```
- If some defects found by the "solution evaluation" are
   ↪ reasonable and some are unreasonable, then your rating
   ↪ should be \(0.5\)

Next, if the "solution evaluation" points out no errors or
   ↪ defects, or all defects found by the evaluation are
   ↪ reasonable, then you should do the following things:

- Analyze whether "expression errors" exist in the "solution
   ↪ evaluation" (**expression analysis**) or whether "solution
   ↪ evaluation" gives a wrong score according to the rules for "
   ↪ solution evaluation" (**score analysis**). If yes, you
   ↪ should rate the "solution evaluation" with \(0.5\); if no,
   ↪ your rating should be \(1\)

Your output should follow the format below:

Here is my analysis of the "solution evaluation":
... // Your analysis here.

Based on my analysis, I will rate the "solution evaluation" as:
\\boxed{{...}} // where ... should be a numerical rating of the "
   ↪ solution evaluation" (0, 0.5, or 1, and nothing else) based
   ↪ on the criteria above.

---

Here is your task input:

## Problem
{question}

## Solution
{proof}

## Solution Evaluation
{proof analysis}
```

A.4. Proof Refinement Prompt

```
{proof_generation_prompt}

## Candidate Solution(s) to Refine
Here are some solution sample(s) along with their correctness
   ↪ evaluation(s). You should provide a better solution by
   ↪ solving issues mentioned in the evaluation(s), or by re-
   ↪ using promising ideas mentioned in the solution sample(s),
   ↪ or by doing both.

{proof}
{proof analyses}

## Final Instruction
Your final response should follow the format above, including a
   ↪ `## Solution` section followed by a `## Self Evaluation`
```