# 大模型的初步探索

## 交互部分

用block形式，通过tab实现多种功能的分离



## 计算部分

1. 概述：通过对模型回复的观察，模型在有部分示例的情况下，解决小学数学问题，思路基本上都正确，主要问题体现在计算能力的不足。于是，我为其设计了两部分，分别实现不同的功能。
2. “简单的计算部分”：通过示例和对回复中的方程的调包计算，提高模型对特定类型问题的解决能力。
   1. 实现：
      1. 举例教会模型以特定的格式，使用乘法除法取整等能力，并且一定程度上提醒模型给出过程，实现CoT。
      2. 对第一次回复中的方程的进行提取，计算求解并将回复中的解修改为正确的解，并且截断，再次发送，得到一个语句通顺的完整回复。
   2. 难点：
      1. 对方程的提取，模型很难以特定格式回复，正则表达式也难以捕捉，只能采取‘走量‘的形式罗列多种形式可能来实现捕捉。
   3. 不足之处：
      1. 未能实现通用的对方程的捕捉，适用范围较窄。
      2. 只能解决部分算例
   4. 在算例上的实验效果：
      1. 在1~3，7、8题上有较好表现
      2. 4、6题难以解决，以第九题为例子的学习效果较差导致第十题依旧不会
      3. 第五题完全在胡言乱语，没想到任何解决思路。
3. “复杂的两步计算”：
   1. 概述：通过对least\_to\_most\_prompt方法的借鉴，在“简单的一部计算”的基础上进行优化。
   2. 实现：分三个stage
      1. stage1：通过特定prompt诱导模型回答要解决这个问题的一个子问题subq。
      2. stage2：提取原问题中的condition，让模型求出subq。
      3. stage3:在stage2的部分回复的基础上，让模型补全回答。
   3. 优势：
      1. 解决了前一步的计算结果要在后面使用的问题
      2. 能更好地让模型’think step by step’
   4. 在算例上的实验效果：
      1. 能在‘简单的一步计算‘的基础上，解决4，6题。
4. 失败的尝试：
   1. 让模型以solve()的形式包括所有的计算，便于提取出来计算，后因为模型无法理解solve的嵌套而告终。
   2. 模型在接受多个示例时，会理解为同一个例子，然后多生成一些奇奇怪怪的东西，后面尝试在每一个例子后面加特殊标记[end]来表示例子的结束，后面发现直接例1，例2效果更好。

## 简单的大模型落地尝试

1. 概述：目前提供的参数量有限的模型，我将其理解为一个自然语言和编程接口的“缝合器”，最合适的应用场景就是用一些日常任务，这里选择用自然语言实现股票开盘收盘价格查询为实现目标。
2. 实现：
   1. 用akshare库实现对实时股市信息的查询
   2. 用大模型，通过prompt诱导模型提取输入的股票名称，开始日期，终止日期等有效信息并以Json格式返回，便于使用。
   3. 通过对gradio的LinePlot方法的学习实现图像化输出。
3. 效果示例：



1. 碎碎念：
   1. 最开始的设想：语言的表达范围很强，但是冲击性不够，或者说不够‘形象‘。我想说，如果能够将文本提取关键词接入ai绘画，并配上ai语音，实现将文字高效转化为视频的功能。but，电脑垃圾跑不了stable diffusion还找不到api，上来就寄了。

## 思考

我看小说，看漫画，看视频。有一种很明显的感觉，优秀的作品数量，小说>漫画>视频。为什么？创作成本所致。而从传播速度和传播量来看，视频>漫画>小说。

这种倒挂某种程度上导致我们接受到最多的都是由视频传播的垃圾信息，而短视频靠其相对低的制作成本，成为我们收到最多的垃圾。

那大模型能不能成为破局者，将大模型落地，结合stable diffusion等ai绘画程序，我能不能抹平三者的制作成本？

再换个角度想，语言本质上是对一切真实与虚拟事物的描述，但是其本质还是人类编撰出来的‘二级信息‘，“语言的边界决定世界的外延。”

那么我们的所见，所闻，以图像和声音传播的为传播形式的信息，才是我们接触到的真正的世界，而这些信息传播路径，能否成为新的AGI路线？通过大量数据喂养，未来能否见到能像我们一样通过眼睛和耳朵感知世界的ai？希望在我有生之年能见到真正的智能涌现的那一刻。