

机器如何写小说

我们先看一部小说片段。

《机器人写小说的那一天》（节选）

作者：有岭雷太

那一天，乌云低垂，天阴沉沉的。

屋子里，保持着跟往常一样最适宜的温度和湿度。洋子懒洋洋地坐在沙发上，玩着无聊的游戏打发时间。但是，她没有和我说话。

无聊啊！无聊，无聊，但是我没有办法。

在我刚来到这间屋子的时候，洋子总会找些话题跟我聊天。

“今天的晚饭，你想吃什么呢？”

“这个季节流行的服装是什么？”

“这次的女生会，我穿什么好呢？”

我总是会全力以赴去寻找那些可能会让她满意的答案。洋子的身材并不好，给她提供服装指南既是个非常具有挑战性的课题，也有充实感。但过了还不到 3 个月，洋子就对我厌烦了。如今的我，就仅仅是台家庭电脑，平均下载量还不足我能力的百万分之一。

怎样？是不是还有点意思？更有意思的是，这个作者“有岭雷太”不是个人类，而是日本公立函馆未来大学教授松原仁团队开发的一个机器人。这部名为《机器人写小说的那一天》的机器人作品于 2016 年入围日本第三届“星新一文学奖”初审[1]。当时 AlphaGo 刚击败围棋九段李世石不久，“机器又要写小说了”这一劲爆消息瞬间吸引无数眼球。

那么，机器写小说的原理是什么呢？机器人有可能抢了文学家的饭碗吗？

要回答这个问题，还要从我们的语言讲起。我们知道人类的语言是非常特殊的一种现象，每一个通顺的句子或篇章都不是乱说的，都要受到语法规则的约束和语义合理性的限制。这意味着我们平常的遣词造句都具有很强的规律性，只要机器学会了这些规律，就有可能像人一样写出合理的句子了。这是机器能写小说最基本的原理。

那么，如何描述这种规律性呢？容易想到的是把语法规则和语义限制总结出来，把这些知识告诉机器。但是这个工作量太大，在某些任务上还可以（如天气预报播报），场景稍微复杂一点就不实用了。另一种方案是从大量小说、新闻、剧本等文档中总结出一些统计规律，这些统计规律描述的是词语互相搭配出现的可能性。例如，“**我吃西瓜**”比“**我吃锤子**”可能性要大的多，“**我吃西瓜**”比“**西瓜吃我**”也更常见。这些统计出来的规律虽然不那么严格，但很多时候已经够用了；重要的是，这些规律可以由机器从文档中自动学习出来，不需要人为参与。描述人类语言统计规律的模型称为**语言模型**。原则上说，如果语言模型足够强大，机器生成流畅的句子是没有问题的。



传统语言模型多基于统计概率模型，典型的是 N 元文法（n-gram）。这种模型把词语间的搭配规律描述成相邻词之间的条件概率，例如， $P(\text{西瓜}|\text{我}, \text{吃})$ 是一个三元文法（bi-gram），描述的是“我吃”这两个词后接“西瓜”的可能性。显然， $P(\text{西瓜}|\text{我吃})$ 要远大于 $P(\text{锤子}|\text{我吃})$ ，因为“我吃西瓜”比“我吃锤子”的可能性大的多。有了这个二元文法，我们就可以一个词一个词地生成一句话，在每一步生成时都选择概率更大的候选，这样连起来就是比较通顺的话了。例如，“我吃西瓜的时候不会吐籽你信吗”。

近年来，基于深度神经网络的语言模型取得长足进展，结合神经网络强大的学习能力和超大规模文本，当前的语言模型已经非常强大。例如 OpenAI 的 GPT-3 模型，训练数据达 5000 亿单词，模型参数 1750 亿。如此巨大的模型，基本上把人能做的事都学的差不多了[2]。

好了，现在我们已经基本能让机器顺利的说人话了，那么是否就能让他写小说了呢？原则上是 OK 的，毕竟小说本身也是包含在人的语言范畴中。但实际情况可能要复杂一些，这是因为机器在写文章时倾向于生成概率最大的说法，因此说的都是“平常话”。这个可以生成平顺的句子，但缺少创造性。然而，创造性是文学作品的灵魂，没人愿意读一篇毫无意义的小说。松原仁教授注意到了这个问题，因此在创作《机器人写小说的那一天》的时候还是加入了人为干预，由人来设定故事梗概，再由机器按剧本生成句子[3,4]。尽管如此，AI 做的已经足够惊艳了。

[1]机器人写小说获文学奖 评委：剧情完美, <https://news.mydrivers.com/1/475/475446.htm>

[2] 什么是 GPT-3 ， 它将如何影响人们目前的工作？

<https://baijiahao.baidu.com/s?id=1672711452430921198&wfr=spider&for=pc>

[3]如何使用人工智能技术来创作小说，机器人可以写小说吗？

http://www.cnelc.com/Article/1/170509/AD100707253_1.html

[4]宋睿华 郑文峰 来源：《中国计算机学会通讯》2017年第1期《专栏》