## Section 4 - Comparison of Results: Mistral vs. Qwen2

**Structure and Clarity**

**Mistral:** Presents reasoning in a linear, step-by-step format, with numbered steps Observation, Hypothesis, Conclusion

The explanation is straightforward, focusing on deducing a logical cause (the glasses being on the head) with minimal exploration of alternative scenarios.

**Qwen2:** Breaks reasoning into conceptual sections like assessmentexploration and resolution**.** Offers a more nuanced exploration of scenarios, such as sensory issues, fatigue, and environmental factors, adding richness to the reasoning process.

**Observation:**

* Mistral adopts a deductive reasoning approach, arriving at a specific conclusion efficiently.
* Qwen2 takes a divergent reasoning approach, exploring multiple possibilities before converging on the conclusion.

**Depth of Reasoning**

**Mistral** Focuses on logical deduction:

* John couldn't find his glasses because they were already on his head, a conclusion directly tied to observable behaviors.
* It doesn’t delve into alternative reasons why John might fail to locate the glasses.

**Qwen2** Provides contextual reasoning:

* Explores psychological and sensory factors, such as misperception, fatigue, or blurriness, as potential contributors to John's confusion.
* Considers more human-like reasoning patterns, including external environmental influences.

**Observation:**

* Mistral provides a concise, logical explanation focused solely on solving the problem.
* Qwen2 adds complexity by integrating human-like contextual details, making its response feel more relatable but less concise.

**Style and Tone**

**Mistral:**

* Maintains a formal, academic tone.
* Prioritizes clarity and precision over storytelling or empathy.

**Qwen2:**

* Adopts a more conversational and reflective tone.
* Incorporates elements of empathy and real-world reasoning by acknowledging possible human errors or environmental influences.

**Observation:**

* Mistral's tone is analytical and straightforward, suitable for tasks requiring formal reasoning.
* Qwen2’s tone is exploratory and human-like, better suited for conversational or creative applications.

Practical Utility

**Mistral:**

* Best suited for scenarios where a concise explanation is needed.
* Useful in contexts like: Automated reasoning systems Structured analysis

**Qwen2:**

* More effective in applications requiring user engagement or creative reasoning, such as: Educational tools, Conversational assistants

**Observation:**

* Mistral excels in formal, structured environments.
* Qwen2’s broader reasoning is better for scenarios where multiple perspectives or creative exploration is required.

**Overall Comparison Table**

|  |  |  |
| --- | --- | --- |
| Aspect | Mistral | Qwen2 |
| Structure | Linear, step-by-step | Conceptual, exploratory |
| Reasoning Style | Deductive | Divergent and contextual |
| Clarity | Clear and concise | Rich but slightly verbose |
| Focus | Logical conclusion | Exploration of multiple scenarios |
| Tone | Formal and academic | Reflective and empathetic |
| Utility | Best for structured tasks | Best for conversational engagement |

**Discussion**

The choice between Mistral and Qwen2 depends on the context of use:

1. For applications requiring precision and brevity—such as debugging, structured problem-solving, or decision-making systems—Mistral is more effective.
2. For tasks requiring engagement, human-like reasoning, or creative exploration—such as conversational AI or storytelling—Qwen2 provides richer and more nuanced responses.

Both models showcase their strengths:

* Mistral focuses on concise, logical reasoning.
* Qwen2 provides a human-like exploration of scenarios, which can sometimes dilute focus but feels more relatable.

# Section 5 - hidden assumption or direct reasoning

# Summary Tables – Mistral model

**Sentence: John couldn't find his glasses while they were on his head.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | Based on John's claim, it logically follows that his vision or awareness might be impaired. |
| 2 | Hidden Assumption | Assumes John has searched usual places thoroughly and would recognize his glasses if visible. |
| 3 | Direct Reasoning | Concludes that the observed behavior aligns with the fact that the glasses were on his head. |

**Sentence: After the rain, Sarah grabbed her umbrella before leaving the office.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Hidden Assumption | Assumes Sarah's act of grabbing an umbrella is directly connected to the occurrence of rain, without external confirmation. |
| 2 | Direct Reasoning | It logically follows that Sarah was at the office during the rain because she only grabbed the umbrella when leaving. |
| 3 | Hidden Assumption | Assumes Sarah's departure coincides with the rain stopping or subsiding, without explicitly confirming the timing. |

**Sentence: The coffee was too hot to drink, so I added an ice cube.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | The reasoning directly follows from the observation that excessively hot liquids are uncomfortable or unsafe to consume. |
| 2 | Hidden Assumption | Assumes that cooling the coffee is the only solution, without verifying whether the drinker might prefer alternative approaches (e.g., drinking carefully). |
| 3 | Direct Reasoning | The decision to add an ice cube and subsequently evaluate the coffee's temperature logically follows the goal of cooling the coffee quickly. |

**Sentence: Tom can't go to his sister's wedding because he's studying abroad in Japan.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | The reasoning directly follows from the information provided, which explicitly states Tom's location. |
| 2 | Hidden Assumption | Assumes that studying abroad inherently imposes travel limitations, though no specific evidence is provided. |
| 3 | Direct Reasoning | The conclusion logically follows from the premise that Tom is in Japan and cannot simultaneously be at the wedding. |

**Sentence: She put the groceries away before the ice cream melted.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | The observation directly follows from the result that the ice cream did not melt, implying it was kept cold. |
| 2 | Hidden Assumption | Assumes that the ice cream was immediately placed in a cool place, even though the exact steps are not stated. |
| 3 | Direct Reasoning | The conclusion logically follows from the premise that the ice cream remained intact, implying efficient action. |

**Sentence: The plant died because Jenny forgot to open the curtains for a week.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | The reasoning directly follows from the fact that plants need sunlight for photosynthesis, which was disrupted by Jenny's actions. |
| 2 | Hidden Assumption | Assumes that the lack of sunlight alone was sufficient to cause the plant’s death without considering other potential factors. |
| 3 | Direct Reasoning | The conclusion logically ties the plant's death to the absence of sunlight, assuming other potential causes have been ruled out. |

# Summary Tables – QWEN2 model

**Sentence: John couldn't find his glasses while they were on his head.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | This reasoning directly follows from John's belief that he cannot find the glasses, inferring a misperception due to his vision issues. |
| 2 | Hidden Assumption | Assumes John's inability to locate the glasses is influenced by factors such as fatigue or external conditions, without verifying these possibilities. |
| 3 | Direct Reasoning | The conclusion logically follows from observing the position of the glasses on John's head and reconciling this with his misunderstanding. |

**Sentence: After the rain, Sarah grabbed her umbrella before leaving the office.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | The reasoning follows directly from the purpose of an umbrella and Sarah's anticipation of rain. |
| 2 | Hidden Assumption | Assumes that Sarah had easy and immediate access to the umbrella, though this is not explicitly stated. |
| 3 | Direct Reasoning | This reasoning directly ties Sarah’s decision to grab an umbrella to her preparedness for potential rain. |

**Sentence: The coffee was too hot to drink, so I added an ice cube.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | The observation directly follows from the experience of tasting the coffee and recognizing its excessive heat. |
| 2 | Hidden Assumption | Assumes cooling the coffee is the best solution without verifying alternatives or considering different preferences. |
| 3 | Direct Reasoning | The reasoning logically follows from the identified solution (adding an ice cube) and its immediate implementation. |

**Sentence: Tom can't go to his sister's wedding because he's studying abroad in Japan.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | The reasoning directly connects the spatial distance and logistical challenges of international travel to Tom's inability to attend. |
| 2 | Hidden Assumption | Assumes that time zone differences make participation impractical, though this depends on specific timing and flexibility in commitments. |
| 3 | Direct Reasoning | The conclusion follows directly from Tom's prioritization of academic commitments over personal events like a wedding. |

**Sentence: She put the groceries away before the ice cream melted.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | The reasoning follows logically from her awareness of time-sensitive conditions and her actions to manage them promptly. |
| 2 | Hidden Assumption | Assumes that putting away the ice cream promptly is the only way to prevent melting, without considering other potential factors like storage method. |
| 3 | Direct Reasoning | The reasoning ties her action directly to preserving the integrity of the ice cream and maintaining its usability for future consumption. |

**Sentence: The plant died because Jenny forgot to open the curtains for a week.**

|  |  |  |
| --- | --- | --- |
| Reasoning number | Reasoning Type | Explanation |
| 1 | Direct Reasoning | The reasoning establishes a straightforward cause-and-effect relationship between the lack of sunlight and the plant's death. |
| 2 | Direct Reasoning | This reasoning is directly based on scientific knowledge about the necessity of sunlight for photosynthesis in plants. |
| 3 | Hidden Assumption | Assumes that no other factors (e.g., water, pests, or soil quality) contributed to the plant's death, without explicit verification. |

# Section 6 - Find a sentence that at least one of the models fails to find the reasoning chain. Explain why it fails.

## MISTRAL RESULTS

**This Model is SOTA (State of the Art)**

1. **Logical Deduction**: The models consistently follow logical reasoning patterns, using observed facts to derive valid conclusions.
2. **Contextual Understanding**: They handle nuanced scenarios with implicit details (e.g., the need for light, the use of umbrellas).
3. **Common-Sense Knowledge**: They leverage real-world knowledge to interpret scenarios (e.g., glasses on the head, ice cream melting).

**Why No Sentence Failed**

1. The provided sentences are straightforward and rooted in common-sense reasoning.
2. The reasoning challenges did not involve ambiguous, abstract, or domain-specific knowledge that could confuse the models.
3. The scenarios have clear cause-effect relationships, making it easier for the models to identify reasoning chains.

**Conclusion**

The models successfully reasoned through all sentences, indicating that they are indeed state-of-the-art at finding reasoning chains for common-sense scenarios. Failures may arise in more complex or ambiguous cases, such as:

* Abstract reasoning (e.g., philosophical or metaphorical contexts).
* Missing information requiring external knowledge.

## QWEN2 RESULTS

**Failed Sentence 4: Tom can't go to his sister's wedding because he's studying abroad in Japan.**

* **Reasoning**: The model explores spatial distance, time zone differences, and Tom's prioritization of studies over attending the wedding.
* **Evaluation**: While the reasoning steps are valid, the inclusion of "time zone differences" as a factor for virtual attendance is unnecessary and distracts from the core explanation: Tom’s physical location and study commitments make travel unfeasible.
* **Weakness**: Reasoning number 2 includes irrelevant details (time zones), leading to a diluted focus on the primary reasoning.

**Weaknesses**:

1. The model occasionally overanalyzes simple scenarios (e.g., Sarah grabbing an umbrella, Tom’s wedding constraints) by introducing irrelevant factors like time zones or external preparation efforts.
2. For John's glasses, the reasoning includes unnecessary complexity, detracting from clarity and focus.

**Conclusion**

The Qwen2 model does not outright fail to find reasoning chains for any sentence, but it struggles with:

* **Overanalyzing simple scenarios**, leading to bloated explanations.
* **Introducing irrelevant factors**, which dilute the reasoning's clarity.

These weaknesses, while not failures, demonstrate room for improvement in prioritizing simplicity and relevance. The model still performs at a high level and is capable of handling common-sense reasoning effectively.