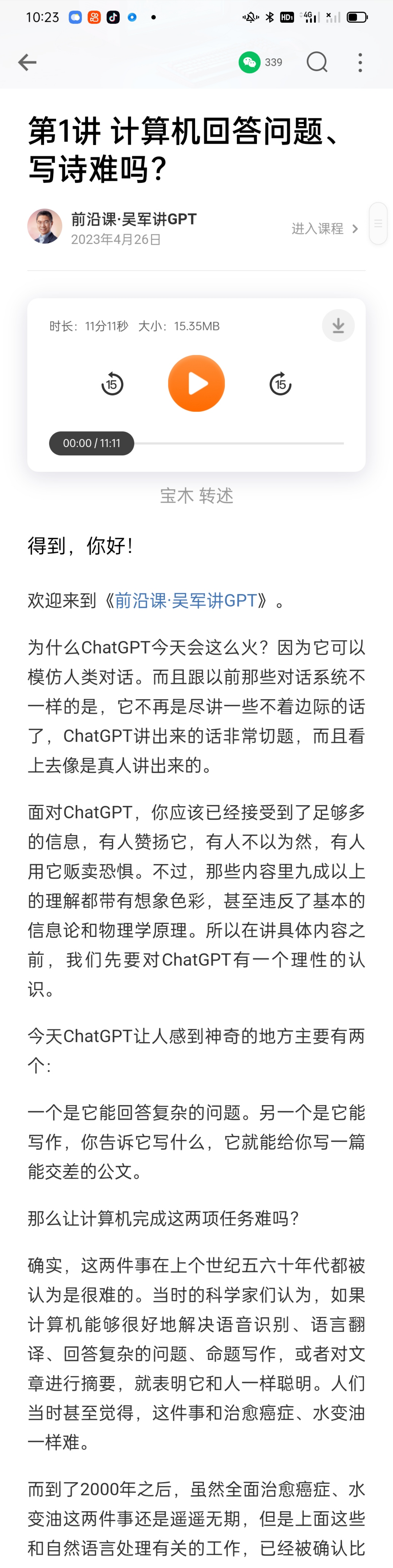

第1讲 计算机回答问题、写诗难吗？

前沿课·吴军讲GPT
2023年4月26日

宝木 转述

得到，你好！

欢迎来到《前沿课·吴军讲GPT》。

为什么ChatGPT今天会这么火？因为它可以模仿人类对话。而且跟以前那些对话系统不一样的是，它不再是尽讲一些不着边际的话了，ChatGPT讲出来的话非常切题，而且看上去像是真人讲出来的。

面对ChatGPT，你应该已经接受到了足够多的信息，有人赞扬它，有人不以为然，有人用它贩卖恐惧。不过，那些内容里九成以上的理解都带有想象色彩，甚至违反了基本的信息论和物理学原理。所以在讲具体内容之前，我们先要对ChatGPT有一个理性的认识。

今天ChatGPT让人感到神奇的地方主要有两个：

一个是它能回答复杂的问题。另一个是它能写作，你告诉它写什么，它就能给你写一篇能交差的公文。

那么让计算机完成这两项任务难吗？

确实，这两件事在上个世纪五六十年代都被认为是很难的。当时的科学家们认为，如果计算机能够很好地解决语音识别、语言翻译、回答复杂的问题、命题写作，或者对文章进行摘要，就表明它和人一样聪明。人们当时甚至觉得，这件事和治愈癌症、水变油一样难。

而到了2000年之后，虽然全面治愈癌症、水变油这两件事还是遥遥无期，但是上面这些和自然语言处理有关的工作，已经被确认比

想象得要容易得多。

发刊词里提到，2010年，Google的杰夫·迪恩等人实现了第一个真正实用的深度学习系统——Google大脑，用它构建出了第三代语言模型。

两年后，也就是2012年，我在Google启动了计算机问答的项目。经过两年的工作，到了2014年的时候，无论是利用语言模型回答问题，还是写作，都已经大部分得到了解决。我们来看看当时计算机可以做到什么样的水平。

计算机回答问题

先看计算机自动问答。这是我2014年离开谷歌前负责开发的自动问答系统，我来试着问它一个问题：

天为什么是蓝色的？

回答是：阳光在透过大气层到达地球时会发生折射，空气中的气体让不同颜色的光散射到各个方向。 蓝光因为波长短，比其他颜色折射率高。

你看，这个答案既简明、又准确，远远比绝大部分物理老师能够写出的答案要好。

再问一个和生活相关的：

怎么清洗洗碗机？

回答是：将1杯白醋装入洗碗机安全碗中，并将其放在空洗碗机的底部。将洗碗机设置为在热水循环中运行。 醋会分解任何残留的食物、油脂、肥皂渣、残留物和任何其他残留的污垢。

我们可以继续问，比如：如何烤蛋糕？从太空上看，天为什么是黑色的？

> 问题：从太空上看，天为什么是黑色的？
>
> 回答：在太空或月球上，因为没有大气可以散射光。太阳光沿直线传播，没有散射，所有颜色都聚集在一起。 朝太阳看，我们因此看到明亮的白光；而向远处看，我们只会看到虚空的黑暗。
>
> 问题：如何烤蛋糕？
>
> 回答：第1步：准备烤盘。 …
> 第2步：让配料达到室温。 …
> 第3步：预热烤箱。 …
> 第4步：将干成分搅拌在一起。 …
> 第5步：混合黄油和糖。 …

> 第6步：加入鸡蛋。 …
> 第7步：交替添加干湿成分。 …
> 第8步：将面糊倒入烤盘中烘烤。…

这几个问题的结果，都来自刚才提到的那个自动问答系统。当然，原版的问题和答案都是英文的，我用Google翻译给翻译成了中文，中间改动的字数不超过10个。因此，你可以基本上认为，上面所有的结果都是计算机产生的。

不知道你对这几个回答是否满意，我猜想，就算知识面比较广的大学毕业生，也未必能够比得过这个问答系统。事实上，这个系统虽然是我做的，但它对于很多问题的回答，显然比我个人给出的更好。

类似地，在2021年，Google的工程师开发了一个解决奥数问题的人工智能系统，它的得分比那些有计算机博士学位的开发者的得分还要高不少。

到了今天，ChatGPT产生的结果类似，质量也差不多，不过它能够回答的问题要多一些。而且，它能回答更多的“难的问题”。

什么叫难的问题？对计算机来说，所谓的容易问题，就是跟事实有关的问题，比如“是谁”“什么时候”“在哪里”“发生了什么事”等等。容易问题的回答，早在2008年前后就实现了。而所谓“难的问题“，是指”为什么”和“怎么做”这两大类问题。2014年，机器能够完美回答40%的“难的”问题，而现在，ChatGPT能回答的难的问题要多得多。

关于计算机是如何回答问题的，我们后面还会讲细节。通过这个例子，我想提醒你的是，**让计算机回答问题并不是一件高不可攀的事情，在今天，让它做得比人好是完全能够办得到的。**

计算机写作

接下来，我们来看看计算机写作。

2022年年底，我参加了美国一个媒体从业者的活动，几家大报纸的记者编辑都在。我们谈到了计算机写作的能力，他们再次确认，今天相当多的纸媒文章，特别是财经类的文章，主要是计算机写的。当然，写什么是人定的，写完之后的编辑修改也是由人完成的。在这个过程中，人和机器要来回来去进行好几次的迭代，也就是修改稿件。

你看，今天计算机的写作水平不仅超出大多数人的想象，甚至比很多人的写作水平都要

高。

但是计算机写作本身并不神秘。我在Google时做过一个“玩具”，编了个程序来写唐诗。我给你念念这个程序写的两首诗，你不妨看看计算机的水平怎么样。

第一首：

> 空愁走百川，微露贵乡还。
> 故园人不见，远望忆长安。

这首诗是怎么写出来的呢？非常简单，我只是用李白的诗做训练数据，训练了一个语言模型，然后花两天时间写了个程序。我给了它一个关键词“忆长安”，然后计算机就写出来了。

再看第二首：

> 东城淡日初破晓，薄云远帆送客棹。
> 临亭桥边渔家忙，镜山湖东春色早。

第二首诗稍微复杂一点，我用了更多的唐诗数据，训练了一个更大的模型，在写作程序中还加了一些限制和规则，然后让计算机产生了这首诗。

不过，这两首诗都有很多改进的空间。首先由于今天很多汉字的读音和古代是有所不同的，我没有使用平仄信息，这两首诗可能不符合平仄的要求。另外，当时我只是好奇，想看看计算机能写出什么样的诗，所以编这个程序花的时间不多，像词性等信息都没有加进去。如果真想让计算机写唐诗，稍微做得细致一点，写出来的会比这两首漂亮得多。

无独有偶，2019年，我和我的师兄郭毅可教授进行交流时发现，他也写了一个写唐诗的程序，原理和我的程序大致相同，我在后面还会详细介绍。

ChatGPT的能力

当然，ChatGPT能做的事情，远比我和郭教授花几天时间写的程序能做的事情多得多。但是我想说明的是，它并不是像很多局外人想得那样，一夜之间实现了很多过去没有的功能。可以说，它就是科技发展顺理成章的结果。

为什么2023之前，我们没有看到ChatGPT这样的问答产品呢？

今年3月底，我约Google负责搜索和人工智能的高级副总裁谈了ChatGPT的事情。他告

诉我——

ChatGPT能做的事情，Google的产品和服务其实三年前就能做了，只是没有包装成ChatGPT这样的产品而已。因为公司认为，有更重要的事需要人工智能来做。

此外，提供不受限制的问答服务，对于Google这样的公司来讲，还有很多法律问题。比如，Google如果就法律问题给出建议，它会惹很多麻烦，搞不好会赔上百万美元。但是至少在今天，大家还不会为难OpenAI，因为和它这样的初创公司打官司挣不到多少钱的。

那今天ChatGPT被炒得那么火，它的能力到底怎么样？实事求是说，有做得好的，也有不尽如人意的。

现在人们使用ChatGPT的目的大致可以归结为三个：

第一个是信息查询，过去Google是给出有答案的网页链接，但今天ChatGPT是直接给出答案。

第二个是让它帮助写作业，这是很多大中学生使用它的原因。

第三个则是写一些应付差事的汇报，这主要是职场上的人士使用。

这三件事的核心是两个，一个是理解自然语言，明白人的意图；另一个是产生自然语言的文本，满足人的要求。

在理解自然语言这方面，ChatGPT做得比较好，但是在产生内容方面的表现就不稳定了。有时让人眼睛一亮，有时也让人大跌眼镜，特别是当你对它刨根问底时，得到的答案、内容经常让人哭笑不得。

至于为什么它的表现不稳定，这和它之前使用的训练数据有关，也和人对它提出的不同要求有关。

虽然GPT自称是可以自己产生训练数据，但实际上，它要依赖人为产生的初始数据，才能产生新的数据。没有好的初始数据，它产生的数据也是乱七八糟的，当然，训练出的模型质量也就无法保证了。

比如我们前面提到，计算机很早就能帮助人类写财经文章，这是因为这方面的数据多，而且这类文章格式固定。今天一些人觉得，用ChatGPT写每周汇报很方便，也是因为这种汇报的格式很固定。关于数据和格式的重要性，我们后面还会讲到。

另外，两位刚毕业的从事机器学习的博士和我讲了如何更好地利用ChatGPT写作，这个我们后面还会讲到。大家可能也注意到了，ChatGPT写出来的文章好坏可以相差很远，这里面是有技巧的。

由于ChatGPT是一个新鲜玩意，人们常常会在它表现好的时候给它点赞，而那些不好的表现，就被不知不觉地给忽略了。

关于ChatGPT所产生的内容质量，微软、Google等公司内部都做了详细的评测，当然它们都没有公布结果。不过，根据《大西洋月刊》的报道，ChatGPT今天更像是一个玩具，而不是一个能够替代人的工具。

下节预告

由于ChatGPT是基于语言模型的自然语言处理系统，它的优势和局限性都来自语言模型本身。因此，我们要了解ChatGPT，就需要多少知道一些语言模型的原理。

下一讲，我们就来说一说，语言模型这个出现了很多次的概念，到底是什么。我是吴军，我们下一讲见。

划重点 添加到笔记

1.ChatGPT是基于语言模型的自然语言处理系统。
2.让计算机回答问题并不是一件高不可攀的事情，在今天，让它做得比人好是完全能够办得到的。
3.我们让ChatGPT做的事情，核心有两个，一个是理解自然语言，明白人的意图；另一个是产生自然语言的文本，满足人的要求。

《前沿课·吴军讲GPT》
成为搞懂GPT的少数人
作家，计算机科学家
吴军
版权归得到App所有，未经许可不得转载

首次发布: 2023年4月26日