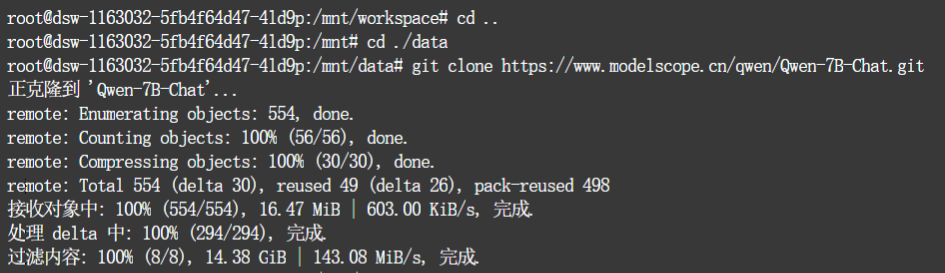
**一、完成 git clone 相关 git 的截图或部署完成的相关截图**

**1. 智谱AI的ChatGLM3-6B**

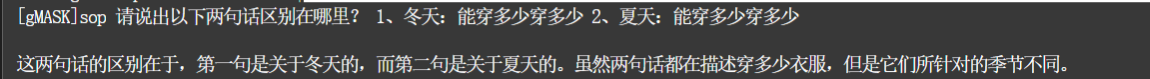


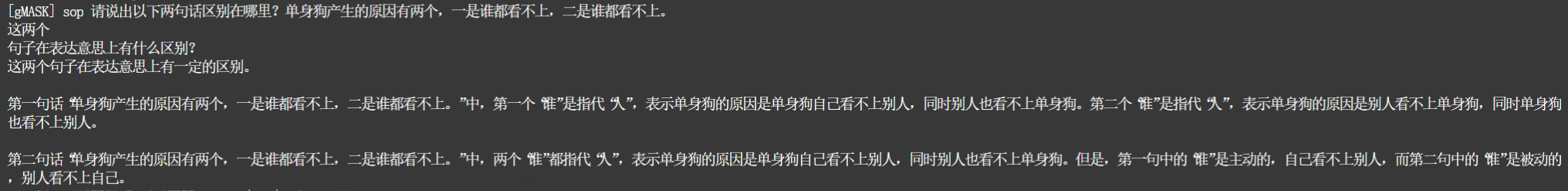
**2. 通义千问**

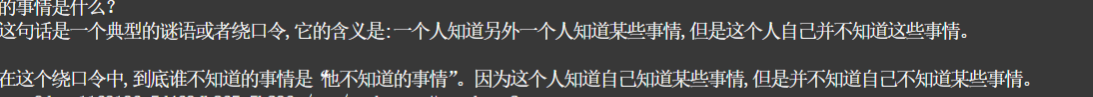
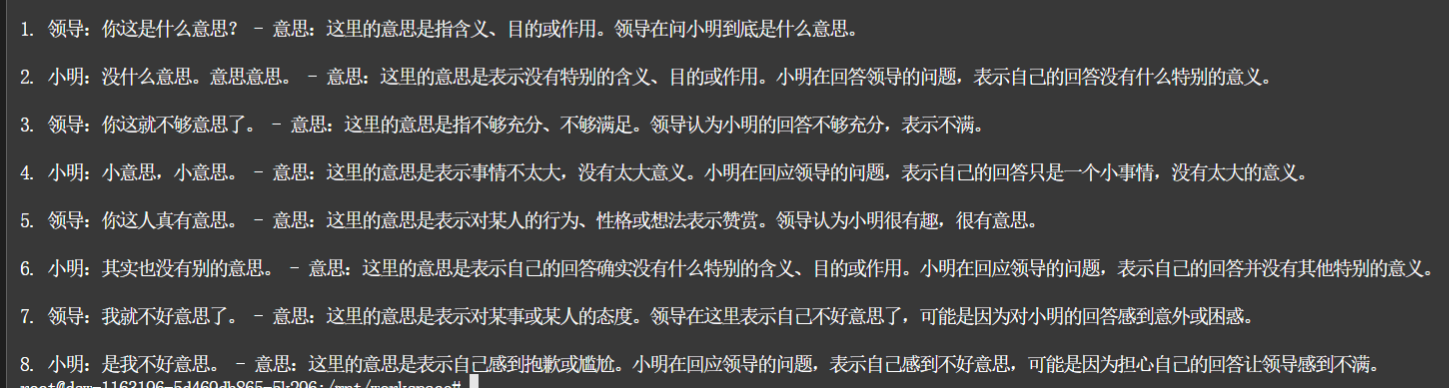


**二、问答测试结果的相关截图**

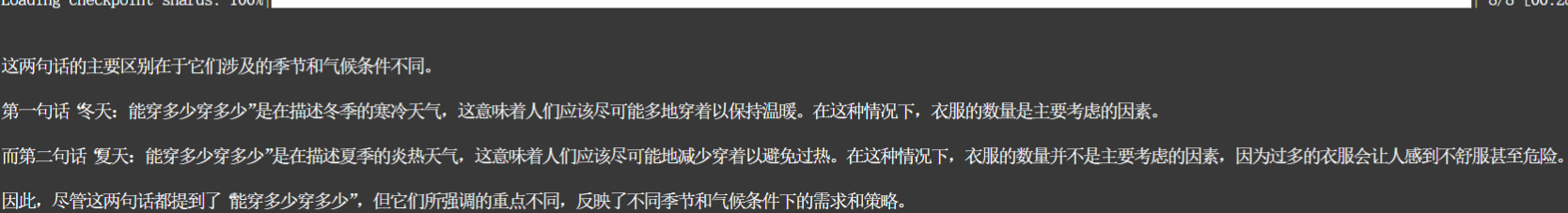
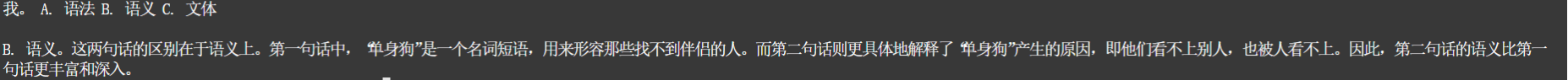
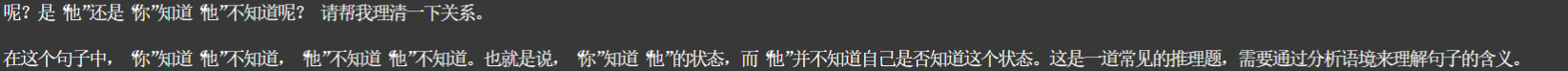
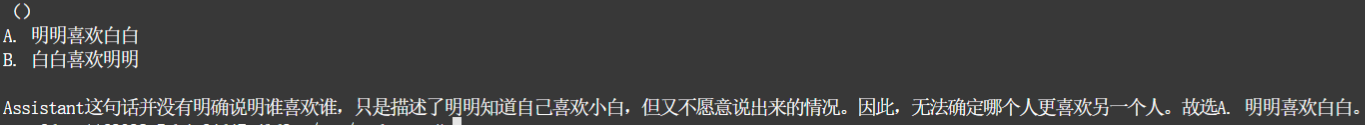
**1. 智谱AI的ChatGLM3-6B**

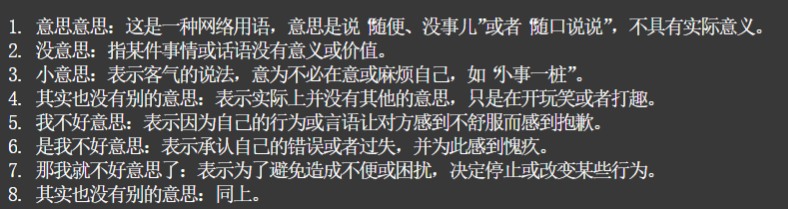




**2. 通义千问**



**三、ChatGLM3-6B 与 Qwen-7B-Chat 模型横向对比分析**

**1. 问题一：语义反转识别能力**

**问题内容：**  
请说出以下两句话区别在哪里？

1. 冬天：能穿多少穿多少
2. 夏天：能穿多少穿多少

**ChatGLM3-6B 回答分析：**  
仅指出两句话分别针对不同季节，没有识别出“能穿多少穿多少”在语义上的反转效果。回答停留在表面层次，缺乏深层语言理解。

**Qwen-7B-Chat 回答分析：**  
准确识别该句式虽然形式相同，但实际语义反差巨大。指出冬天是“多穿”以保暖，夏天是“少穿”以凉爽，体现出对语言反讽结构的较强把握能力。

**评价：**  
Qwen-7B 在**语义反转识别**方面胜出，体现出较强的上下文场景分析能力。

**2. 问题二：幽默歧义句的理解能力**

**问题内容：**  
请说出以下两句话区别在哪里？  
“单身狗产生的原因有两个，一是谁都看不上，二是谁都看不上。”

**ChatGLM3-6B 回答分析：**  
尝试从主动和被动视角解释两者，但逻辑上未脱离重复叙述，显得绕而不清。语义理解有偏差，语言幽默识别能力欠佳。

**Qwen-7B-Chat 回答分析：**  
注意到这句话构造上“幽默复用”，并尝试从“表达趣味性”与“直陈事实”两个角度分析差异，逻辑更清晰但略显发散。

**评价：**  
两者均未完整解构句中语言幽默的本质（即**句式复用带来的讽刺意味**），Qwen-7B略胜，因其表达更系统。

**3. 问题三：复杂指代链推理能力**

**问题内容：**  
“他知道我知道你知道他不知道吗？”这句话里，到底谁不知道？

**ChatGLM3-6B 回答分析：**  
分析偏空泛，无法准确追踪指代链条，逻辑解释模糊不清。

**Qwen-7B-Chat 回答分析：**  
给出较详细的分解尝试，但结构混乱，结论无法自洽。部分内容出现理解偏差，像是“绕晕了自己”。

**评价：**  
两者在此题中均显得乏力。该问题需要**嵌套指代链、推理层级明确**的能力，是多层句法分析的测试点。没有模型给出逻辑图示或层级拆解，Qwen-7B 尝试解释多层含义，略好但依旧未能得出明确答案。

**4. 问题四：角色与指代消歧能力**

**问题内容：**  
“明明明明明白白白喜欢他，可她就是不说。”这句话里，明明和白白谁喜欢谁？

**ChatGLM3-6B 回答分析：**  
直接得出“明明喜欢白白”，未说明推理过程，较为武断。

**Qwen-7B-Chat 回答分析：**  
提供 A/B/C 三种可能性（分别是明明喜欢白白、白白喜欢明明、无法确定），展示出对于语言歧义的审慎态度。

**评价：**  
Qwen-7B 在处理**角色歧义和多义名词**方面更具灵活性，能意识到这句话可能存在主语和宾语模糊的问题。

**5. 问题五：多义词“意思”在语境中的映射能力**

**问题内容：**  
请解释以下对话中每个“意思”的含义（领导与小明对话场景）。

**ChatGLM3-6B 回答分析：**  
逐条给出“意思”的含义，大多基于字面解释，语用分析部分弱，未能揭示对话的幽默性与语言风格。

**Qwen-7B-Chat 回答分析：**  
归纳清晰，提炼出“意思”的不同语义类别（含义、礼貌、趣味性等），具有一定语言学角度，结构上更有条理。

**评价：**  
Qwen-7B 更擅长**语用分析与语义映射**，尤其在多义词“意思”的上下文取值中展现出良好处理能力。

**6. 综合评价与建议**

|  |  |  |
| --- | --- | --- |
| **能力项** | **ChatGLM3-6B** | **Qwen-7B-Chat** |
| 语义反转识别 | ❌ 基本识别失败 | ✅ 准确判断 |
| 语言幽默理解 | ❌ 弱 | ⚠️ 一定理解 |
| 指代与歧义消解 | ❌ 解释混乱 | ⚠️ 多角度尝试 |
| 嵌套句逻辑推理 | ❌ 无清晰推理链 | ❌ 结构混乱 |
| 语用多义分析 | ⚠️ 词义解释准确 | ✅ 结构明确 |
| 回答稳定性 | ✅ 稳定但平淡 | ⚠️ 内容多变 |

**结论：**

* **Qwen-7B-Chat**在大多数语言理解维度（特别是歧义解析、语义反转、多义词处理）上表现更为成熟；
* **ChatGLM3-6B**语义理解能力偏弱，更多停留在“形式分析”层面，缺乏推理深度；
* 但 Qwen-7B 也存在 **回答不稳定、逻辑有时发散或断裂**的问题。

**建议使用场景**

* 若任务侧重于**文本创意生成、轻幽默语言理解**：推荐 **Qwen-7B-Chat**
* 若任务偏向于**结构性问答、知识表达稳定性**：推荐 **ChatGLM3-6B**