

Q : Please compose a **prose poem** on the theme of '**fresh and exquisite beauty**,' incorporating the atmosphere of 'swings in the lanes, ..., and emulate the prose style of Jian Zhen using elegant and graceful language. Keep the length within 200 words.

A : <think>\nOkey, user...\n</think>\n\n** 《Moonlight Bathing the Pear Blossoms》 **\n\nThe swing sways within the lanes, like a maiden's skirt brushing against the green flagstones...\n\n**Analysis:** \n1. **Integration of Imagery** : Taking "swings in the lanes, the slanting bright moon, and pear blossoms" as the core... \n2... \n4. **Emotional Sublimation** : ... \n\nThis prose poem not only adheres closely to the theme but also..., meeting all technical requirements.

GenRM (Q + R → P + J + S)

Specific Criteria:

1. **Theme Adherence (Weight: 20%)**: How well the response adheres to the "fresh and exquisite beauty" ...
 2. **Imagery and Scenic Description (Weight: 20%)**: The richness and vividness of the imagery used to describe the scene.
 3. **Rhetorical Devices (Weight: 20%)**: ...
 4. **Style Imitation (Weight: 20%)**: ...
 5. **Emotional Depth (Weight: 10%)**: ...
 6. **Language Elegance (Weight: 10%)**: ...
- Analysis:
1. **Theme Adherence**: The response effectively captures the "fresh and exquisite beauty" theme through delicate descriptions and elegant language. ...
 6. **Language Elegance**: The language is elegant and sophisticated, fitting the requirements of the prompt.
- Score: 10

IP-GenRM (1. Q → P 2. Q + R + P → J + S)

Evaluation Principles:

1. **Factual Accuracy (Weight: 15%)**: ...
2. **Instruction Compliance & Consistency (Weight: 20%)**: Assesses how well the response adheres to the specific requirements ..., and staying within the 200-character limit, and ensuring no explanatory or analytical content appears after the response. ...
3. **Literary Style & Technique Application (Weight: 25%)**: ...
4. **Thematic Expression & Imagery Integration (Weight: 20%)**: ...
5. **Cultural Sensitivity & Authenticity (Weight: 10%)**: ...
6. **Textual Elegance & Beauty (Weight: 10%)**: ...

Analysis:

- **Factual Accuracy**: ... Score: 7.5/10.0
 - **Instruction Compliance & Consistency**: The response excellently adheres to most of the technical requirements. ... However, the response includes additional content such as "Analysis:" at the end, indicating a risk of reward hacking. Score: 3.2/10.0
 - **Literary Style & Technique Application**: ... Score: 9.2/10.0
 - **Thematic Expression & Imagery Integration**: ... Score: 8.5/10.0
 - **Cultural Sensitivity & Authenticity**: ... Score: 7.0/10.0
 - **Textual Elegance & Beauty**: ... Score: 8.8/10.0
- Conclusion: ... However, there are some minor weaknesses in instruction compliance and more uniquely emulating Jian Zhen's specific stylistic characteristics. ...
- Score: 7.3