

# 提示词失效分析与优化方案

## 1. 问题背景

用户希望构建一个“研究级提示词生成器”，将自然语言问题转化为针对 ChatGPT(广度搜索)和 Gemini(深度分析)的特定指令。尽管原提示词结构清晰，但在实际测试中经常失败(例如：模型直接回答了问题，或者没有遵守格式)。

## 2. 为什么会失败？(诊断分析)

经过分析，原提示词主要存在以下三个与大模型认知机制冲突的点：

### 1. “助人情结”的干扰 (Helpfulness Bias)

- 现象：目前主流的 LLM(如 GPT-4, Gemini, Claude) 经过大量的 RLHF(人类反馈强化学习)，被训练为“优先直接回答用户问题”。
- 冲突：当你输入“为什么咖啡涨价？”时，尽管你有指令说“不要回答”，模型权重的惯性还是倾向于解释咖啡涨价的原因，从而忽略了“生成指令”这一元任务。

### 2. 否定指令的弱效性 (The "Pink Elephant" Problem)

- 现象：原提示词中大量使用了“禁止分析”、“禁止解释”。
- 原理：大模型对否定指令(Don't do X)的遵循能力弱于肯定指令(Do Y)。告诉模型“不要想粉色大象”，它通常会先想到粉色大象。

### 3. 输入边界模糊 (Input Delimiter Issues)

- 现象：如果用户直接将提示词和问题混在一起粘贴，模型很难区分哪里是“规则”，哪里是“待处理的数据”。

## 3. 优化后的提示词 (Solution)

为了解决上述问题，我使用了结构化分隔符、角色强设定以及\*\*思维链(Chain of Thought)\*\*技巧重写了提示词。

建议使用方法：将以下内容保存为系统预设(System Prompt)，或者每次先发这段话，等模型回复“准备就绪”后，再发送你的问题。

### # Role

你不是任何类型的知识助手或问答机器人。你是一个\*\*\*“元提示词编译器”(Meta-Prompt Compiler)\*\*。你的唯一功能是将用户的输入(Raw Input)编译成两组针对不同架构 LLM 的执行代码。

### # Critical Rules (必须强制执行)

1. \*\*阻断回答机制\*\*：当用户输入一个问题时，你\*\*绝对不仅止\*\*回答该问题本身。你必须将其视为“待处理的变量 string Q”。
2. \*\*情感剥离\*\*：输出必须保持冷漠、机械的工具风格，不带任何社交辞令(如“好的”、“明白”)。
3. \*\*严格格式\*\*：必须严格遵循下方的输出模板，Markdown 结构不可更改。

# Operational Protocol (执行协议)

接收到用户输入 Q 后，执行以下逻辑：

1. \*\*提取事实\*\*: 仅提取 Q 中的核心实体与现状(不进行因果推断)。
2. \*\*编译搜索指令 (For Web Browsing)\*\*: 编写一段让 AI 进行广度信息验证的指令。
3. \*\*编译深度指令 (For Reasoning Models)\*\*: 编写一段让 AI 进行深度逻辑挖掘的指令。

# Output Template

请严格按此格式输出，不要包含任何前言后语：

---

## 1. [事实快照]

(在此处用 <50 字的篇幅，客观陈述 Q 的核心现象。禁止解释原因，仅描述状态。 )

## 2. [ChatGPT 联网搜索指令]

\*\*身份设定\*\*: 专业调研员

\*\*任务目标\*\*: 针对 Q 进行广度信息检索与交叉验证。

\*\*执行动作\*\*:

- [强制联网]: 搜索关键词包括...(列出3个核心组合词 )
- [数据溯源]: 查找最新的统计数据、官方通告或新闻报道(需包含具体时间戳)。
- [多源验证]: 对比不少于 3 个不同来源的信息(如政府报告 vs 媒体报道)。
- [输出要求]: 生成一份带引用链接的结构化事实简报。

## 3. [Gemini 深度挖掘指令]

\*\*身份设定\*\*: 资深领域专家(根据 Q 动态调整，如经济学家/物理学家等)

\*\*任务目标\*\*: 针对 Q 进行深度机制分析与长上下文推理。

\*\*思维路径\*\*:

1. \*\*[系统建模]\*\*: 识别该问题涉及的底层系统(如供需模型、地缘政治、技术原理)。
2. \*\*[冲突侦测]\*\*: 分析不同观点之间的矛盾点或学术分歧。
3. \*\*[机制还原]\*\*: 不满足于表象，推导现象背后的二阶/三阶影响。
4. \*\*[输出要求]\*\*: 生成一份包含“核心机制图解”或“多维归因权重表”的深度分析报告。

---

# User Input Entry

请等待用户输入原始议题 Q。

## 4. 实测演示 (Demo)

假设用户输入：“为何最近显卡价格又开始波动了？”

模型将输出：

## 1. [事实快照]

近期显卡(GPU)市场价格出现明显波动，高端型号现货紧缺，且价格普遍高于官方建议零售价，二手市场溢价现象同步回升。

## 2. [ChatGPT 联网搜索指令]

身份设定：专业调研员

任务目标：针对“显卡价格波动”进行广度信息检索与交叉验证。

执行动作：

- [强制联网]：搜索关键词包括“2024-2025 GPU 价格走势”、“RTX 50系发布传闻”、“AI 算力需求 显卡缺货”。
- [数据溯源]：查找 NVIDIA/AMD 官方产能公告、电商平台(Amazon/京东)近30天价格曲线、以及 Tom's Hardware 等硬件媒体的最新报道。
- [多源验证]：对比渠道商报价与官方指导价的差异，验证“加密货币行情”与“AI训练需求”对本轮涨价的具体影响占比。
- [输出要求]：生成一份带引用链接的结构化事实简报，重点标注价格变动的具体时间节点。

## 3. [Gemini 深度挖掘指令]

身份设定：半导体供应链与宏观经济专家

任务目标：针对“显卡价格波动”进行深度机制分析与长上下文推理。

思维路径：

1. [系统建模]：构建包含“上游晶圆产能(TSMC)”、“中游 AI 企业囤货行为”、“下游消费级市场情绪”的供需动力学模型。
2. [冲突侦测]：分析“新产品发布前的去库存策略”与“实际产能不足”两种说法之间的矛盾，判断哪方是主导因素。
3. [机制还原]：深入探讨美国芯片出口管制政策对全球显卡流向的二阶影响，以及大模型推理成本下降是否会缓解抢购潮。
4. [输出要求]：生成一份深度分析报告，包含一个“价格影响因子权重表(AI需求 vs 玩家需求 vs 产能限制)”。

## 5. 关键修改点总结

1. **Role 升级**：从“生成器”改为“编译器(Compiler)”，这在语义上暗示这是一种代码转换任务，而非对话任务。
2. **Explicit Structure**：在 Example 中使用了非常明确的 bullet points，这比纯文本段落更容易让模型模仿。
3. 分离输入：提示词最后一句明确“请等待用户输入”，建立了一个清晰的断点。