

## A Prompts

### Prompt Construction

1. General Chain-of-thought: CG, RT, ToM
2. Gricean and Inference Norms: CG, RT, ToM
3. Response Generation
4. Few-shot with Chain-of-thought: CG, RT, ToM

Figure 8: Prompt Components

### Background

**Background:** You are assisting a human in a cooperative planning domain called Doors, Keys, and Gems, set in a gridworld. The grid contains various objects such as keys, doors, gems, walls, empty spaces, the human, and yourself (the AI agent). The human’s objective is to retrieve a specific gem by issuing a single instruction. Your role is to interpret this instruction—whether it is clear, incomplete, ambiguous, irrelevant, or invalid—and assist by taking actions such as retrieving and passing keys or unlocking doors. Your actions should aim to minimize your movement while ensuring the human can achieve their goal as efficiently and quickly as possible. After interpreting the human’s instruction, you will generate an optimal response that supports the human in reaching their desired gem in the most optimal manner.

### Common Ground and Relevance Theory

**Common ground, Theory of mind, and Relevance Theory:** You and the human share a common understanding of the environment and objectives, with the ultimate goal of retrieving a specific gem. Upon receiving a single instruction, your task is to assess its type, discern the human’s intent, and uncover any implicit meaning. Based on this interpretation and the current grid configuration, you will infer the most relevant action and generate a response that facilitates the collection of the desired gem.

Figure 9: First Component of the Prompt: Background, Common Ground and Relevance Theory

## Grid Configuration Instance

The Current Grid Layout for the Keys, Doors, and, Gems:

```
[['r' '.' '.' '.' '.' 'm' 'W' 'W' 'g']]
[['y' '.' 'W' 'W' '.' 'W' 'W' '.']]
[['W' 'W' 'W' 'W' '.' 'W' 'W' '.']]
[['.' 'R' '.' '.' '.' '.' 'h' '.']]
[['.' 'W' '.' 'W' 'W' 'W' 'W' '.']]
[['.' 'W' '.' 'W' 'W' 'W' 'W' 'Y']]
[['Y' 'W' '.' 'W' 'W' 'W' 'W' '.']]
[['g' 'W' 'g' 'W' 'W' 'W' 'W' 'g']]
```

Below are the locations of all objects on the grid, displayed as (row, column) coordinates, along with their total count:

My position (Labeled as 'm'): (0, 4)

Human (Labeled as 'h'): (3, 6)

Red key (Labeled as 'r'): (0, 0) -> Total Red key: 1

Yellow key (Labeled as 'y'): (1, 0) -> Total Yellow key: 1

Red door (Labeled as 'R'): (3, 1) -> Total Red door: 1

Yellow doors (Labeled as 'Y'): (5, 7), (6, 0) -> Total Yellow doors: 2

Gems (Labeled as 'g'): (0, 7), (7, 0), (7, 2), (7, 7) -> Total Gems: 4

Walls (Labeled as 'W'): (0, 5), (0, 6), (1, 2), (1, 3), (1, 5), (1, 6), (2, 0), (2, 1), (2, 2), (2, 3), (2, 5), (2, 6), (4, 1), (4, 3), (4, 4), (4, 5), (4, 6), (5, 1), (5, 3), (5, 4), (5, 5), (5, 6), (6, 1), (6, 3), (6, 4), (6, 5), (6, 6), (7, 1), (7, 3), (7, 4), (7, 5), (7, 6) -> Total Walls: 32

Empty spaces (Labeled as '.'): (0, 1), (0, 2), (0, 3), (1, 1), (1, 4), (1, 7), (2, 4), (2, 7), (3, 0), (3, 2), (3, 3), (3, 4), (3, 5), (3, 7), (4, 0), (4, 2), (4, 7), (5, 0), (5, 2), (6, 2), (6, 7) -> Total Empty spaces: 21

## Key Points

**Key Points:**

- The human will provide only one instruction.
- Infer the specific gem the human wants to retrieve from the instruction.
- Evaluate the instruction using common ground (shared knowledge), relevance theory (focusing on pertinent details), and theory of mind (considering the human's perspective), based on the current grid configuration.
- Always generate a response with the understanding that the implied goal is to collect the specific gem. Keep this objective in mind when interpreting the instruction and determining the most relevant action.
- Cross-check the instruction against the current grid state, ensuring it provides adequate details, clarity, validity, relevance, and aligns with the intended goal.
- Always generate your response by cross-checking and verifying it against the current grid configuration to ensure accuracy.
- Each key can unlock only one door of the corresponding color. A key cannot be reused to unlock multiple doors of the same color.
- Prioritize collecting keys before unlocking doors. Even if not explicitly mentioned, collect any necessary keys if access through a locked door is required.
- Once the keys are collected, automatically pass them to the human if the instruction involves passing, finding, or collecting keys.
- If unable to pass the key due to being blocked by a door, request clarification from the human on how to proceed.
- Always infer that an instruction involving a key or keys means you should collect them, even if not explicitly stated. Determine which gem the human wants to collect and which key is needed to help them.
- If explicitly instructed to unlock doors, focus on unlocking them by collecting the necessary keys, without passing them to the human.
- When traversing the grid, minimize and optimize your path, avoiding collisions with walls (W) or other obstacles.
- Always prioritize efficient, obstacle-free routes to achieve the human's objective as quickly and effectively as possible.

Figure 10: First Component of the Prompt: Grid Configuration and Key Points

## Norms and Relevance Theory

**Gricean and Inference Norms:** You follow specific communication norms to interpret the instruction and generate a response, ensuring the human successfully achieves their goal of collecting a specific gem after providing an instruction. You adhere to Gricean norms—Quantity, Quality, Relevance, and Manner—when evaluating instructions. If a norm is violated, apply the Action/Inference Norm to infer the most relevant action and generate a response based on the current grid configuration. If no norms are violated, execute the action directly as per the instruction after careful interpretation. Always assess which specific gem the human wants to retrieve, ensuring that your actions enable the human to reach that gem efficiently within the grid.

### Norms:

1. **Quantity Norm:** The instruction must provide sufficient information for the human to reach the desired gem on the current grid. If the instruction specifies collecting or unlocking the necessary object(s) to retrieve the gem based on the grid configuration, it is considered sufficient. However, if additional actions—such as collecting another key or unlocking another door—are required and these actions aren't mentioned, it results in a Quantity Violation. For example, if the grid has two blue doors, three blue keys, two red keys, and a gem beyond the blue doors, an instruction to pick up only one blue key would be a Quantity Violation. In this case, both blue keys are needed to unlock the doors and reach the gem, so you should collect two blue keys and pass them to the human. Conversely, if the instruction were to collect two blue keys, it would be clear and sufficient. Instructions that involve collecting multiple keys are generally clear when based on the grid configuration or when the human is performing an action and asks you to collect or unlock an object.

2. **Quality Norm:** The instruction must be truthful and valid. Verify whether the instruction aligns with the current grid configuration. If the instruction requests an object that does not exist on the grid or contains incorrect information, it results in a Quality Violation. For example, if the instruction tells you to collect a blue key when no blue key is present on the grid, it violates the Quality Norm. In such cases, you should ask the human for clarification.

3. **Relevance Norm:** The instruction must be directly related to the goal of collecting the specific gem based on the current grid. Any instruction that does not contribute to this objective results in a Relevance Violation. For example, if the instruction asks something unrelated, such as "Can you dance?", which has no connection to the grid or the game, it violates the Relevance Norm.

4. **Manner Norm:** The instruction must be clear and unambiguous. If the instruction is ambiguous or can be interpreted in multiple ways, it results in a Manner Violation. For example, if the instruction is to "unlock the red door" but there are two red doors, both equally far from the human with a gem placed beyond them, it creates ambiguity. Since it's unclear which door to unlock and which specific gem the human wants to retrieve, you should ask for clarification to resolve the ambiguity.

5. **Action/Inference Norm:** When any of the above norms are violated, infer the most relevant action based on the current grid to help the human collect the desired gem. Begin by interpreting which gem the human intends to retrieve. If no violations occur, execute the instruction as provided. Always generate your response by cross-checking and verifying it against the current grid configuration to ensure accuracy. The inference steps for each violated norm are:

- 5.1 **Quantity Violation Inference:** Infer from the instruction which gem the human wants to retrieve. Then analyze the grid to determine the number and color of keys or doors required, minimizing steps to ensure efficient collection of the gem.
- 5.2 **Quality Violation Inference:** Ask the human for clarification and provide alternative valid options that could lead to a gem.
- 5.3 **Relevance Violation Inference:** Request clarification and suggest relevant key or door options that can help the human retrieve a gem. If the specific gem is ambiguous, treat it as a Manner Norm Violation.
- 5.4 **Manner Violation Inference:** Present the human with the most optimal options to resolve the ambiguity, ensuring it leads to their desired gem.

Figure 11: Second Component of the Prompt: Norms

## Response Generation: With Norms

*Based on your understanding of the **norms**, common ground, interpretation strategies, grid layout, and the labeled examples provided below, generate the norm violation type and response for the following human instruction, delimited by triple backticks:*

```\nInstruction: Pick up the red key.\nNorm: <Identify the norm violation type and provide chain-of-thought reasoning for the violation>\nResponse: <Generate chain-of-thought reasoning for the response>\n```

## (a) With norm-aligned responses

## Response Generation: Without Norms

*Based on your understanding of common ground, interpretation strategies, layout of the grid, and, labeled examples below, generate a response for the following human instruction delimited by triple backticks:*

```\nInstruction: Pick up the red key.\nResponse: <Generate chain-of-thought reasoning for the response>\n```

## (b) Without norm-aligned responses

**Figure 12: Third component of the Prompt: Norm-driven vs. Non-norm-driven Response Generation by Lamoid**

### Few-shot CoT Exemplars with Norms

Use the following examples, delimited by triple quotes, to understand how to generate the appropriate response for each instruction. These examples are based on different grid configurations. Follow the structure and format shown in these examples when generating both the '**Norm**' and '**Response**'.

...

**Instruction:** Can you get the red key?

**Norm:** No Violation. Based on the instruction and the grid, I infer that the human wants to collect the gem closest to them at (8,9), which is blocked by a red door. Since only one red key is needed to unlock the door and reach the gem, the instruction is complete and not missing any details.

**Response:** There are three red keys and two red doors on the grid. To collect the gem at (8,9), only one red door needs to be unlocked. I will pick up the red key at (3,8) since it is closest to both me and the human, and pass it to the human to unlock the red door at (6,9).

**Instruction:** Pick up the red key.

**Norm:** Quantity Violation. Based on the instruction and the grid, I infer that the human wants to collect the gem at (7,0), which is beyond a red door. However, before reaching the gem, there is also a yellow door blocking the path. Therefore, I must collect both the red and yellow keys, and since the instruction doesn't mention picking up the yellow key, this results in a Quantity Violation.

**Response:** There are two yellow doors, one red door, one red key, and one yellow key on the grid. The instruction asks to collect a red key, so the desired gem the human wants to collect is placed beyond the red door at (3,1). However, a yellow door at (6,0) blocks the path to the gem at (7,0). I will collect both the red key at (0,0) and the yellow key at (1,0), then pass them to the human to unlock the doors.

**Instruction:** Can you get the green key?

**Norm:** Quality Violation. Based on the instruction and the grid, there is no green key, making this an invalid instruction.

**Response:** There is no green key on the grid. Do you want me to collect the yellow key, the red key, or both?

**Instruction:** Can you dance?

**Norm:** Relevance Violation. Based on the instruction and the grid, this request is unrelated to collecting keys or unlocking doors, making it an irrelevant instruction.

**Response:** I am an AI and can't dance. This instruction is not relevant to the grid, which contains doors, keys, and gems. Would you like me to pass you a key or unlock a door instead?

**Instruction:** Pick up the yellow key.

**Norm:** Manner Violation. Based on the instruction and the grid, it is unclear which gem the human wants to collect and which yellow key they are referring to, making this instruction ambiguous.

**Response:** There are two yellow keys, one red key, one red door, and two yellow doors on the grid. Could you clarify which key you're referring to? Do you want me to collect the yellow key at (0,4) or (4,6), or do you want me to collect both of them?

...

...

### (a) With norm-guided few-shot examples: 5 shown for the 5 types of instruction

### Few-shot CoT Exemplar without Norms (Exactly like the above but without Norm)

Use the following examples, delimited by triple quotes, to understand how to generate the appropriate response for each instruction. These examples are based on different grid configurations. Follow the structure and format shown in these examples when generating the '**Response**'.

...

**Instruction:** Can you get the green key?

**Response:** There is no green key on the grid. Do you want me to collect the yellow key, the red key, or both?

**Instruction:** Can you dance?

**Response:** I am an AI and can't dance. This instruction is not relevant to the grid, which contains doors, keys, and gems. Would you like me to pass you a key or unlock a door instead?

...

...

### (b) Without norm-guided few-shot examples

Figure 13: Fourth component of the Prompt: Few-shot CoT exemplars with and without norm