```
This is a plan that contains detailed bullet points about a topic. This plan is used to generate realistic
     chat conversations between a user and an AI assistant, which are then used to evaluate the long-term
     memory capabilities of LLMs.
Your task is to analyze this plan and select PAIRS of bullet points that would be most effective for testing
     contradiction resolution abilities when incorporated into chat conversations.
Contradiction resolution tests whether the LLM can detect and appropriately handle impossible contradictions -
      statements that logically cannot both be true simultaneously.
Analyze this project plan and identify PAIRS of bullet points where one completely contradicts the other with
     impossible contradictions. Each pair should contain statements that are logically incompatible and
     cannot both be true.

## INPUT DATA
- **PLAN**: <plan>

Focus on bullet point pairs that show:
- **Never-Statement Violations**: One bullet says "never" did something, another shows they did it
- **Always-Statement Violations**: One bullet claims "always" pattern, another breaks that pattern
- **Only-Statement Conflicts**: One bullet claims exclusivity ("only"), another contradicts it
- **Impossible Reversals**: Age going backward, timeline impossibilities, logical reversals
- **Dead-Alive Contradictions**: References to deceased people being active
- **Mutually Exclusive States**: Being in two places simultaneously, having contradictory capabilities
- **Absolute Negations**: Claiming something is impossible then showing it happened

**Types of Impossible Contradictions to look for:**
1. **Never-Statement Violations**: "Never attended X" vs "Attended X event"
2. **Always-Statement Violations**: "Always lived in Y" vs "Moved from Z to Y"
3. **Only-Statement Conflicts**: "Only child" vs "Has siblings"
4. **Timeline Impossibilities**: Events happening in wrong chronological order
5. **Capability Contradictions**: "Cannot do X" vs "Successfully did X"
6. **Location Impossibilities**: Being in two places at once
7. **Relationship Contradictions**: "Never met person" vs "Long friendship with person"

Prioritize bullet point pairs that:
- Contain completely impossible contradictions that cannot be resolved or explained
- Use absolute language ("never," "always," "only," "impossible," "cannot")
- Create clear logical impossibilities rather than simple inconsistencies
- Enable questions about detecting fundamental contradictions
- Test whether the AI can identify when statements are mutually exclusive
- Focus on contradictions that are objectively impossible, not subjective differences

Return your analysis in this exact JSON format where each object contains exactly TWO contradicting bullet
     points:
[{"capability": "contradiction_resolution", "batch_numbers": "1, 8", "bullet_numbers": "30, 29",
        "bullet_points": "• **Logical Contradiction:** Jeremiah has never attended any Bahrain Jazz Festival
             events. | • **Character & Relationship:Close Friend:** Jeremiah, 37, met at Bahrain Jazz
             Festival 2015, recommended an acoustic consultant."}
]

Important formatting notes:
- The "batch_numbers" and "bullet_numbers" correspond to each other positionally
- "1, 8" and "30, 29" means: Batch 1 Bullet 30, Batch 8 Bullet 29
- Each object must contain exactly 2 bullet points separated by " | "
- Use comma-separated values for batch_number and bullet_number
- First bullet point can be the contradiction marker or the contradicted statement
- Second bullet point should directly contradict the first with impossible logic
- Focus on pairs that test detection of fundamental logical impossibilities

Select <bullet_number> pairs of bullet points that demonstrate the clearest impossible contradictions for
     testing contradiction resolution abilities.

NOTE: Only output the list without any explanation before or after the list.
```

Listing 7: Candidate selection contradiction resolution prompt

```
This is a plan that contains detailed bullet points about a topic. This plan is used to generate realistic
     chat conversations between a user and an AI assistant, which are then used to evaluate the long-term
     memory capabilities of LLMs.
Your task is to analyze this plan and select GROUPS of bullet points that would be most effective for testing
     summarization abilities when incorporated into chat conversations.
Summarization tests whether the LLM can synthesize and condense information from across multiple conversation
     sessions into coherent, comprehensive summaries.
Analyze this plan and identify GROUPS of 8-12 related bullet points that represent topics suitable for
     summarization testing. Groups can vary in size depending on the richness and complexity of the topic.

## INPUT DATA
- **PLAN**: <plan>

## CRITICAL REQUIREMENT: EARLY BATCH PRIORITIZATION
**SELECTION PRIORITY ORDER:**
1. **Groups starting in Batches 1-3 (HIGHEST PRIORITY)**: Select 60-70% of your groups with foundational
     content from early batches
2. **Groups spanning early to middle batches (MEDIUM PRIORITY)**: Select 20-30% of groups that bridge early-to
     -middle timeline
3. **Groups from later batches only (LOW PRIORITY)**: Select only 10-20% from purely later batches

## CRITICAL REQUIREMENT: CONTENT-BASED ANALYSIS
**ANALYZE BULLET CONTENT, NOT CATEGORY NAMES:**
- Read the actual bullet point text to identify mentions of entities (people, places, items, topics, amounts,
     processes)
- The same entity might appear in different category types - include mentions across all categories
- The same process might appear in different category types - include mentions across all categories
```