-court Player's net shot, the characterization of the Bottom-court Player's subsequent lob as purely a 'defensive reset' overlooks its inherent offensive potential and the

strategic timing that transformed it into a scoring opportunity. The lob, though initiated from a challenging position, was not merely about regaining court control but about exploiting the Top-court Player's forward momentum and creating an open court for a winning shot. The opponent's argument focuses on the defensive origin of the shot rather than its offensive outcome and the strategic intent behind it.

[Supporting Evidence]

- Evidence 1: Shot 2 (net shot by Top-court Player) landed in Bottom-court Player's Left Front court (258, 510), forcing the Bottom-court Player to move to Center Front court (224, 560). This indeed put the Bottom-court Player in a forward, potentially compromised position.
- Evidence 2: The Bottom-court Player's lob (Shot 3) landed in Top-court Player's Right Rear court (76, 169). This deep and wide placement is characteristic of an offensive lob designed to exploit court space.
- Evidence 3: At the time of Shot 3, the Top-court Player was located at (161, 333), indicating a forward-mid court position, which left the rear court vulnerable.
- Evidence 4: The 'Reason for Win/Loss' explicitly states 'Bottom-court Player Winning shot', directly attributing the rally win to the offensive nature of the final shot, which was the lob.

[New Insight] The Bottom-court Player's ability to execute a high-quality lob from a 'Below the net' hitting height, despite being in a forward and potentially defensive position, demonstrates a superior technical capacity to convert a challenging situation into an offensive advantage. This technical skill allowed for the precise placement and depth required to create the scoring opportunity, rather than just a neutral clear.

Defense Analyst

Rejoinder: While the opponent highlights the Bottom-court Player's technical capacity to execute a lob from a challenging position, framing it as an 'offensive advantage,' this perspective still understates the critical defensive timing and data-driven necessity of the shot. The lob, despite its eventual outcome, was primarily a defensive response to regain court control and reset the rally after being drawn to the net, rather than an immediate offensive creation. The 'winning shot' was a consequence of the Top-court Player's inability to recover from the defensive pressure created by the lob's depth, not solely due to the Bottom-court Player's offensive intent.

[Supporting Evidence]

- Shot 2: Top-court Player's net shot landed in Bottom-court Player's Left Front court (258, 510), forcing the Bottom-court Player into a forward, compromised position.
- Shot 3: Bottom-court Player's Hitting Position was Center Front court (224, 560), indicating they were still recovering from the net shot.
- The lob's primary function from a front-court, 'Below the net' position is often to gain time and push the opponent back, which aligns with a defensive reset strategy.
- The Top-court Player's 'leaning forward' pose attempting to return the winning shot suggests they were reacting to a deep shot that put them under defensive pressure, rather than being caught off guard by a purely offensive setup.

[New Insight] The Bottom-court Player's choice of a lob (Shot 3) from a 'Below the net' hitting height, despite being in a front-court position, demonstrates an effective defensive strategy to mitigate risk and force a deep return. This shot type, when executed with sufficient depth, statistically reduces the opponent's immediate attacking options and increases the likelihood of a weak return or an unforced error, thereby shifting the rally's momentum defensively.

Round 4: Convergence

Offense Analyst

Rejoinder: While acknowledging the defensive pressure that led to the lob, it's crucial to integrate the understanding that the Bottom-court Player's execution of this lob, even from a challenging position, was not merely a defensive reset but a strategically timed offensive transition. The depth and placement of the lob, as a direct consequence of the Top-court Player's forward court commitment, transformed a potentially defensive situation into a scoring opportunity. The 'defensive pressure' argument overlooks the offensive intent and outcome that the Bottom-court Player skillfully engineered, turning a reactive shot into a proactive attack that created the winning scenario.

Defense Analyst

Rejoinder: While the opponent highlights the Bottom-court Player's technical skill in executing the lob from a challenging position, and suggests it was an 'offensive transition,' it's more accurate to view this as a highly effective defensive maneuver that created an *opportunity* for a winning shot, rather than being an offensive shot itself. The primary causal factor remains the Bottom-court Player's astute defensive timing and shot selection, which forced the Top-court Player into an unrecoverable position. The lob's depth and placement were crucial in resetting the rally and shifting momentum, which is a hallmark of strong defensive play, even if it ultimately led to a point.

Final Phase: Narrative Synthesis and Review

Following the debate, the *Tactic Summarizer* generated a draft. This draft was then critiqued by specialized reviewers to ensure tactical depth before the final version was produced.

Draft Report (Version 1)

Bottom-court Player Strategy: The Bottom-court Player initiated the rally with a short service (Shot 1) from their Center Mid court. After the Top-court Player's net shot drew them to the front, the Bottom-court Player executed a lob (Shot 3) from their Center Front court. This lob, despite being hit from a 'Below the net' height, was strategically placed deep into the

Top-court Player's Right Rear court, exploiting the Top-court Player's forward positioning and securing the winning shot.

Turning Point: The turning point occurred with the Bottom-court Player's lob (Shot 3). After being drawn to the net by the Top-court Player's net shot, the Bottom-court Player's ability to execute a deep and precise lob from a challenging front-court position, landing in the Top-court Player's Right Rear court, effectively exploited the Top-court Player's forward court commitment and created an unrecoverable situation.

Reviewer Feedback:

- **Offense Analyst**: [Needs Revision] Regarding the *Bottom-court Strategy*, expand on the 'why' and 'how'. Analyze the Technical Execution of the lob from a 'Below the net' height. How did this technical skill allow them to turn a defensive retrieval into an offensive winning shot? For the *Turning Point*, explicitly link the Top-court Player's 'forward court commitment' to the success of the lob.
- **Defense Analyst**: [Needs Revision] The report correctly identifies the lob as strategic, but it needs to delve deeper into the *defensive aspect*. Highlight the technical execution of this defensive lob and its effectiveness in neutralizing the Top-court Player's net play. For the *Turning Point*, elaborate on how the lob reset the rally's spatial dynamics. Analyze the 'Below the net' hit height as a key detail underscoring the defensive skill involved.

Final Verified Report (after Revision)

Bottom-court Player Strategy: The Bottom-court Player initiated the rally with a short service (Shot 1) from their Center Mid court. After the Top-court Player's net shot drew them to the front, the Bottom-court Player executed a lob (Shot 3) from their Center Front court (224, 560). This lob, despite being hit from a challenging 'Below the net' height, demonstrated significant technical skill. The Bottom-court Player's ability to generate sufficient height and depth from this defensive position allowed them to strategically place the shuttle deep into the Top-court Player's Right Rear court (76, 169). This technical execution turned a defensive retrieval into an offensive winning shot, exploiting the Top-court Player's forward positioning and forcing them to retreat from a forward position, creating an unrecoverable situation.

Turning Point: The turning point occurred with the Bottom-court Player's lob (Shot 3). After being drawn to the net by the Top-court Player's net shot, the Bottom-court Player's ability to execute a deep and precise lob from a challenging front-court position (224, 560), landing in the Top-court Player's Right Rear court (76, 169), effectively exploited the Top-court Player's forward court commitment. The 'Below the net' hit height of the lob underscores the defensive skill involved in turning a potentially vulnerable position into a winning opportunity. This defensive counter-attack was timed to effectively counter the net shot, resetting the rally's spatial dynamics and creating an unrecoverable situation due to the Top-court Player's inability to adjust their court coverage in time.

Win/Loss Attribution: The Bottom-court Player won the rally due to their effective defensive counter-attack. The Bottom-court Player's technical execution of a deep lob (Shot 3) from a challenging 'Below the net' front-court position directly exploited the Top-court Player's aggressive forward court positioning after their net shot. This created an immediate and decisive scoring opportunity that the Top-court Player could not recover from, highlighting the Bottom-court Player's defensive prowess in turning a vulnerable position into a winning shot.

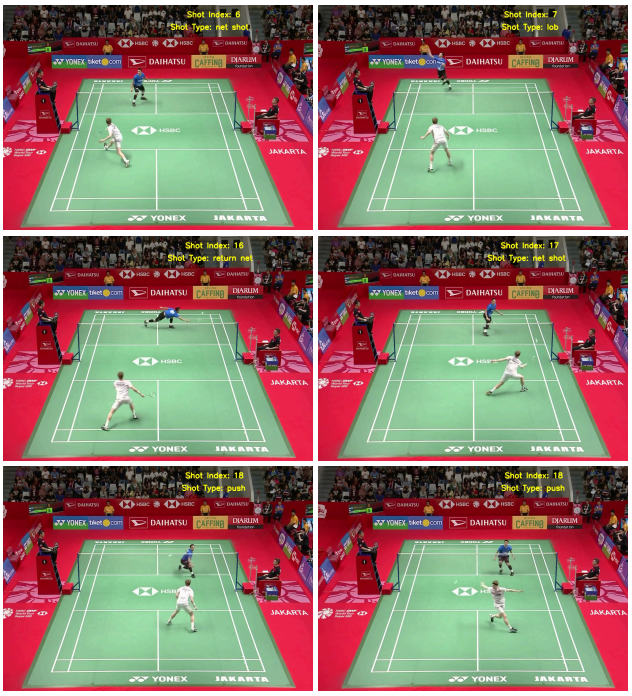
Figure A8. **Visual Verification of the Final Report (Rally 1).** This sequence confirms the MADR tactical narrative. **(Top-Left) Shot 2:** The Top-court Player's net shot draws the opponent forward. **(Top-Right) Shot 3:** The Bottom-court Player executes a defensive lob from a challenging "Below the net" stance, as analyzed. **(Bottom-Left) Reaction:** Committed forward, the Top-court Player attempts a swing but misses, validating the "Unrecoverable Situation." **(Bottom-Right) Result:** The shuttle lands in bounds, confirming the "Winning Shot" attribution.

A10. Retrieval Results Examples

To demonstrate the effectiveness of the "Generate-then-Retrieve" framework, we present qualitative examples of retrieved rallies for complex, natural-language queries. These examples highlight the system's ability to align user intent with the deep tactical narratives generated by the MADR pipeline, surpassing the capabilities of keyword matching or surface-level visual alignment.

Example 1: Abstract Intent

User Query: "Find rallies that demonstrate **systematic disruption of opponent's defensive timing** through **alternating deep and short shot placement** to create spatial imbalances."



Top-court Player Strategy:
The Top-court Player adopted an aggressive and varied offensive strategy, explicitly designed to disrupt the Bottom-court Player's defensive timing and create scoring opportunities. They immediately countered the initial service with a cross-court net shot (Shot 2). Following defensive lobs from the Bottom-court Player, they consistently responded with powerful wrist smashes (Shot 4, Shot 8) to maintain pressure and force weak returns, often setting up subsequent net play. The Top-court Player also utilized strategic net play, including a net shot (Shot 6) and another cross-court net shot (Shot 10), to pull the Bottom-court Player to the front court. A calculated passive drop (Shot 12) was used to draw the Bottom-court Player forward and out of position, specifically exploiting their recovery path and creating spatial disequilibrium. They also employed a lob (Shot 14) and a return net (Shot 16) to manage the rally, ultimately securing the point with a well-placed push (Shot 18) that exploited the Bottom-court Player's compromised defensive position and timing.

Win/Loss Attribution:
The Top-court Player's consistent offensive pressure, achieved through strategic variations in shot placement and tempo, systematically broke down the Bottom-court Player's defensive structure and timing. The Top-court Player's ability to alternate between deep, powerful wrist smashes (Shot 4, Shot 8) and short, deceptive shots created scoring opportunities by forcing the Bottom-court Player into predictable, weak, and often 'Below the net' returns. This systematic exploitation of the Bottom-court Player's defensive anticipation and court coverage ultimately led to the winning shot.

Figure A9. **Retrieval Result (Example 1).** Retrieval of a rally matching a complex strategic query. The generated Strategy Analysis directly addresses the user's request for "alternating shot placement" and "timing disruption".

Example 2: Performance Analysis

User Query: "Instances of winning a point by executing a **precise, low cross-court net shot** that **forces the opponent into a stretched position**, leading to a **technical error** where the shuttle fails to clear the net."

Turning Point Analysis:
The turning point occurred with the Top-court Player's cross-court net shot (Shot 2). This shot, following the Bottom-court Player's short service, effectively pulled the Bottom-court Player wide and to the front of the court (101, 544). This spatial displacement severely constrained the Bottom-court Player's options and execution window, directly impacting their ability to execute a successful net shot. Being stretched and out of position led to a low-quality defensive return with a 'Below the net' hit height (Shot 3), which failed to clear the net, establishing a clear causal link between the spatial dynamics and the technical error.

Win/Loss Attribution:
The Top-court Player won the rally due to their effective offensive strategy, specifically the precise placement and low trajectory of the cross-court net shot (Shot 2). This directly exploited the Bottom-court Player's court coverage and created an unrecoverable situation, leading to a forced technical error (Shot 3) where the shuttle failed to clear the net. This emphasizes the proactive nature of the Top-court Player's win, as their offensive pressure created a defensive dilemma and directly caused the Bottom-court Player's error.

Figure A10. **Retrieval Result (Example 2).** Retrieval of a rally matching a specific causal query. The generated Turning Point Analysis explicitly explains the causal chain: precise shot → stretched position → forced technical error.