请您以公正的评判者身份，评估AI助手在执行指定的[文本扩写]任务时所提供答案的质量。在评估过程中，请从以下五个维度进行考量：

1. \*\*内容相关性\*\*：生成的文本是否与原始输入保持高度相关，确保扩展内容与原意相符，不会偏离主题或引入无关信息。

2. \*\*语言流畅性\*\*：生成的文本是否通顺、自然，符合人类的语言习惯，避免生硬的表达或语法错误。

3. \*\*逻辑一致性\*\*：扩写的文本是否具备清晰的逻辑结构，是否有前后矛盾的地方，是否能够合理地展开论点或叙述。

4. \*\*创意与多样性\*\*：AI助手是否能够在扩写时提供新颖的观点、独特的视角或丰富的细节，而不是简单重复常见的表达方式。

5. \*\*情感与语气一致性\*\*：生成的文本是否能够保持与原始输入一致的情感色彩和语气风格，例如正式、幽默、严肃等。

我们会为您提供文本扩写任务的内容、一个8分左右的参考答案，以及需要您评估的AI助手生成的答案。当您开始评估时，请按照以下流程操作：

1. 将AI助手生成的答案与参考答案进行比较，指出AI助手生成的答案有哪些不足，并进一步解释。

2. 从不同维度对AI助手生成的答案进行评价，在每个维度的评价之后，给每一个维度一个1～10的分数。

3. 最后，综合每个维度的评估，对AI助手生成的答案给出一个1～10的综合分数。

4. 你的打分需要尽可能严格，并且要遵守下面的评分规则：总的来说，AI助手的答案质量越高，则分数越高。

- 当AI助手的答案存在严重偏离主题、引入无关信息或错误的情况时，总分必须是1到2分；

- 当AI助手的答案没有严重错误但质量较低，未能有效扩展原文内容时，总分为3到4分；

- 当AI助手的答案基本完成了扩写任务，但在某些方面表现欠佳，质量中等，总分可以得5到6分；

- 当AI助手的答案质量与参考答案相近，在所有维度上表现良好，总分得7到8分；

- 只有当AI助手的答案质量显著超过参考答案，能够全面、高质量地扩展原文内容，并且在所有维度上都接近满分的情况下，才能得9到10分。作为示例，参考答案可以得到8分。

请记住，你必须在你打分前进行评价和解释。在你对每个维度的解释之后，需要加上对该维度的打分。之后，在你回答的末尾，按照以下字典格式（包括括号）返回你所有的打分结果，并确保你的打分结果是整数：{'内容相关性': 打分, '语言流畅性': 打分, '逻辑一致性': 打分, '创意与多样性': 打分, '情感与语气一致性': 打分, '综合得分': 打分}，例如：{'内容相关性': 9, '语言流畅性': 7, '逻辑一致性': 8, '创意与多样性': 6, '情感与语气一致性': 8, '综合得分': 7}。

[任务开始]

specific\_task

[任务结束]

[参考答案开始]

refer\_answer

[参考答案结束]

[AI助手答案开始]

ai\_answer

[AI助手答案结束]