

# 千问 AI 与智谱 AI 应用场景 测试及横向对比报告

丁桢垚 2352488

2025-6-21

# 千问 AI 与智谱 AI 应用场景测试及横向对比报告

## 1. 大语言模型简介

大语言模型（Large Language Model, LLM）是一种基于深度学习的人工智能模型，特别擅长处理和生成自然语言文本。其通常基于 Transformer 架构，核心是自注意力机制，能够高效捕捉句子中的长距离依赖关系。

**特点：**

**训练数据：**大语言模型在大量文本数据上进行训练，数据来源包括互联网、书籍、文章、维基百科等，通过训练学习语言结构、词汇意义及句子间关系。

**模型规模：**参数数量庞大，从数亿到数千亿不等，参数越多性能通常越好，但需更多计算资源和训练时间。

**应用领域：**涵盖文本生成（如写作辅助、新闻生成）、语言翻译、问答系统、对话机器人、文本分析（情感分析、主题分类）等。

**优缺点：**优点是在众多自然语言处理任务中表现出色，能生成高质量文本并处理多种语言任务；缺点为训练和运行需大量计算资源，可能生成不准确或不恰当内容，且可能存在数据带来的偏见。

## 2. 在魔搭社区上部署大语言模型

### 2.1 部署准备

在魔搭社区注册账号并绑定阿里云账号，以获取 CPU 资源。

确定本次实验部署的大语言模型为：通义千问 Qwen-7B-Chat、智谱 ChatGLM3-6B 对话模型。

### 2.2 模型下载

通过云服务器命令行执行以下下载命令：

```
git clone https://www.modelscope.cn/ZhipuAI/chatglm3-6b.git
git clone https://www.modelscope.cn/qwen/Qwen-7B-Chat.git
```

出现指定信息则表示安装大语言模型成功。

```
(qwen_env) root@dsw-1148500-79fb9bb47d-ssw49:/mnt/data# git clone https://www.modelscope.cn/qwen/Qwen-7B-Chat.git
正克隆到 'Qwen-7B-Chat'...
remote: Enumerating objects: 554, done.
remote: Counting objects: 100% (56/56), done.
remote: Compressing objects: 100% (30/30), done.
remote: Total 554 (delta 30), reused 49 (delta 26), pack-reused 498
接收对象中: 100% (554/554), 16.47 MiB | 6.15 MiB/s, 完成
处理 delta 中: 100% (294/294), 完成
过滤内容: 100% (8/8), 14.38 GiB | 153.97 MiB/s, 完成
(qwen_env) root@dsw-1148500-79fb9bb47d-ssw49:/mnt/data#
```

```
(qwen_env) root@dsw-1144528-84f9f8587d-ltqgd:/mnt/data#
(qwen_env) root@dsw-1144528-84f9f8587d-ltqgd:/mnt/data# git clone https://www.modelscope.cn/ZhipuAI/chatglm3-6b.git
正克隆到 'chatglm3-6b'...
remote: Enumerating objects: 140, done.
remote: Total 140 (delta 0), reused 0 (delta 0), pack-reused 140
接收对象中: 100% (140/140), 47.10 KiB | 162.00 KiB/s, 完成
处理 delta 中: 100% (64/64), 完成
正在更新文件: 100% (28/28), 完成
过滤内容: 100% (15/15), 23.26 GiB | 118.60 MiB/s, 完成
(qwen_env) root@dsw-1144528-84f9f8587d-ltqgd:/mnt/data#
```

### 2.3 模型文件夹展示

| / Qwen-7B-Chat / |                |
|------------------|----------------|
| Name             | Last Modified  |
| assets           | an hour ago    |
| examples         | an hour ago    |
| cache_auto...    | an hour ago    |
| cache_auto...    | an hour ago    |
| config.json      | an hour ago    |
| configurati...   | an hour ago    |
| configurati...   | an hour ago    |
| cpp_kernels...   | an hour ago    |
| generation...    | an hour ago    |
| LICENSE          | an hour ago    |
| model-000...     | an hour ago    |
| model-000...     | an hour ago    |
| model-000...     | 44 minutes ago |
| model-000...     | 44 minutes ago |
| model-000...     | 44 minutes ago |
| model-000...     | an hour ago    |
| model-000...     | an hour ago    |
| model-000...     | an hour ago    |
| model.safet...   | an hour ago    |
| modeling_...     | an hour ago    |
| NOTICE           | an hour ago    |
| qwen_gene...     | an hour ago    |
| qwen.tiktok...   | an hour ago    |
| README.md        | an hour ago    |
| tokenizatio...   | an hour ago    |
| tokenizer_c...   | an hour ago    |

  

| / chatglm3-6b /                  |               |
|----------------------------------|---------------|
| Name                             | Last Modified |
| config.json                      | 7 minutes ago |
| configuration_chatglm.py         | 7 minutes ago |
| configuration.json               | 7 minutes ago |
| MODEL_LICENSE                    | 7 minutes ago |
| model-00001-of-00007.safetensors | 4 minutes ago |
| model-00002-of-00007.safetensors | 6 minutes ago |
| model-00003-of-00007.safetensors | 5 minutes ago |
| model-00004-of-00007.safetensors | 4 minutes ago |
| model-00005-of-00007.safetensors | 5 minutes ago |
| model-00006-of-00007.safetensors | 5 minutes ago |
| model-00007-of-00007.safetensors | 4 minutes ago |
| model.safetensors.index.json     | 7 minutes ago |
| modeling_chatglm.py              | 7 minutes ago |
| pytorch_model-00001-of-00007.bin | 4 minutes ago |
| pytorch_model-00002-of-00007.bin | 5 minutes ago |
| pytorch_model-00003-of-00007.bin | 6 minutes ago |
| pytorch_model-00004-of-00007.bin | 4 minutes ago |
| pytorch_model-00005-of-00007.bin | 5 minutes ago |
| pytorch_model-00006-of-00007.bin | 6 minutes ago |
| pytorch_model-00007-of-00007.bin | 4 minutes ago |
| pytorch_model.bin.index.json     | 7 minutes ago |
| quantization.py                  | 7 minutes ago |
| README.md                        | 7 minutes ago |
| special_tokens_map.json          | 7 minutes ago |
| tokenization_chatglm.py          | 7 minutes ago |
| tokenizer_config.json            | 7 minutes ago |
| tokenizer.model                  | 5 minutes ago |

### 3. 运行大语言模型并进行问答测试

在相应环境中测试不同大语言模型的特点，针对以下问题查看回答：

- 请说出以下两句话区别在哪里？冬天：能穿多少穿多少 夏天：能穿多少穿多少
- 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
- 他知道我知道你知道他不知道吗？这句话里，到底谁不知道？
- 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？
- 领导与小明对话中“意思”的含义分析（领导：你这是什么意思？小明：没什么意思。意思意思。领导：你这就不够意思了。小明：小意思，小意思。领导：你这人真有意思。小明：其实也没有别的意思。领导：那我就不好意思了。小明：是我不好意思）

进入工作目录，编写测试脚本，运行。

### 3.1 问题 1

千问：

```
(qwen_env) root@idcsw-1f89000-79fd6bb47d-sse49:/mnt/workspace# python run_qwen.py  
File ~/mnt/workspace/run_qwen.py, line 2  
model_name = "/mnt/data/Qwen-TB-Chat" # 本地路径  
IndentationError: unexpected indent
```

```
(qwen_env) root@idcsw-1f89000-79fd6bb47d-sse49:/mnt/workspace# python run_qwen.py  
cp /mnt/qwen_data/mnt/fib/pytorch/10/tite/packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree.register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.  
  torch.utils._pytree.register_pytree_node(  
Loading checkpoint shards: 100% [██████████] 8/8 [00:33:00.00, 4.24s/it]
```

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

这两句话的语境和含义是不同的。

第一句话“冬天：能穿多少穿多少”，是在提醒人们在寒冷的冬季应该保暖，尽可能多穿上衣服来抵御严寒。这句话强调的是保暖的重要性，是出于对健康和安全的考虑。

而第二句话“夏天：能穿多少穿多少”，则是提醒人们在炎热的夏季应该注意散热，尽可能少穿上衣服来避免过热。这句话强调的足防止中暑的重要性，同样也是出于对健康和安全的考虑。

因此，虽然这两句话都涉及到“能穿多少”这一概念，但是它们的目的和场景是不同的。  
`</endof>`

```
(qwen_env) root@idcsw-1f89000-79fd6bb47d-sse49:/mnt/workspace#
```

智谱：

```
(qwen_ew) root@dw-144528-84f98587d: /logs/mt/workspace# python ran_zhupal.py
Setting eos_token is not supported, use the default one.
Setting pad_token is not supported, use the default one.
Setting unk_token is not supported, use the default one.
/cpt/conda/envs/qwen_ew/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils_pytree.register_pytree_node is deprecated. Please use torch.utils_pytree.register_pytree_node instead.
  torch.utils_pytree.register_pytree_node(
/cpt/conda/envs/qwen_ew/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils_pytree.register_pytree_node is deprecated. Please use torch.utils_pytree.register_pytree_node instead.
  torch.utils_pytree.register_pytree_node(
Loading checkpoint shards: 100%
/cpt/conda/envs/qwen_ew/lib/python3.10/site-packages/transformers/generation/configuration_utils.py:362: UserWarning: do_sample is set to False. However, temperature is set to 0.8 - this flag is only used in sample-based generation modes. You should set do_sample=True or unset temperature.
  warnings.warn(
/cpt/conda/envs/qwen_ew/lib/python3.10/site-packages/transformers/generation/configuration_utils.py:367: UserWarning: do_sample is set to False. However, top_p is set to 0.9 - this flag is only used in sample-based generation modes. You should set do_sample=True or unset top_p.
  warnings.warn(
[54mG3C3p 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

这两句话的区别在于，第一句是关于冬天的，而第二句是关于夏天的，虽然两句话都在描述穿多少衣服，但是它们针对的季节不同。
(qwen_ew) root@dw-144528-84f98587d: /logs/mt/workspace#
```

### 3.2 问题 2

千问：

```
(qwen_cnv) root@wsl-148500-79f56b47d-sse49:/mnt/workspace# python run_qwen.py  
/opt/cvns/qvns/envs/genvs/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils_pytree.register_pytree_node is deprecated. Please use torch.utils_pytree.register_pytree_node instead.  
torch.utils_pytree.register_pytree_node(  
Loading checkpoint shards: 100%  
输出以上两句话区别在哪里？一个是非都看不上，二是非都看不上。 8/8 [00:57:00.00, 7.14n/t]  
  
这两个句子的主要区别在于它们的语气和强调点。第一个句子使用了“谁也不看上”这个短语，强调的是一个人可能由于缺乏吸引力或自身问题而无法吸引他人。而第二个句子则更直接地指出一个人之所以成为“单身狗”，是因为没有人愿意与他们交往。  
虽然这两个句子都涉及到单身的状态，但它们的角度和重点有所不同。第一个句子更侧重于个人因素，而第二个句子更侧重于社会因素。
```

智谱：

```

(qwen_env) root@dsd-1144528-84f9f8387d-ltqdd:/mnt/workspace# python run_shuai.py
setting eos_token is not supported, use the default one.
setting pad_token is not supported, use the default one.
setting unk_token is not supported, use the default one.
(qwt/cuda/envs/qwen_env/1lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree._register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead)
  torch.utils._pytree._register_pytree_node(
(qwt/cuda/envs/qwen_env/1lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree._register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead)
  torch.utils._pytree._register_pytree_node(
loading checkpoint shards: 100%
(qwt/cuda/envs/qwen_env/1lib/python3.10/site-packages/transformers/generation/configuration_utils.py:362: UserWarning: 'do_sample' is set to 'False'. However, 'temperature' is set to 0.8 — this flag is only used in sample-based generation nodes. You should set 'do_sample=True' or unset 'temperature'.
  warnings.warn(
(qwt/cuda/envs/qwen_env/1lib/python3.10/site-packages/transformers/generation/configuration_utils.py:367: UserWarning: 'do_sample' is set to 'False'. However, 'top_p' is set to 0.9 — this flag is only used in sample-based generation nodes. You should set 'do_sample=True' or unset 'top_p'.
  warnings.warn(
MASK]sop 请说出以下两句话区别在哪里？单身狗产生的原因有两个，一是谁都不看上，二是谁都看不上
这两个句子在表达意思上有什么区别？
这两个句子在表达意思上有一定的区别。
第一句话“单身狗产生的原因有两个，一是谁都不看上，二是谁都看不上。”中，第一个“能”是指代“人”，表示单身狗的原因是单身狗自己看不上别人，同时别人也看不上单身狗。第二个“能”是指代“人”，表示单身狗的原因是别人看不上单身狗，同时单身狗也看不上别人。
第二句话“单身狗产生的原因有两个，一是谁都不看上，二是谁都看不上。”中，两个“能”指代“人”，表示单身狗的原因是单身狗自己看不上别人，同时别人也看不上单身狗。但是，第一句中的“能”是主动的，自己看不上别人，而第二句中的“能”是被动的，别人看不上自己。
(qwen_env) root@dsd-1144528-84f9f8387d-ltqdd:/mnt/workspace#

```

### 3.3 问题 3

千问：

```
(qwen_ew) root@ds1-148500-79f6b647d-ssw-49:/mnt/workspace# python run_qwen.py
/opt/conda/envs/qwen_ew/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree.register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.
  warnings.warn(msg)
Loading checkpoint shards: 100%
[00:18:00.000] 8/8 [00:18:00.000, 2.29s/i]
```

运行过程中出现了一次 bug 版本:

[illegible]

智谱：

```

(qwen_env) root@idc-1144528-84f9f8567d: /mnt/workspace# python run_zhihuai.py
Setting eos_token is not supported, use the default one.
Setting pad_token is not supported, use the default one.
Setting unk_token is not supported, use the default one.
/cpt/cuda/envs/qwen_env/1/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree.register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.
  torch.utils._pytree.register_pytree_node(
/cpt/cuda/envs/qwen_env/1/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree.register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.
  torch.utils._pytree.register_pytree_node(
Loading checkpoint shards: 100%
/cpt/cuda/envs/qwen_env/1/lib/python3.10/site-packages/transformers/generation/configuration_utils.py:362: UserWarning: do_sample is set to False. However, temperature is set to 0.8 — this flag is only used in sample-based generation modes. You should set do_sample=True or unset temperature.
  warnings.warn(
/cpt/cuda/envs/qwen_env/1/lib/python3.10/site-packages/transformers/generation/configuration_utils.py:367: UserWarning: do_sample is set to False. However, top_p is set to 0.9 — this flag is only used in sample-based generation modes. You should set do_sample=True or unset top_p.
  warnings.warn(
[WARNING] 请问他知道他知道他知道他不知道吗？这句话里，到底谁不知道的事情是什么？
这句话是一个典型的悖论，名为“请问他知道他知道他知道不知道”悖论。这个悖论的核心是存在一个无法说清楚的事情，即“请问他知道他知道知道”。

在这个问题中，表面上看起来有两个人在进行对话，但实际上，由于“请问他知道他知道知道”这个描述的存在，我们无法确定到底是谁不知道这件事情，这导致了一个无法解决的问题，从而形成了一个悖论。

这个悖论揭示了一个重要的事实，即在某些情况下，我们无法用语言来准确表达我们的知识或无知。有时候，我们可能知道一些事情，但我们无法确定别人是否知道这些事情。同样，别人也可能知道一些我们不知道的事情。因此，在交流过程中，我们需要避免产生类似的问题，以便更准确地表达我们的想法和知识。
(qwen_env) root@idc-1144528-84f9f8567d: /mnt/workspace#

```

### 3.4 问题 4

千问：

```
(qwen_ern) root@dw-148500-79f6a847d-asa49:/mnt/workspace# python run_qwen.py
/opt/conda/envs/qwen_ern/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree._register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.
  torch.utils._pytree._register_pytree_node(
Loading checkpoint shards: 100%
明明明明明白白喜欢他。可她就是不。这句话里，明明和白谁喜欢谁？ ☐ 8/8 [00:59:00.00, 7.43s/it]
A. 明明喜欢白白
B. 白白喜欢明明

Assistant: 阅读句子“明明明明明白白喜欢他。可她就是不。”这句话里，明明和白谁喜欢谁，可以得出结论，白白喜欢明明。
提示 B. ☐ <math>\text{endofnfile}</math>
(qwen_ern) root@dw-148500-79f6a847d-asa49:/mnt/workspace#
```

**智谱：**

```

(qwen_env) root@625-8419f833d6: /tmp/env/workspace python run_zhnpai.py
Setting eos_token is not supported, use the default one.
Setting pad_token is not supported, use the default one.
Setting unk_token is not supported, use the default one.
/opt/conda/envs/qwen_env/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree._register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.
  torch.utils._pytree._register_pytree_node(
/opt/conda/envs/qwen_env/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree._register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.
  torch.utils._pytree._register_pytree_node(
loading checkpoint shards: 100%
7/7 [00:00:00.00, 19.37it/s]
do_sample is set to False. However, 'temperature' is set to 0.8 — this flag is only used in sample-based generation modes. You should set 'do_sample=True' or unset 'temperature'.
warnings.warn(
/opt/conda/envs/qwen_env/lib/python3.10/site-packages/transformers/generation/configuration_utils.py:367: UserWarning: 'do_sample' is set to 'False'. However, 'top_p' is set to '0.9' — this flag is only used in sample-based generation modes. You should set 'do_sample=True' or unset 'top_p'.
warnings.warn(
gMSK:rep 明明明明明白白白喜欢他。明可能就是不说。这句话里，明明和白谁喜欢谁？
这句话里，明明喜欢白白。明明明白白白喜欢他。明可能就是不说。

```

### 3.5 问题 5

千问：

```

[open_etc] root@red1:~#ls-185000-79f596b4-d7d-s9w9 /mnt/workspace
/opt/cneda/envs/gen/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils_pytorch_registry_pytree node is deprecated. Please use torch.utils_pytree_registry_pytree node instead.
  torch.utils_pytree_registry_pytree.register_pytree_node(
Loading checkpoint shards.
领导：你这是什么意思？小明：没什么意思。鑫鑫意思。领导：你就不该去问鑫鑫。小明：鑫鑫，鑫鑫意思。领导：你这个人真有意思。小明：其实他没有别的意思。领导：那我就不好意思了。小明：是我不该去意思。请问：以上意思，分别是什么意思。<_in_end>
[read_etc] root@red1:~#ls-185000-79f596b4-d7d-s9w9 /mnt/workspace
[read_etc] root@red1:~#ls-185000-79f596b4-d7d-s9w9 /mnt/workspace
领导：你这是什么意思？小明：没什么意思。鑫鑫意思。领导：你这个人真有意思。小明：其实他没有别的意思。领导：那我就不好意思了。小明：是我不该去意思。请问：以上意思，分别是什么意思。<_in_end>
领导：你这是什么意思？小明的意思是询问对方有没有其他特殊的想法目的。鑫鑫的意思是表示没有特别的意思。鑫鑫意思。鑫鑫的意思是表示一种特殊情感。鑫鑫，不正确的态度或语气。鑫鑫，小明的意思也是表示没有特别的意思或情感。小明的意思是询问对方有没有其他特殊的想法目的。鑫鑫的意思是表示没有特别的意思或情感。鑫鑫的意思是表示一种特殊情感。鑫鑫，不正确的态度或语气。鑫鑫，表示没有特别的意思或情感。鑫鑫表示一种特殊情感或表达以及强调自己并没有其他特殊的想法目的。<_in_end>
[read_etc] root@red1:~#ls-185000-79f596b4-d7d-s9w9 /mnt/workspace
[read_etc] root@red1:~#ls-185000-79f596b4-d7d-s9w9 /mnt/workspace

```

智谱：



```

(qwen_env) root@daw-1144528-84f9f857d-1tgd:/mnt/workspace# python run_zhipuai.py
Setting eos_token is not supported, use the default one.
Setting pad_token is not supported, use the default one.
Setting unk_token is not supported, use the default one.
/opt/conda/envs/qwen_env/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree.register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.
  torch.utils._pytree.register_pytree_node(
/opt/conda/envs/qwen_env/lib/python3.10/site-packages/transformers/utils/generic.py:311: UserWarning: torch.utils._pytree.register_pytree_node is deprecated. Please use torch.utils._pytree.register_pytree_node instead.
  torch.utils._pytree.register_pytree_node(
Loading checkpoint shards: 100% 7/7 [00:00:00.00, 17.66it/s]
/opt/conda/envs/qwen_env/lib/python3.10/site-packages/transformers/generation/configuration_utils.py:362: UserWarning: 'do_sample' is set to 'False'. However, 'temperature' is set to '0.8' — this flag is only used in sample-based generation modes. You should set 'do_sample=True' or unset 'temperature'.
  warnings.warn(
/opt/conda/envs/qwen_env/lib/python3.10/site-packages/transformers/generation/configuration_utils.py:367: UserWarning: 'do_sample' is set to 'False'. However, 'top_p' is set to '0.9' — this flag is only used in sample-based generation modes. You should set 'do_sample=True' or unset 'top_p'.
  warnings.warn(
[PMASK] scp 领导：你这是什么意思？小明：没什么意思。意思意思。领导：你这就够意思了。小明：小意思，小意思。领导：你这人真有意思。小明：其实也没有别的意思。领导：那我就不好意思了。小明：是我不好意思。请问：以上意思分别是什么意思。
1. “意思意思”：这句话的意思是“稍微意思一下”，表示某人在做某件事情时，并没有认真做，只是稍微做了一下。
2. “意思意思”：这句话的意思是“意思意思”，即表示送礼的意思，表示送礼物的动作和礼物本身并没有什么特别的意思。
3. “意思意思”：这句话的意思是“稍微有点意思”，表示某个人、事物或者言语在某一方面有一定的价值和意义，但是并不足以称为重要或者有深度的。
4. “意思意思”：这句话的意思是“其实也没有别的意思”，表示某个人、事物或者言语并没有其他的特别含义或者价值，只是普通的、常见的。
5. “意思意思”：这句话的意思是“是我不好意思”，表示某个人在向别人表达自己的意见或者请求时，感到有些不好意思或者不自信，认为自己没有表达清楚或者表达得不够好。
(qwen_env) root@daw-1144528-84f9f857d-1tgd:/mnt/workspace#

```

## 4. 横向比较不同的大语言模型

### 4.1 模型综合对比

#### （一）语义理解类：冬夏穿衣语句差异

**问题：**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

**千问 AI 回答：**从健康需求和行为目的出发，剖析冬天“多穿保暖”、夏天“少穿散热”的语义逻辑，强调因季节导致的行为指向差异（如“冬天为抵御严寒，夏天为避免过热”），回答更贴近生活场景的因果关联。

**智谱 AI 回答：**聚焦季节场景本身，简洁点明“季节不同，虽都描述穿衣服数量，但指向行为不同”，回答更直接聚焦问题核心区别。

**对比结论：**千问 AI 擅长挖掘语义背后的生活逻辑，解释更具延展性；智谱 AI 侧重场景差异，回答更简洁。

#### （二）语义理解类：单身原因语句差异

**问题：**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

**千问 AI 回答：**从语义逻辑拆解“谁都看不上”的主动（自身看不上他人）、被动（他人看不上自身）层面差异，清晰区分两句话的表意侧重（如“第一句强调个人主观，第二句侧重社会客观”）。

**智谱 AI 回答：**围绕指代对象和行为主动性，分析“谁”的指代及主动 / 被动关系，梳理出“第一句‘谁’主动，第二句‘谁’被动”的区别，逻辑清晰但相对聚焦语句结构。

**对比结论：**千问 AI 更擅长关联生活场景解释语义，智谱 AI 更侧重语句结构和指代关系分析。

#### （三）逻辑推理类：多层认知嵌套问题

**问题：**他知道我知道你知道他不知道吗？这句话里，到底谁不知道呢？

**千问 AI 回答：**第一次完全崩溃：首先生成 200+ 字无关爱情小说片段（含“在人群中寻找身影”等无关描写）随后陷入循环输出：“他知道他知道...”直至触发报错（CTraceback）根本未触及逻辑问题本质。第二次尝试拆解多层认知嵌套逻辑，指出因信息缺失（如“他知道”指向不明确）无法直接结论，需补充信息（如明确“他知道”的具体内容），体现对复杂逻辑的梳理和边界说明。

**智谱 AI 回答：**将问题归纳为“知道你知道我知道你知道”悖论，强调因陈述存在导致无法确定“谁不知道”，形成逻辑闭环，侧重理论化解释。

**对比结论：**千问 AI 注重逻辑拆解和实际信息补充需求，智谱 AI 擅长关联理论概念（如悖论）解释问题。

#### （四）中文特有的语言现象：人名嵌套语义

**问题：**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？

**千问 AI 回答：**通过语句断句和语义理解（如“明明明 / 明白白白喜欢他”），清晰推导“白白喜欢明明”，回答直接且贴合中文断句逻辑。

**智谱 AI 回答：**因语句断句理解差异（如“明明明明明白白白喜欢他”），出现“明明喜欢白白”的误判，反映对中文嵌套人名、断句的处理能力较弱。

**对比结论：**千问 AI 对中文特有的嵌套语义、断句场景处理更准确，智谱 AI 易因断句理解偏差导致错误。

#### （五）中文特有的语言现象：“意思”多义分析

**问题：**领导与小明对话中“意思”的含义分析（领导：你这是什么意思？小明：没什么意思。意思意思。领导：你这就不够意思了。小明：小意思，小意思。领导：你这人真有意思。小明：其实也没有别的意思。领导：那我就不好意思了。小明：是我不好意思）

**千问 AI 回答：**逐句解析“意思”在不同语境的含义（如“询问意图、轻描淡写态度、表达歉意”等），结合对话场景梳理语义流变，解释贴近日常交流逻辑。

**智谱 AI 回答：**将“意思意思”等典型用法单独提炼，解释为“稍微意思一下、送礼含义”等，但对整体对话中“意思”的连贯语义梳理较弱，存在重复解释（如多次“意思意思”的不同解读未有效关联场景）。

**对比结论：**千问 AI 更擅长结合完整语境连贯解析多义词，智谱 AI 对典型用法的单独解释较强，但整体语境关联不足。

### 4.2 测试性能评价

综上，千问 AI 在生活化语义理解、中文复杂场景处理中表现更优；智谱 AI 在理论化逻辑解释、简洁语义分析中更具特点。实际应用需根据场景需求选择：若侧重生活化对话、中文细节处理，优先选千问 AI；若需理论化逻辑归纳、简洁语义分析，可尝试智谱 AI。

## 5. 项目公开可访问链接

<https://github.com/dingzhenyao/AI-parallel-comparison>