

Magma-8B Multi-Technique Analysis Report

Generated: 2025-06-11 08:34:01

Model: Microsoft Magma-8B

Task/Prompt: How can you go to the first floor?

Input Image



Analysis Results

STANDARD Technique

Direct prompting without additional reasoning steps

Response:

You can take the green staircase on the left side of the room.

COT Technique

Chain-of-Thought: Step-by-step reasoning before final answer

Response:

To get to the first floor, you would need to walk down the green staircase on the left side of the image. The staircase leads from the second floor to the first floor, providing access to the lower level of the house.

COD Technique

Chain-of-Draft: Draft thoughts followed by refined response

Response:

To reach the first floor, take the green staircase located in the living room.

COT_CONSISTENCY Technique

Self-Consistency: Multiple reasoning paths with majority voting

Majority Vote Result:

The first floor is accessible via a set of green stairs located in the living room.

All Samples:

Sample 1: The first floor is accessible via a set of green stairs located in the living room.

Sample 2: To get to the first floor, you can use the green stairs located on the left side of the room. The stairs will take you down to the first floor where you can explore more rooms and find what you need.

Sample 3: To get to the first floor, you would need to use the green stairs next to the wooden rocking chair. The stairs lead down to the lower level of the house.

Sample 4: There is a green staircase in the background that leads to the first floor.

Sample 5: To go to the first floor, you would need to find the green staircase on your right side of the room and climb up the stairs.

Analysis Summary

This report compares four different prompting techniques applied to the Magma-8B vision-language model. Each technique has different strengths:

- **Standard:** Fast and direct, good for simple tasks
- **CoT:** Better reasoning for complex tasks requiring step-by-step thinking
- **CoD:** Iterative refinement for improved response quality
- **CoT-Consistency:** Most robust but computationally expensive, reduces hallucinations