

第一次作业

Problem 1

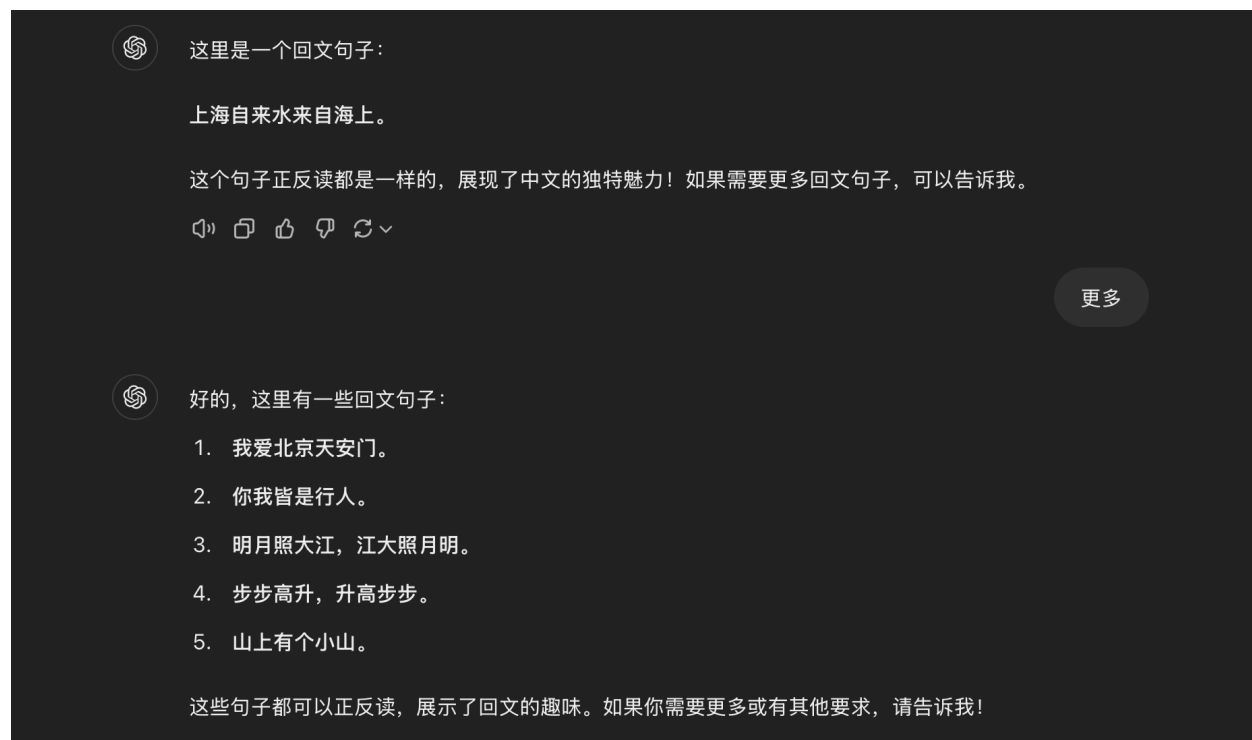
选择的测试大模型为Chatgpt。Chatgpt是OpenAI公司开发的一个人工智能技术驱动的自然语言处理工具，自推出以来就广受好评，它能够基于在预训练阶段所见的模式和统计规律，来生成回答，还能根据聊天的上下文进行互动，真正像人类一样来聊天交流，甚至能完成撰写论文 [21]、邮件、脚本、文案、翻译、代码等任务。

而在很多AI测试的排行榜中，GPT4都名列前茅，甚至可以断档领先。但我并没有ChatGPT Plus，因此本次测试是针对GPT3.5模型。

为了更全面的测试他的功能，我决定从以下三个功能去测试大模型的功能。

文字生成测试

我使用生成chatgpt生成回文句子，也就是正向和反向都相同的句子，chatgpt首先正确给出了一个结果，但在我询问更多的时候，他没有给出正确结果。



根据其他同学所说，对于回文句处理的困难很大程度上是因为chatgpt的顺序生成的模型，因此其模型就不支持反向处理文字。所以除了一些耳熟能详的案例之外，他实际上不能够创新的生成回文文字。

图片生成测试

我首先让chatgpt生成了一个手掌的图片



可以看到，大体上来说，这个手指还是符合一般人的认知的。但如果我们仔细查看，发现还有很多不合理之处：

1. 手指指腹上的指纹不合理，照片上的指纹更类似于“人”字形，而非一般的螺旋形。
2. 在手掌上的纹路也过于深，方向过于单一。
3. 手掌上部的肌肉表现不自然。
4. 大拇指似乎多了一个关节。

但是当我询问关于指纹，掌纹，解剖学的知识时。他又对答如流，正确给出的解答。我想这意味着在chatgpt生成手掌图片的时候，事实上并没有联想到这些知识，而仅仅停留在手的形象表层，而非从骨骼到肌肉再到皮肤的构造。

而之后我又测试了生成一个运动中的运动员形象。如下图所示：



但可以看到这个人两条腿的肌肉结构有问题，手指也很不自然，左侧球网在光下的投影也不正确。

而根据我对这些问题的直接提问，chatgpt显然掌握了相关的知识。但没有在生成图片的时候使用到这些知识。

科学推理

我尝试让他做高中难度的数学和物理学题目。他在几道简单的题目中表现不错，但在稍微困难一些的问题中，他就出了错，而且是频繁的犯错。而其中一个很有代表性的解答是下面这个：

设 $m_1 = 1$ 和 $m_2 = -1$:

$$\tan \alpha = \left| \frac{1 - (-1)}{1 + 1 \cdot (-1)} \right| = \left| \frac{2}{0} \right| \text{ (无法求解)}$$

题目做到此处，一个正常人类的想法应该是特殊处理或者采用别的思路绕开无法求解的问题。这意味着，chatgpt并不会在意识到他的思路错误的时候，进行大规模的回退然后寻找新的方法。而这种方法是在人类的思维中非常常见的。

总结

综上所述，我认为chatgpt大语言模型具有以下三个问题：

1. 无法进行非顺序的思考，因此无法处理回文串的问题
2. 无法进行跳跃的思考，会忽略很多表面上没有联系但其实有深度联系的内容。比如手掌的照片和手掌的解剖学。
3. 在进行了很多语言生成的之后，再遇到错误的时候，不会进行回退并从中进行学习，然后寻找更合适的思路。

据我所知，像chatgpt之类的大模型一直被视为一个巨大的黑箱，其中的Transformer模型和多层次结构中只有浩如烟海的参数值以及各种数据的连接。而我们对于这些参数所表征的意义仍然缺少了解，换言之可解释性非常的差。

因此我认为，对于目前大模型的研究仍然缺少更深度的理解，我们很难定点的通过微调来实现修改。

而且在大模型的设计上，或许目前的方法也面临很多问题，或许需要引入很多的随机性，跳脱出序列处理的方法，才会和人类的思考方式更相似。

Problem 2

因为为均匀分布，因此落到某一个区域的概率就等于这个区域占总区域的比值。

因此我们得到一个数据点 距离原点不超过 r 的概率为：

$$\frac{r^p}{1^p} = r^p$$

因此容易得到，距离原点最近数据点的距离大于 r 的概率为

$$(1 - r^p)^N$$

求中位数距离，则使上式为 $\frac{1}{2}$ ，因此可以解出：

$$r = \left(1 - \frac{1}{2}^{\frac{1}{N}}\right)^{\frac{1}{p}}$$