

Music Generation from Statistical Models

In this paper Conklin presents several statistical methods for the analysis and generation of music. He argues that the traditional breakdown of music into the categories of analytic, where one studies the competence of a work, and synthetic, where one studies the generation, is misguided. He shows this by describing several current statistical approaches that can effectively describe or classify music. He finishes the argument by showing how these same models can also be used in the synthesis of new music and presents several techniques for doing so.

The goal of a statistical model is to assign a probability to any given piece of music with the property that any piece that is of high quality or likely to be in the class/style being modeled. For example, a good statistical model of Jazz should assign a high probability to a piece by John Coltrane and a low probability to a piece by Steve Reich. Actual parameter values for an individual model are trained using a set of example instances and evaluated using held out data or using a cross-validation scheme.

The most common modeling techniques are context models, which include Markov, hidden Markov, n-gram, and finite state models. They are popular because there are well established methods for training, they allow efficient processing, it is easy to compute the overall probability from them and it is easy to generate new music from them. However they also suffer several drawbacks. Although they are easy to train they require sufficient training data which in general there is never enough, and requires complex smoothing techniques to obtain decent results. Also while good at local dependencies these types of context models are not effective at capturing long distance dependencies. Other, more complex models can be used to try to alleviate some of these problems. History based models, which can “remember” earlier vantage points to use in conditioning the current event, can be used to help smooth data. Other models such as probabilistic context free grammars and dependency grammars can also be used to try to capture more long distance relationships. However, with these models comes more complexity in terms of computation, development and in the model.

The previous sections outlined how statistical models could be used to classify genres of music or provide a fitness function of sorts as to the quality of a particular piece. The following sections described four different methods for using these models to produce new music from these models. The first is simply by taking a random walk through the model. Although this approach produces optimal decisions locally it can be shown that it is not guaranteed to produce an optimal solution globally. Using Viterbi decoding is a way to improve on random walks by using a dynamic programming solution that is capable of finding a more globally optimal solution. However, time complexity becomes an issue and not all models are easily adapted to Viterbi decoding. Other techniques include Gibbs sampling and pattern based sampling which try to address these issues.

There are two aspects that are relevant to the discussion that I do not feel

are particularly well addressed in the paper. The role in which humans influence the modeling process is only mentioned in passing when discussing the training method for the models. However, it is of extreme importance to the ability of the system to do what it is designed to do, because the model is only able to learn distinctions in the music if humans are able to consistently make the same decisions. Namely that the classes trying to be trained are well defined and humans have no problem separating them, and thus preventing “noisy” training data. It is nice to have a theory of why A is different than B but if given unlabelled instances of pieces and two people can not agree on the classification the computer will not be able to either. Also something mentioned but not well addressed is the ability for the a system to deviate from the norm as a potential for making new and interesting work. A statistical model such as these would find it very difficult to create a “Classical” work that incorporated “Jazz” themes because of the low probabilities it would assign.