

基于马尔科夫链的音乐模仿和生成

叶昊洋 袁愉清 熊珏翔 白天宇 唐元宸

May 30, 2022

1 一阶马尔科夫链（无时值版本）

本节主要仿照课上所讲，用了一阶马尔科夫链进行乐曲生成，音频、乐谱、代码均在附件文件夹中，正文不再赘述。

2 高阶马尔科夫链（无时值版本）

本节对四四拍曲目分别使用一阶二阶四阶马尔科夫链，讨论不同阶马尔可夫链对于音乐生成的影响。

2.1 基本想法

关于时值的研究将在下一节讨论，本节我们选取节奏较为简单的《东方萃梦想》的部分片段，该片段仅有四分音符和八分音符以及少量出现十六分音符的片段，将 midi 文件分解并将伴奏部分简化，得到原曲的乐谱与 mp3 文件（文件【原曲】）。

2.2 概率转移矩阵

为了更好的符合节拍，我们取原曲主旋律出现的 11 种音高分别对应八分音符与四分音符，以及 3 个两拍的相对无规则的音符片段，共 $11 \times 2 + 3 = 27$ 个变量。生成一阶、二阶与四阶概率转移矩阵。其中四阶转移矩阵太过庞大（282,429,536,481 个元素），我们使用数据结构 *map* 进行存储与处理。整个转移矩阵较大，绝大多数元素为 0，可视化难度较高，此处便放弃展示。

2.3 乐谱生成

生成音乐过程中，以八分音符为单位，会出现因四分音符跨越重音拍的情况，此处强行将四分音符压缩为八分音符，若为特殊片段，则直接删除。以此方式生成 16 小节并强行用 A_4 全音符作为第 17 小节结尾，得到主旋律。

伴奏以二分音符为单位，统计两拍中出现的音高，枚举和弦并比较相符合的程度生成伴奏，若同一小节的两个二分音符相同，则以 80% 的概率合为全音符。

生成的乐谱和 mp3 文件均在附件文件夹中（文件【一/二/四阶马尔科夫链】）

2.4 结果分析

聆听了三份音乐，我们可以发现一阶马尔科夫链生成的音乐与原曲几乎没有相似（仅风格类似）的片段，二阶马尔科夫链之中已经出现部分和乐曲相似的片段，而四阶马尔科夫链已几近和原曲一致。（其中 15 和 16 小节全为 A_4 ，已无法生成其他音高，故我们将旋律在第 15 小节终止）

我们猜想，当概率转移矩阵中 1 的出现次数较多时，即某个状态必定转移到另一个状态的情况数较多时，容易生成两个乐曲片段相似甚至相同的情况。

为验证此猜想，我们统计了三个概率转移矩阵中 1 的出现频率，结果显示，一阶马尔科夫链转移矩阵的 1 的出现频率仅为 6.8%，而二阶便达到了 58.3%，四阶甚至达到了 99.8%。与我们提出的猜想相符合。

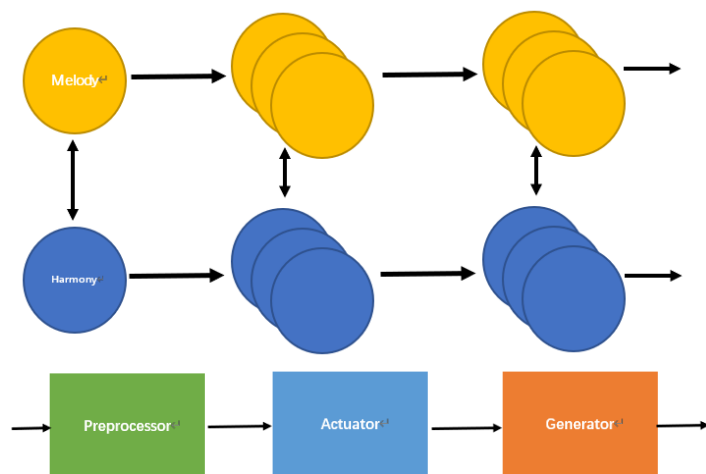


Figure 1:

我们采用一种平行式生成的方案，用两种不同概率转移矩阵来生成音乐并在时值上尽可能对齐，使其小节的意义更加清晰

2.5 结论

对于我们的实验结果，我们作出结论，马尔科夫链的阶数越高，生成音乐与乐曲的相似程度越高，到达某个阶数后，生成的音乐将与原曲完全一致。这一结果不难理解，随着阶数的增加，所需要统计的概率值成指数增加，为了得到正确的分布，所需要的样本量也会成指数增加，对于较短的片段，高阶马尔科夫链可以很轻易的做到过拟合。于是在使用马尔科夫链生成音乐时，选取恰当的阶数，是让生成的音乐既不与原曲高度相似，又能提取出原曲一定优美旋律的关键。对于阶数较高的马尔科夫链，作为音乐生成可能效率不高，但在做音乐风格识别的时候，得益于大样本的支持也许可以有很好的效果。

2.6 反思不足

对于结果，我们进行了反思，我们选取的乐曲片段仅 25 个小节，样本空间较小，可能会造成实验结果的偶然性。此外，对于生成乐曲音符的强制改动，同样会对结果造成一定的影响。由于时间原因，未能对此试验进行优化，此后的学习中，我们将继续对此的探索。

3 考虑时值、和弦的马尔科夫链音乐生成

本节研究考虑时值和和弦和声的一阶马尔科夫链音乐生成。

为了让马尔科夫链能够运用到时值上，我们统计了每个音符自身可能的时值概率。而为了能够产生和声，我们一方面将 midi 文件进行分解，分为 melody（旋律）和 harmony（和声）部分，分别统计概率转移矩阵进行生成音乐；另一方面，我们对旋律线上的音符时值进行可能的重叠，从而让节奏和音响结构上富有更多可能性。

3.1 基本想法

我们使用一种平行式生成的方法来组合出一种新的歌曲，架构和流程显示在 Figure 1 中。具体的生成流程分为 Preprocessor + Actuator + Generator，分别进行文件预处理和统计以及最后的生成，同时我们也集成了可视化概率转移矩阵的 visualization 模块。

因为和声部分的相对简单性和普适性，忽略和声和旋律匹配有关的概率转移矩阵不会对生成歌曲的总体和谐性造成很大影响，而针对 Halo 歌曲的两种生成结果（详细乐谱请看附件，01 表示关联生成、02 表示无关联平行生成）也证实了这一观点。

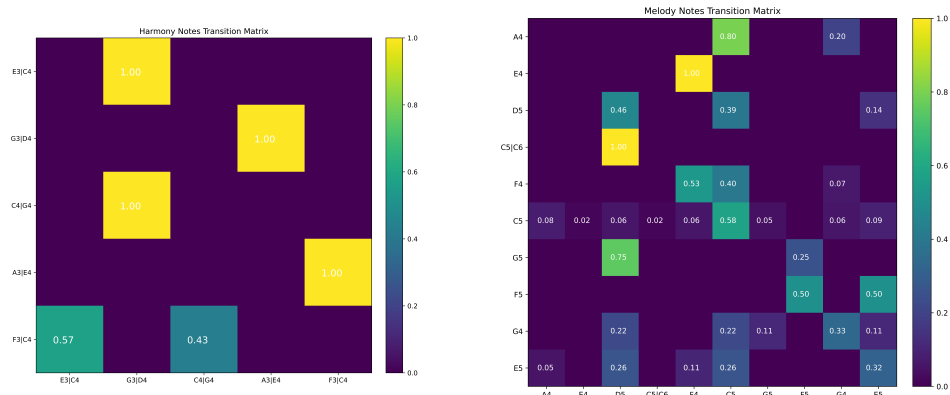


Figure 2:
All Too Well 的和声概率转移矩阵（左）和旋律转移矩阵（右），可以看到两者分布都具有一定程度的对称性，但在多样性上有巨大差异

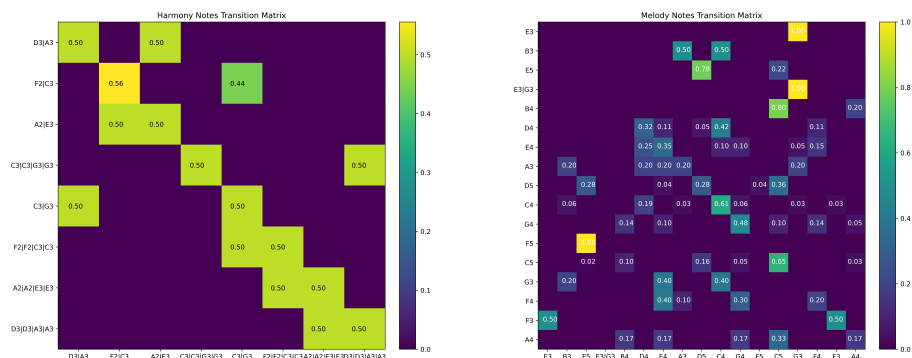


Figure 3:
Halo 的和声概率转移矩阵（左）和旋律转移矩阵（右），图像总体分布特征和 ATW 类似

3.2 概率转移矩阵

为了更好地分离旋律与和声伴奏，我们主要选取流行歌曲的 midi，我们分别统计了三首歌（Taylor Swift——All Too Well、Beyonce——Halo、Harry Styles——Late Night Talking）的音符概率转移矩阵和时值转移矩阵。每种类型的矩阵都分别针对旋律和和声伴奏分别进行了统计然后平行生成。

3.2.1 音符概率转移矩阵

Figure 2 和 Figure 3 是前两首歌音符概率的可视化结果，完整结果在附件中，可以看到和声部分的多样性明显低于旋律部分，但两者都具有一定对称性，可以想见这样一种简单对称性是听感上容易被大众接受的一个原因或者说是 pop 的固有特征，而对称性的破坏则为旋律和音响结构的多样性带来可能。

3.2.2 时值概率转移矩阵

时值的概率转移矩阵并没有被可视化。这里有两种方案，一种是统计时值之间的转移概率，另一种则是统计音符和其相应时值的概率转移矩阵。两种方法各有优劣，第一种能够更好地保有原先的节奏性，但忽略了和音符间的联系；第二种则和音符结合更紧密，在风格上可能和原曲会更加相似，但由于音



Figure 4:
乐谱结果一览，完整结果见附件

符数量众多，音符-时值概率转移矩阵相对比较稀疏。我们主要采用第二种方式，因为通过音符之间的转移矩阵，仅考虑音符和时值的转移矩阵也可以一定程度上保有原先的节奏型从而在相似度和多样性上有更好的平衡。

3.3 乐谱生成

我们在 Figure 4 中展示了三首歌共五种的乐谱结果。

Halo 的第一份乐谱表示关联生成的结果、第二份表示平行式生成的结果。

ATW 的第一份乐谱是运行了两次旋律生成和一次和声生成的结果（代码当时有 bug），但结果意外的还不错就放上来了，第二份乐谱是正常生成的结果，因为和声部分本来可能性就比较小，所以使用平行生成的方案没有让音乐变得过于不和谐。

LNT 的乐谱因为和声部分有较高的复杂性，因此在平行式生成后可以发现与原曲有较大的偏差，同时不和谐的和声出现频率也有所增加。

4 结论

此次实验主要探索研究了一阶高阶马尔科夫链以及考虑了时值和和声的马尔科夫链音乐生成模型，探究了他们各自的特点以及在风格和相似度上的情况。高阶马尔科夫链的结果和预期一致，随着阶数的提高，相似度和不断升高，直到最后和原曲一致。而在时值和和声的马尔科夫链方面，因为组合的多样性，方案数可以有很多，在多样性和相似性上不同人可以做出不同的平衡，本文主要使用平行的生成方案，在生成结果多样性较为丰富，但和原曲的相似性以及整体和谐性的角度可能表现略差，除此之外，也简单对其他生成方案进行了尝试并进行了对比。