

AI 如何成为诗人？

诗词是中华民族的文化瑰宝。不论是“长风破浪会有时，直挂云帆济苍海”的豪迈，还是“月上柳梢头，人约黄昏后”的婉约，不论是“衣带渐宽终不悔，为伊消得人憔悴”的刻骨深情，还是“肝胆洞，毛发耸，立谈中，死生同，一诺千金重”的男儿义气，诗词伴随着中华文明走过了五千年的风雨历程。

诗词为什么有如此强大的精神感染力？一个很重要的原因可能是情感与音律的共振式表达。我们知道诗词都是有一定格式限制的，这些限制包括字数、节拍（律）、合音（韵）等，这些限制使句子具有音乐美感。另一方面，每一首能够传承下来的名篇佳作都充满了作者的真情实感，因此很容易让读者有身临其境的代入感。基于诗词的上述特点，诗词创作具有极高的挑战性，只是极少数才情俱佳的诗人才能写出打动人心作品。想要机器写出一首好诗来，还是很困难的。

如果把诗词创作比作跳舞的话，第一步首先需要按规则把动作摆出来，然后才是熟练和优美。因此，早期的诗词生成方法多是机械式的，以能按格律要求生成连贯的句子为主，缺乏对词句语义的理解。近年来，基于深度学习的自然语言处理技术兴起，以语义为基础的诗词生成方法逐渐成为主流，极大提高了生成的自然度和连贯性。

1. 机械式诗词生成方法

早期的生成算法大多解决“诗言诗语”的问题，即按格律要求，将合适的字和词按合理的顺序组合起来。典型的是词句拼凑法：给定一个主题，在大量现有诗词中搜索相关句子，将这些句子打碎后，挑选可以连接在一起的片段，基于诗词规则组合起来，即生成一首新诗。

拼凑法的基本理念是“熟读唐诗三百首，不会作诗也会吟”，只要见过的样例足够多，就可以通过对已有样例的重组来生成新诗。然而，这种机械拼凑的方法显然过于机器化了，既没有对句子意思的理解，也没有对规则的学习，因而生成的诗有明显的“破句”，且语义不清晰，不连贯，欣赏价值较低。

2. 统计模型法

为了提高句子的流畅度，研究者提出各种概率模型方法，其中统计机器翻译模型（SMT）最具有代表性。在这种方法里，从一句诗到下一句诗被认为是个机器翻译过程，这一过程既考虑了句子间的对应，也考虑了生成句子的连贯性，因此在一定程度上提高诗词生成的质量。然而，这一方法仅利用了字面上的统计规律，没有形成对语义的理解，生成的诗句依然有明显的斧凿痕迹。图 2 给出微软的 SMT 系统生成的一首诗，其中输入的主题词为“醉”。



图 1: 基于 SMT 的诗词生成, 主题词为“醉”。图片来自微软的 SMT 线上作诗系统 (<http://duilian.msra.cn/jueju/>)。

3. 基于深度神经网络的诗词生成

机械式生成方法对语义缺乏理解, 生成的诗词语义连贯性不强, 容易出现没有意义的“破句”。近年来, 神经网络 (DNN) 的出现使得对句子进行语义编码成为可能, 由此启发了基于深度学习的诗词生成方法。和拼凑法不同, 该方法将用户想要生成的内容通过 DNN 映射到一个语义空间, 基于该空间中的语义表达进行诗句生成。

2014 年, 爱丁堡大学的 Zhang 等人首先提出了基于神经网络的古诗生成方法[1]。他们首先用拼凑法生成第一句诗, 之后以第一句诗为条件, 利用一个递归神经网络 (RNN) 一步步生成余下的句子。上述生成过程是分步式的, 不利于扩展到宋词等较灵活的体例。Wang 等人提出了基于注意力机制的序列对序列模型来解决这一问题[2]。这一模型将整首诗看成一个完整的汉字序列 (包括断句符号), 利用 RNN 逐字生成整个序列。图 2 是基于该模型生成宋词的一个例子。

菩萨蛮
哀箏一弄湘江曲，
风流水上人家绿。
小艇子规啼，
不堪春去时。
花前杨柳下，
红叶满庭洒。
月落尽成秋，
愁思欲寄留。

图 2: 基于序列对序列模型生成的宋词

2016 年 3 月，清华大学语音语言技术中心做了一组实验，实验中该中心研发的作诗机器人“薇薇”和一些网络诗人就同一主题进行创作，并邀请北大、社科院等单位的诗词专家进行品评。实验结果发现，在得分最高的十篇作品里，薇薇的作品占了三篇，而且得分最高的一首作品正是薇薇创作的。同时，该研究还发现，薇薇的作品中有 31% 被专家认为是人写的。因此，研究组宣布薇薇通过了图灵测试。

近两年来，基于神经网络的诗词生成方法取得长足进展。得益于大量诗词数据的积累，当前诗词生成已经比较成熟，出现了包括“九歌”在内的优秀做诗系统[3]，至少生成较为平顺的句子是没有太大问题了。然而，诗词本质上还是个人情感的表达，只会生成应景的平顺句子还无法达到优秀诗人的境界。如何让机器生成真正具有创造性的诗句，这不仅是诗词生成本身的问题，更是对人工智能边界的探索，还需要仔仔细细去研究。

1. X. Zhang and M. Lapata, "Chinese poetry generation with recurrent neural networks," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 670–680.
2. Qixin Wang, Tianyi Luo, Dong Wang, Chao Xing, "Chinese Song Iambics Generation with Neural Attention-based Model", IJCAI 2016
3. <https://jiuge.thunlp.cn/>