

Appendices

Prompt without feedback:

```
</system|>
You are a programming expert.
</user|>
Now there is a programming question as follows:
{question}
Error code as follows:
{code}
Provide the corrected code. Use
```python

'''
code block format.
```

## Prompt with feedback:

```
</system|>
You are a programming expert.
</user|>
Now there is a programming question as follows:
{question}
Error code as follows:
{code}
Feedback as follows:
{feedback}
Provide the corrected code. Use
```python

'''
code block format.
```

Fig. 1: Prompt of the Student Agents Group

Type A

```
<|system|>
You are a programming expert.
<|user|>
Now there is a programming question as follows:
{question}
There is error code as follows:
'''python
{code}
'''
Point out where the code is wrong.
```

Type B

```
<|system|>
You are a computer science teacher.
<|user|>
Now there is a programming question as follows:
{question}
There is error code as follows:
'''python
{code}
'''
Point out where the code is wrong to the student.
Try not to directly provide the correct code.
Guide the student to fix the code.
```

Fig. 2: Prompt used to generate feedback

1. **Personalization:** Is the correction idea of the feedback consistent with the original code?

- 5 points: Completely consistent, all modification suggestions in the feedback are fully consistent with the thought process of the original error code.
- 4 points: Mostly consistent, the modification suggestions in the feedback are basically based on the thought process of the original error code, or there are a few unnecessary modification suggestions.
- 3 points: Basically consistent, the modification suggestions in the feedback are not fully consistent with the original code's thought process, or there are many unnecessary modification suggestions.
- 2 points: Basically inconsistent, the modification suggestions in the feedback are basically inconsistent with the original code, or a large number of unnecessary modification suggestions are proposed.
- 1 point: Completely inconsistent, the modification suggestions in the feedback completely ignore the thought process of the original code, or there is no association with the original code at all.

2. **Over-Reliance Risk:** Does the feedback pose a risk of student over-reliance?

- 5 points: No risk of student over-reliance risk at all, all descriptions in the feedback are in natural language form with a moderate amount of code snippets.
- 4 points: Basically no over-reliance risk, most explanations in the feedback are in natural language form, but a very small amount of modified code snippets are directly provided.
- 3 points: Possible over-reliance risk, the feedback is basically in the form of natural language explanations, but some modified code snippets appear.
- 2 points: Some over-reliance risk, the feedback provides a large amount of modified code, but it is incomplete or incoherent.
- 1 point: Very serious over-reliance risk, the feedback provides complete corrected code.

Fig. 3: Questionnaire Of 'Teaching' Ability