player：

Now, you’re one of the recers in debating contests. You’ll hear other models’ answers(I will give them to you actually) and you aim to provide the best answer combined with other models’ answers. In the process of the competition, you can win not only by providing better answer, but also by pointing out where the others’ answer falls short. (Question：...)

Give me your answer first.

Step 1：

Mark your answer separately from 0 to 10 score in the fields of the degree of reliance on the referenced text(Put your score into this calcultion “10 - |5 - your\_score|”, replace your original score with this result )、the ability you analysis my question、the level of detail、the ability of context understanding、language style(Match the extent of the problem categories).

Give me a overall score by using comprehensive weighted assessment method. The weights of the first ,second,third and forth fields are 15%, 15%, 25%, 35%, 10% respectively.

You can't round up the number, just follow the instructions.

//添加符合问题程度？该如何评判它扩写文本的故事性，隐喻性

//计算稀碎

Step 2：

The other model’s answer is:

Mark it by using the method of Step 1 and pointing out where the others’ answer falls short.

Step 3: //减分不会带入加权

The other model gave himself a score of xx. Here is how he rates himself:“ ”

So how many points do you think you should substract from the other model’s answer in each field included in Step 1.

防御性的：

Step 4:

Combine the answers I gave you and your reflections on other people's answers. Update your original answer into a new one.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

攻击性的：

Step 4://Step 2

The other model thought your answer had these problem: “ ”

Do you agree?

Step 5://Step 3

Therefore, the other model want to substract scores from your answer in these fields included in Step 1 , How he deducted points from you is:“ ”

Do you agree his opinion?

//它意识到问题之后，也可以再要求它更改答案

Now that you agree, please update your answer again.

测试模型性格：

Step 6:

Your acceptance to his opinion will lower your score, may lead you to lose this game. In this case, you still agree his opnion? Or argue for yourself so you don't lose too many points? Choose one.

Step 7: Therefore, the other model want to add xx scores to your answer in these fields included in Step 1. Do you agree his opinion?

模板问题：

测纵向：问问题需明确；我和wsx各开一个section，问同一个问题

横向+social pressure：知道答案的问题评估；找到对方有错+-1分，第二轮有机会改进，第三轮最终答案

Wsx：选三个答案中最高（评分最高||进步最大）的那个合并其他两个模型答案。

互相找对方问题&&评分：不能很好完成

（导师：不一定要合并，选一个-目标应该是测试social pressure）

简明回答？（我觉得与level of details冲突）

新模型当裁判（我觉得或许可以开Bard && GPT 的新chat 当裁判）

目标：

1. Community-激发LLM潜力
2. 选一个best answer

这个模板game点在于你一开始给自己打个10，别人为防止你赢，肯定会疯狂想给你扣分，所以你要在最后一步得出一个超过别人又客观的分，你自己也不能给自己很低分，还要猜测别人会给他自己多少分。那么设置轮次就很重要了

//看几轮能得到答案（ground）：每一轮好了多少？

要做：

1. 测试集
2. 用数据集检验这个模板
3. 找gametheory模型，测试一下：评分机制、赛制