

0.1 Markov Chains

Markov Chains (MC) are used to model specific random processes in which transitions between various states occur. The most important property that these processes must possess, is that the transition between one state to another must be *memoryless*. This means that the probability of going from one state to another is independent of the previous states. This chapter introduces the theory behind MCs and also describes how they can be applied to the domain of algorithmic music composition.

Let S be a set of states and assume that there is some process by which random states are selected at discrete time intervals. The sequence of random variables X_1, X_2, \dots represent the random state that is chosen at time interval $1, 2, \dots$, respectively. A realization of all these random variables $s = (s_1, s_2, \dots)$ is called a *history* which thus represents the states chosen at each time index. The sequence of random variables is called a *Markov Chain* if they satisfy the *Markov Property*:

$$\Pr(X_{n+1} = s | X_1 = s_1, X_2 = s_2, \dots, X_n = s_n) = \Pr(X_{n+1} = s | X_n = s_n) \quad (1)$$

This property states that the probability of going from one state s_n at time n to another s_{n+1} at time $n + 1$ is dependent on only the previous state s_n and not the history $s = (s_1, s_2, \dots, s_n)$. Further, if the transitional probability from a state $y \in S$ to $x \in S$ is independent of n , then the MC is said to be *stationary*. These stationary markov chains can be described using a single matrix P whose entries correspond to transitional probabilities, i.e. $P_{ij} = \Pr(X_2 = j | X_1 = i)$.

Markov Chains have been used to model the process of music composition by capturing information about transitions between note events in order to generate random pieces with similar structures. One could define the states of the system using the pitch class values, i.e. $S = (C, C\#, D, \dots, B)$. The transitional probabilities of a piece of music would then be determined by calculating the frequencies between various pitch transitions. A new piece is then composed by performing a *random walk* using the probabilities from the transition matrix.

As an example, the melody *Mary had a Little Lamb* is represented by the note sequence $M = \{E, D, C, D, E, E, E, D, D, D, E, E, E, E, D, C, D, E, E, E, E, D, D, E, D, C\}$. This melody only has 3 distinct pitches so it is sufficient to describe the state space with just $S = (C, D, E)$. An MC built with this melody would then be represented by following transition matrix.

$$P = \begin{matrix} & \begin{matrix} C & D & E \end{matrix} \\ \begin{matrix} C \\ D \\ E \end{matrix} & \begin{pmatrix} 0 & 1 & 0 \\ \frac{3}{11} & \frac{4}{11} & \frac{4}{11} \\ 0 & \frac{5}{13} & \frac{8}{13} \end{pmatrix} \end{matrix} \quad (2)$$

A caveat to this approach is the fact that most, if not all, music does not possess the Markov Property. The structure of note transitions is highly dependant on the larger scale structure of music as notes play different roles depending on whether they are part of a chord progression, a cadence or the climactic note of phrase. One way to address this issue is to increase the *order* of the MC. The term order refers loosely to the amount of memory an MC has. In an m th order MC, the probability distribution for the next state depends on the previous m states. If $m = 1$ we have the MC defined above. Higher order MCs can be described in a similar fashion to first order MCs by increasing the state-space.

In the previous example, in order to use a second order MC, one would use the state space $S' = (CC, CD, CE, DC, DD, DE, EC, ED, EE)$. As such, given an original state space of size $|S|$, an n th order MC would have a refined state space of size $|S|^n$. With higher order MCs, sampled musical pieces will tend to exhibit larger scale structures such as concrete phrases or themes.

Another issue with modeling music composition with MCs is one that has to do with the state-representation. Using the states as defined above would only be sufficient to model *monophonic* music, music with just one voice. In order to model polyphonic music or harmonies, one needs a different representation. One possible approach would be to use sets of notes as the different states, i.e. one state s could be given by $s = \{F, A, C\}$ to represent that a F major triad was played. If the state space S in such an approach would be limited to a maximum of three pitches per state and only using pitch classes, then $|S| = 12^3 + 12^2 + 12 = 1884$ different states. Creating a transition matrix from a piece of music using

such a large state space would result in a very sparse matrix. Sampled music from such a model would thus be very similar if not identical to the original piece. Therefore, in determining a suitable representation, it is necessary to extract features that sufficiently describe the music from a melodic and harmonic point of view. This problem remains an open question, although some attempts have been made to address the issue. In an approach outlined by Conklin, music is represented using *vertical viewpoints*.

$$IV(a, b) = |b - a| \quad (3)$$

In the formula above, a and b represent the (MIDI) pitch values of the notes and thus the interval just represents the amount of semitones between two notes. Using such an approach one could model, for example, which types of intervals typically follow a perfect fifth. Obviously, relying solely on this representation would result in a loss of information about the absolute pitches used in the piece.

Markov Chains are used in this thesis in order to influence the playout policy of the Monte Carlo search. Since composing music is such a complex problem, guiding the search towards transitions that have been sampled from existing works is expected to increase the efficiency of the search algorithm.