An algorithmic approach to South Indian classical music

Article *in* Journal of Mathematics and Music · May 2019

DOI: 10.1080/17459737.2019.1604845

Harish Seshadri
Indian Institute of Science Bangalore

Music and symbolic dynamics: The science behind an art

Shayan G. Srinivasa
Department of Electronic Systems Engineering
Indian Institute of Science, Bangalore, 560012
Email: shayan.gs@dese.iisc.ernet.in

Harish Seshadri
Department of Mathematics
Indian Institute of Science, Bangalore, 560012
Email: harish@math.iisc.ernet.in

***Abstract*—Music signals comprise of atomic notes drawn from a musical scale. The creation of musical sequences often involves splicing the notes in a constrained way resulting in aesthetically appealing patterns. We develop an approach for music signal representation based on symbolic dynamics by translating the lexicographic rules over a musical scale to constraints on a Markov chain. This source representation is useful for machine based music synthesis, in a way, similar to a musician producing original music. In order to mathematically quantify user listening experience, we study the correlation between the max-entropic rate of a musical scale and the subjective aesthetic component. We present our analysis with examples from the south Indian classical music system.**

***Index Terms*—music signal processing, symbolic dynamics, Markov chains**

I. Introduction

Music is a form of art using sound and silence for communication. For most people, music is an integral part of their culture. The creation, performance and appreciation of music is highly social and culture dependent. This makes the definition of music rather subjective. However, it is widely accepted that good music encompasses core elements such as melody, harmony, rhythm, dynamics, timbre and texture [1],[2],[3]. Melody is a successive arrangement of musical tones that can be perceived by the listener as a single entity in the foreground. Harmony refers to the simultaneous use of pitches and chords (i.e., two or more notes played simultaneously). While harmonic usage is more common in Western music, South-Asian musical forms, particularly, Indian classical genres have little emphasis on this aspect. Rhythm comprises of cyclically repeating beats/patterns with a certain tempo and articulation, and musical sounds and silences are time synchronized with rhythm. The term dynamics refers to the intensity and velocity of sound production. This is a volume modulation technique. Timbre refers to the quality of a musical tone from a psychoacoustic view. For example, the musical tone produced by violin is different from that of flute. Texture refers to the overall quality of music perceived through melody, rhythm and harmonic elements; music comprising of vocal and several instruments can be regarded as having a thick texture. Thus, one can qualitatively distinguish music from noise since music is structured in many ways.

Music theory is linked to many diverse fields such as acoustics, human physiology, psychology and mathematics [4],[5],[6]. There is a strong link between music and mathematics [4],[6]. Musical notes are the key elements constituting a melody. They form the atomic basis for sound perception. These notes can be arranged in different ways within a musical scale. In Western music, an octave comprises of a series of twelve notes. This series of twelve notes also called semitones constitutes a chromatic scale. Contrary to this, Arabic and Persian music comprise of quarter-tones. Indian music uses 7 notes with 5 variant notes that correspond to the 12 tones of the European chromatic scale and are used differently [7]. Theoretically, the 12 tones of Indian music map onto a 22 division of the octave, where, the seven tones are non-linearly placed in the ratio 4:3:2:4:4:3:2 based on consonance [8]. Based on the cardinality of the tones in the ascending and descending octave, there are various melodies such as the seven toned major, six toned major, pentatonic minor, etc. One can have asymmetries in number of tones and the choice of tones during the ascending and descending portions of a chromatic scale leading to various melodies.

We consider the well known heptatonic scale corresponding to the European chromatic scale for theoretical investigation throughout this paper. This assumption does not in any way make the theory restricted towards other genres of music. Following the convention in south-Indian classical music, the seven tones corresponding to the heptatonic scale are labeled as 's'-shadja, 'r'-rishabha, 'g'-gandhara, 'm'-madhyama, 'p'-panchama, 'd'-daivata, and 'n'-nishada [7]. In the Western solfa equivalent, they correspond to 'do', 're', 'mi', 'fa', 'so', 'la', and 'ti'. Once the base frequency for 's' is chosen, the tones 's' and 'p' form an invariant perfect-fifth and the remaining patterns are in a progression relative to the base frequency for tone 's'. Except in the case of 's' and 'p', there are variations around other tones as indicated in Table I. For example, 'm' has 2 variants m_1 and m_2. An overtone of the seven notes and their variants can be made by a shake or quiver of the pure tone. Indian music has at least 15 well defined overtones called gamakas [7]. Overtones have a semantic context in music creation and often add melody to music. The seven tones with their variations, relative frequencies and rankings over an octave are shown in Table I. It is possible to extend the frequencies and rankings to tones above and below the base octave by a relative linear translation of numbers from Table I.

TABLE I

TONE NAMES, NOTATION, RELATIVE FREQUENCY AND RANKINGS OVER AN OCTAVE AS PER SOUTH-INDIAN MUSICOLOGY.

Indian tone name	**Western equivalent**	**Notation**	**Normalized relative frequency**	**Indian musicology ranking**
shadja	C	$\mathbf{s}$	1	1
shuddha rishabha	D flat	$\mathbf{r_1}$	$16/15$	2
chatushruti rishabha	D	$\mathbf{r_2}$	$9/8$	3
shuddha gandhara	E double flat	$\mathbf{g_1}$	$32/27$	3
shatshruti rishabha	D sharp	$\mathbf{r_3}$	$6/5$	4
sadharana gandhara	E flat	$\mathbf{g_2}$	$6/5$	4
antara gandhara	E	$\mathbf{g_3}$	$5/4$	5
shuddha madhyama	F	$\mathbf{m_1}$	$4/3$	6
prati madhyama	F sharp	$\mathbf{m_2}$	$45/32$	7
panchama	G	$\mathbf{p}$	$3/2$	8
shuddha daivata	A flat	$\mathbf{d_1}$	$8/5$	9
chatushruti daivata	A	$\mathbf{d_2}$	$27/16$	10
shuddha nishada	B double flat	$\mathbf{n_1}$	$16/9$	10
shatshruti daivata	A sharp	$\mathbf{d_3}$	$9/5$	11
kaishiki nishada	B flat	$\mathbf{n_2}$	$9/5$	11
kakali nishada	B	$\mathbf{n_3}$	$15/8$	12

With the basic setup from Table I, one can easily generate arbitrary musical scales through combinatorial means. The purpose of this article is to expose the science behind artificial music synthesis similar to the way a musician can create original musical patterns. By translating the lexicographic constraints on musical scales to appropriate constrained graphs, we can obtain the necessary source models for music creation. Music signal processing is an emerging area of research building upon tools from music theory, acoustics and information theory. Music signal processing has many applications such as artificial music synthesis, music signal analysis and source separation. With appropriate source models, we can synthesize music digitally by means of an 'automaton'. Music source models can be used in a music analyzer such as a content retrieval system in a music search engine etc. Also, music models are very useful for separating vocal and various instruments within a thick musical texture for digital enhancement of individual source streams and remixing.

The paper is organized as follows. In section II, we link music theory to symbolic dynamics. By representing musical scales as constrained graphs, we show how musical patterns can be artifically synthesized through symbolic dynamical means. The combinatorial entropy of musical patterns is directly linked with the maxentropic rate of the constrained graph representing the underlying musical scale. By computing this quantity, we can quantify user listening experience as well as a musician's improvisation skills from an information-theoretic perspective. Simply put, not all musical scales are worth listening for a long time since patterns can get repetitive. Scales having higher entropy produce more combinatorial patterns and can be listened to for longer durations. In section III, we introduce the concept of perceptual equivalence of musical scales. By cyclically shifting notes over a musical scale, and fixing the scale reference at the shifted note, one can generate new scales that are perceptually equivalent to one another. This concept is cleverly done by highly skilled musicians during concerts. When a musician introduces perceptual equivalence during a concert rendering, listeners can perceive different melodies at the same time depending upon where they perceive the reference scale. We would like to quantify this special effect mathematically. In section IV, we show examples for modeling a musical automaton followed by conclusions and discussions.

II. MUSIC REPRESENTATION AND SYMBOLIC DYNAMICS

Symbolic dynamics originated in the study of discrete dynamical systems using methods of recurrence, transitivity and graph-theoretic concepts [9]. The applications include coding for constrained systems in magnetic and optical data storage [10], study of lattices in theoretical chemistry [11], automata theory, and many others. Music signal processing is another emerging area based on symbolic dynamics.

Meyer established one the earliest connections of information theory to music [1]. Matthews [12] was one of early contributors in the field of algorithmic computer music creation. Risset [13][14] developed perceptual music synthesis techniques. Beyls [15] considered the musical universe as cellular automata.

Musical patterns can be created from atomic notes that are braided in an aesthetically constrained manner according to a musical scale. In the language of symbolic dynamics, we can visualize musical patterns as multi-level constrained sequences over an alphabet set comprising of musical tones.

Before we begin our discussions, we introduce a few formal definitions. In music, pitch refers to a fundamental auditory attribute of sound by which musical tones can be ordered from low to high over a scale. Pitch is almost synonymous to frequency without any reference to the intensity of the perceived sound. An octave refers to the interval between a musical pitch and its second harmonic (either double of half the pitch frequency). A tone corresponds to a frequency subject to the instrument/voice. The reader must note that tones produced by various instruments bear a certain characteristic even though they are tuned to the same pitch. Let $\mathcal{S}$ denote the set of all tones over an octave. For south-Indian music, using Table I [7], $\mathcal{S} = \{s, r_1, ..., n_3\}$.

Definition 1: Let $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$ denote the forbidden set that disallows certain tones from $\mathcal{S}$ to be paired together. The rules of south-Indian music do not allow the following sets to occur.

- $\mathcal{F}_1 = \{(r_2, g_1), (r_3, g_2), (d_2, n_1), (d_3, n_2)\}$
- $\mathcal{F}_2 = \{(r_3, g_1), (d_3, n_1)\}$

The set $\mathcal{F}_1$ comprises of tones that are accorded the same ranking since they are perceptually close to the human ear. The set $\mathcal{F}_2$ comprises of notes that are highly disconsonant when paired together. From Table-I, even though the relative frequencies for the pairs (r_3,g_1) and (d_3,n_1) are different, they are accorded the same ranking. It must be noted that the ranking scheme in Table I clearly lacks a bijection between tones and their rankings.

Definition 2: A scale $\mathcal{R}_{\mathcal{F}}$ is a constrained sequence of notes over ascending and descending octaves such that any two element subset of musical sequences from $\mathcal{R}_{\mathcal{F}} \notin \mathcal{F}$.

With a few notable exceptions where a musical scale terminates within the base octave, almost all scales span over all octaves. A tone $a \in \mathcal{R}_{\mathcal{F}}$ lying within an octave δ levels higher or lower with respect to the base octave is denoted as $a^{(\pm\delta)}$. In general, the following can be summarized about musical scales:

- Scales can be asymmetric: The tones during ascending and descending portions of an octave need not be the same.
- Scales need not be ordered: The tones within a scale need not be partially ordered according to the rankings in Table-I. In other words, scales can have higher ranked tones preceding lower ranked tones.

Though one can mathematically define a scale to span an arbitrary number of octaves above and below the base octave, professional musicians restrict themselves to one octave above and below the base octave to produce music of good acoustic quality. Human vocal chords are also restrictive to this effect.

Having defined a scale with the above properties, we would like to illustrate the rules pertaining to scales through examples.

Example 1: The following examples of scales provide an insight into the variations of lexical constraints involved with the scale constructions. The scales will be demonstrated via musical clips embedded within the accompanying presentation.

- The pentatonic scale called mohana is defined as $\mathcal{R}^{(a)} = \{s, r_2, g_3, p, d_2, s^{(1)}\}$ during ascent ($s^{(1)}$ indicates that the scale continues to span multiple octaves). The descending portion of the scale is $\mathcal{R}^{(d)} = \{s^{(1)}, d_2, p, g_3, r_2, s\}$, exhibiting symmetry with the ascending scale. The superscripts a and d in $\mathcal{R}^{(a)}$ and $\mathcal{R}^{(d)}$ denote ascending and descending portions respectively. Over a 2 octave span during ascent, this scale comprises of notes taking values from the set $\mathcal{R} = \{p^{(-1)}, d_2^{(-1)}, s, r_2, g_3, p, d_2, s^{(1)}, r_2^{(1)}, g_3^{(1)}, p^{(1)}\}$. In this example, it must be noted that the set of tones constituting the scale are arranged in a strictly increasing order of the tone rankings.
- The scale with ascending notes given by $\mathcal{R}^{(a)} = \{s, r_1, g_3, m_2, p, d_2, p, s^{(1)}\}$ and descending notes given by $\mathcal{R}^{(d)} = \{s^{(-1)}, n_3, d_2, p, m_2, g_3, r_1, s\}$ is called purvikalyani and spans over multiple octaves. This asymmetric scale does not have ordering during the ascent since the tone d_2 appears before p during ascent, but is well ordered in the descending sequence of notes. For this scale, the forbidden set comprises of the following subsequences $\mathcal{F} = \{(p, d_2, a)\}$, a being any tone over a higher octave including $s^{(1)}$.
- An example of a south-Indian music scale that does not span more than an octave is chittaranjani. This scale is defined as $\mathcal{R} : \mathcal{R}^{(a)} = \{s, r_2, g_2, m_1, p, d_2, n_2\}, \mathcal{R}^{(d)} = \{n_2, d_2, p, m_1, g_2, r_2, s\}$.

Based on Table-I, we can obtain different scales with varying levels of sophistication. For instance, using elementary combinatorics, we have 72 different 7 toned major scales. The 72 major scales are very important in Indian musicology since many new aesthetically pleasing scales can be derived from these elementary scales. Each of the 72 major scales are labeled lexicographically so that two consecutive scales differ in at most two tones (9). The first of the 72 major scales starts with the ascending scale $\mathcal{R}^{(a)} = \{s, r_1, g_1, m_1, p, d_1, n_1, s^{(1)}\}$ and descending notes given by $\mathcal{R}^{(d)} = \{s^{(-1)}, n_1, d_1, p, m_1, g_1, r_1, s\}$. The seventy second major has the ascending scale $\mathcal{R}^{(a)} = \{s, r_3, g_3, m_2, p, d_3, n_3, s^{(1)}\}$ and descending notes given by $\mathcal{R}^{(d)} = \{s^{(-1)}, n_3, d_3, p, m_2, g_3, r_3, s\}$. The mapping from the name of the scale to its associated index from 1 to 72 is accomplished by an ancient technique called 'katapayadi' formula which is equivalent to hash functions in today's computer science. For more details on this formula, the interested reader is referred to the paper in [16].

So far we have discussed the constituents of a scale. Rhythm is an important constituent of music. Concert renderings often have melodies with rhythm in the background. In musical language, rhythm is the arrangement of sounds and silences over a repeating measured time interval called a cycle. In theory, the length of a cycle can be arbitrary. However, from Indian musicology perspective about 175 different rhythms are well documented, and compositions are set to those rhythms. Each rhythmic cycle is comprised of beats. A beat duration can be further decomposed into p-adic sub-beats. In other words, it is possible to have sub-beats that are of length 2, 4 and 8, or 3, 6 and 9, or 5, 10 and 15, and so on. Dyadic sub-beats are the most common compared to three or five-adic sub-beats. Higher p-adic representations are mathematically possible but not aesthetically pleasant sounding to the human ear. Expressed mathematically, if a cycle is of length n (i.e., comprising of n beats), each beat i can have $k_i p_i$ sub-beats where k_i is any positive integer and p_i is a radix that can be 2, 3 or 5 etc. In rhythm improvisations, the sub-beats over a beat duration usually follow a Markov random process. It is highly likely that two or three consecutive beat durations have the same number of sub-beats.

The variations in speed for the emission of musical patterns in the foreground is a Markov random process. For music to resonate, it is desirable to have sub-beats of the rhythm in the background be correlated to the duration of emission of musical patterns in the foreground. For reasons discussed before, we would like to model the emission of musical notes using dyadic variations in the speed.

A. Layered Graph Representation

In order to build a musical automaton, we need to represent the creation of musical tones with an underlying random process. This is accomplished using a layered graph representation for musical notes. Figure 1 shows the schematic of a layered constrained graph. The graph is constructed with two purposes in mind:

1) The emission of music tones are of variable duration taking a finite set values.
2) The notes themselves are constrained according to a musical scale spanning two octaves. The two octave restriction is to ensure good acoustic quality.

In order to represent the transitions of scale constrained musical tones with variable duration, we resort to layered constrained graphs. Let $\mathcal{G}_i$ denote the constrained graph for the i^{th} layer. Within a layer i, a musical tone has a time duration of $2^{-(i-1)}$ units. Let $V(\mathcal{G}_i)$ denote the vertices for graph $\mathcal{G}_i$. Each vertex/state corresponds to a tone belonging to the scale. Let $E(\mathcal{G}_i)$ denote the directed edges for graph $\mathcal{G}_i$ indicating valid paths satisfying the scale constraints. Let $\mathcal{G}_i \circ \mathcal{G}_j$ denote the merger of the graphs $\mathcal{G}_i$ and $\mathcal{G}_j$ such that:

- $|V(\mathcal{G}_i \circ \mathcal{G}_j)| = |V(\mathcal{G}_i)| + |V(\mathcal{G}_j)|$.
- For any two vertices $v_1 \in V(\mathcal{G}_i)$ and $v_2 \in V(\mathcal{G}_j)$, a directed edge connects v_1 and v_2 if the scale constraints are satisfied.

The overall constrained graph over n layers is represented by $\mathcal{G} = \mathcal{G}_1 \circ \mathcal{G}_2 \circ .... \circ \mathcal{G}_n$. From a practical angle, typically, $n \leq 3$. Since there is no constraint on the scales in each layer, the graphs $\mathcal{G}_i$ and $\mathcal{G}_j$ are identical except for the emission durations i and j. The order of the Markov random process for building the automaton depends on the constraints of the scale. For example, in the case of the pentatonic scale 'mohana' described earlier, there is no constraint on the notes appearing within the scale. However, from aesthetic view point, abrupt transition of notes from lower to higher octaves and vice versa are not appealing. In such cases, it is best to represent the transitions through a 1^{st} order Markov process and learn the transition probabilities. With a slight abuse of notation on the superscript, each tone $a^{(i)}$ appearing in layer i is represented by a state $s(a^{(i)})$ within that layer. Based on practical data from well known musical compositions set to those scales, the state transitions from $s(a^{(i)})$ to $s(b^{(j)})$ are modeled. The emission of notes at time t is governed by the probability transition $p_{a^{(i)}b^{(j)}} = Pr(s_t = a^{(i)} | s_{t-1} = b^{(j)})$ of the layered digraph.

The layered graph formalism that we described for a 1^{st} order process can be easily generalized to scales represented by higher order Markov processes. Suppose a scale is characterized by a Markov process of order D. Let a block of music tones in layers i and j be denoted by $\mathbf{a}^{(i)} = a_1^{(i)} a_2^{(i)} ... a_D^{(i)}$ and $\mathbf{b}^{(j)} = b_1^{(j)} b_2^{(j)} ... b_D^{(j)}$. Since the scale has memory D, a merger of blocks $\mathbf{a}^{(i)}$ and $\mathbf{b}^{(j)}$ is valid if the concatenated block $\mathbf{a}^{(i)} \odot \mathbf{b}^{(j)}$ satisfies the scale constraint. Accordingly, the states within a graph can be represented by blocks of musical tones consistent with the memory of the Markov process. The states are connected by directed edges if the merger of these blocks denoted by the $\odot$ operator forms a valid sequence across all the layers. A random walk over the layered graph produces all musical tones that satisfy this constraint. We need to model the transition probabilities from real data to accurately represent those transitions that are musically pleasant to hear.

Example 2: The scale 'purvikalyani' described earlier forbids the occurrence of a subsequence of the type $\{pd_2a\}$ only during the ascending scale. In other words, musical tones 'a' belonging to this scale, that are lexicographically ranked higher than 'd_2', are prohibited whenever the tone 'd_2' is immediately preceded by 'p'. Subsequences of the form $\{pd_2s^{(1)}\}$ etc are not allowed, where as, subsequences like $\{md_2s^{(1)}\}$ or $\{pd_2m_2\}$ are allowed. Clearly, the emission of a note 'a' depends on the previous two notes that are emitted in the sequence. The Markov process has memory two. In such cases, a state in each layer is formed by a block comprising of two musical notes. Directed edges are drawn connecting these states if the merger satisfies the scale constraint.

B. Entropy of music

There is a high degree of correlation between the subjective component of music listening experience with the combinatorics of aesthetically well sounding musical sequences. Even if the performing musician is highly professional, some musical scales can be heard for a longer duration of time than the rest. Scales that do not have a high entropy rate often manifest in musical patterns that may seem repetitive over time. Modeling the music source and computing the combinatorial entropy of the musical sequences is a natural way towards quantifying the quality of music improvisation.

The layered graph model that we discussed in section II-A encapsulates the essence of the musical source. Let the layered graph $\mathcal{G}$ have n layers with m states in each layer. Let each state be represented by D musical tones consistent with the memory of the underlying constrained scale. Let $\pi_i^{(l)}$ denote the steady state probability of the Markov chain. Let $H_i^{(l)}$ be the source entropy at each state i within a layer l. Under dyadic duration of musical tone emissions, for a stationary Markov process, the overall source entropy H of a musical scale is given by the following result.

Fact 1: The entropy rate of a musical scale is given by
$$H = \frac{1}{D} \sum_{l=1}^{n} \sum_{i=1}^{m} \pi_i^{(l)} H_i^{(l)} 2^{(l-1)} bits/time.$$

Proof: Following the standard results on entropy rates of stationary Markov chains [9], using the fact that each layer has variable emission duration $2^{-(l-1)}$, the uncertainty within each layer l is given by,

$$H^{(l)} = \sum_{i=1}^{m} \pi_i^{(l)} H_i^{(l)} 2^{(l-1)}. \quad (1)$$

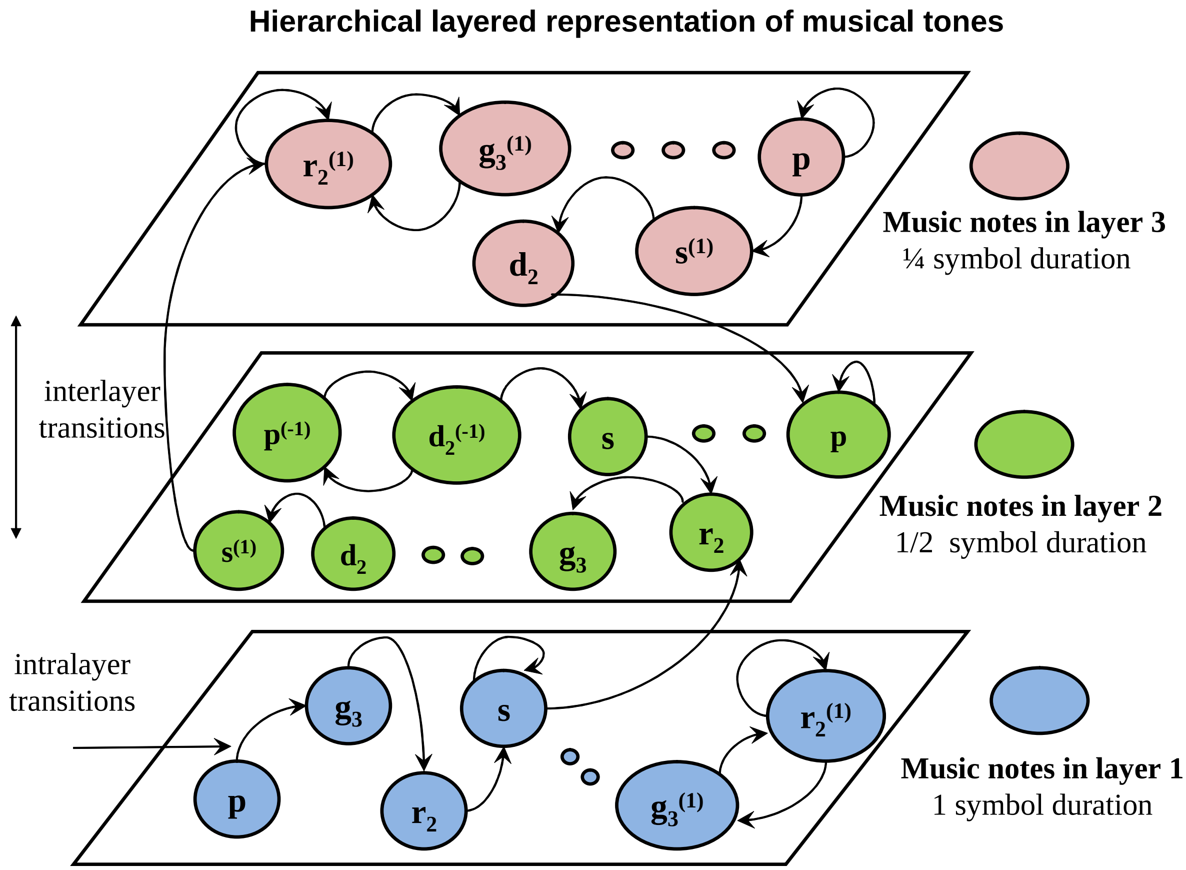

Fig. 1. Schematic of a layered graph representation for dyadic duration of musical notes. There are totally 3 layers labeled 1, 2 and 3 where musical tones are emitted with a duration 1, $\frac{1}{2}$ and $\frac{1}{4}$ time units respectively. Within each layer, the graph is constrained according to a musical scale with tones spanning two octaves. The process is 1^{st} order Markov.

Each state can have D tones corresponding to the memory of the Markov process. The overall entropy rate across all layers is

$$H = \frac{1}{D}\sum_{l=1}^{n} H^{(l)}. \tag{2}$$

■

It is rather intuitive that seven toned major scales have higher entropy compared to a pentatonic scale derived from a seven toned major scale. Similarly, a scale without constraints has higher source entropy than a constrained scale with same number of notes in the scale definition.

III. Perceptual Scale Equivalence

One of the intriguing aspects of music rendering during improvisation is the concept of 'perceptual scale shifting'. During normal rendering the reference tone is set to 's' corresponding to the base pitch. During perceptual scale shifting, a performer can shift the reference tone from 's' to any tone 'a' within the scale and use the tone 'a' as the new reference point for the scale. This operation is essentially equivalent to shifting the original scale to start with a base pitch set equal in frequency to tone 'a'. Under certain conditions, this change of reference appears as an illusion to the listener who can perceive it as a completely different scale. In this section, we describe the necessary conditions for quantifying this effect and outline a simple procedure to enumerate all the scales that form a perceptual equivalent class.

Let $\mathcal{S} := (\mathcal{S}^{(a)}, \mathcal{S}^{(d)})$ and $\mathcal{T} := (\mathcal{T}^{(a)}, \mathcal{T}^{(d)})$ be any two scales with constraints defined over the ascending and descending portions of an octave. Let $n_1^{(a)}$ and $n_1^{(d)}$ be the cardinality of notes over the ascending and descending portions for scale $\mathcal{S}$. Similarly, let $n_2^{(a)}$ and $n_2^{(d)}$ be the cardinality of notes over the ascending and descending portions for scale $\mathcal{T}$. Expressing the scale in terms of musical notes,

$$\begin{aligned} \mathcal{S}^{(a)} &= (s^{(a)}(1)s^{(a)}(2)...s^{(a)}(n_1^{(a)})) \\ \mathcal{S}^{(d)} &= (s^{(d)}(1)s^{(d)}(2)...s^{(d)}(n_1^{(d)})) \end{aligned} \tag{3}$$

$$\begin{aligned} \mathcal{T}^{(a)} &= (t^{(a)}(1)t^{(a)}(2)...t^{(a)}(n_2^{(a)})) \\ \mathcal{T}^{(d)} &= (t^{(d)}(1)t^{(d)}(2)...t^{(d)}(n_2^{(d)})) \end{aligned} \tag{4}$$

Let $\theta(a)$ denote the ranking of a note as per Table I.

Definition 3: Two scales $\mathcal{S}$ and $\mathcal{T}$ are perceptually equivalent to one another if the 1^{st} order finite differences between

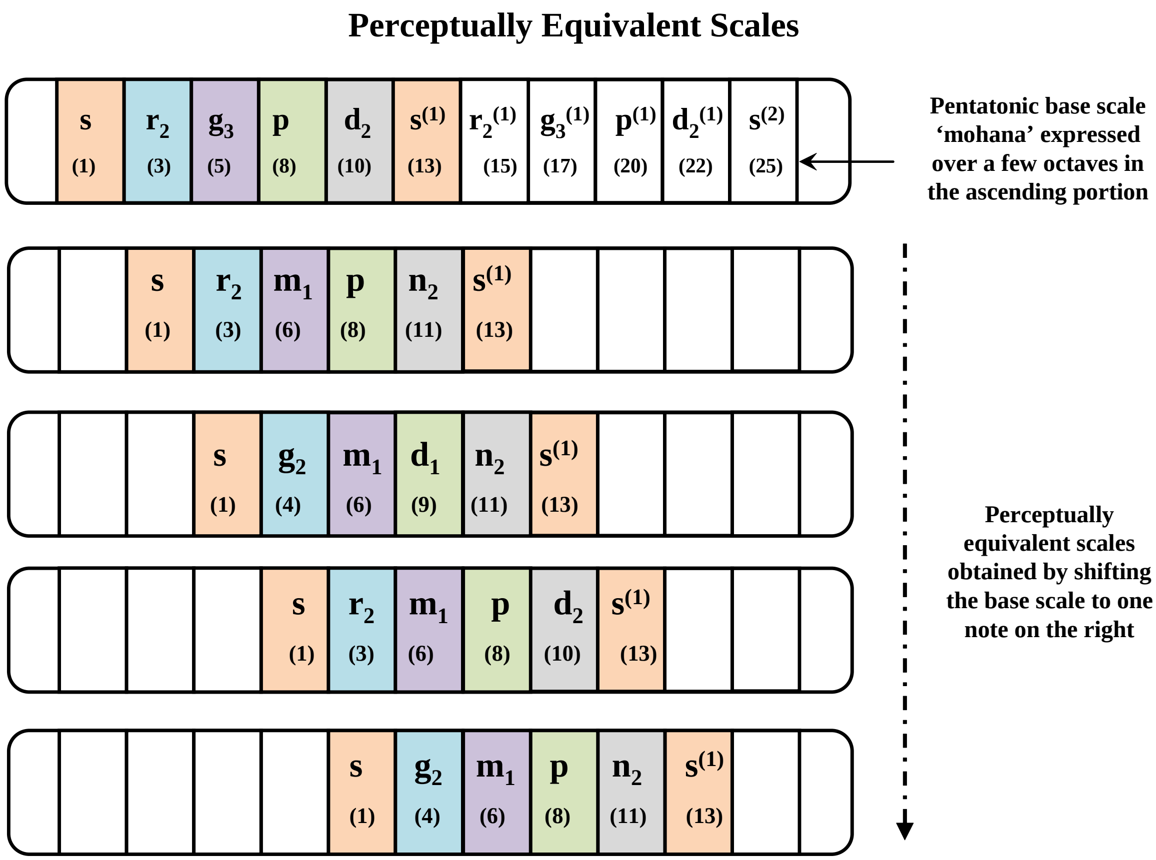

Fig. 2. Illustration of perceptually equivalent scale. The notes and the rankings are indicated over the scales. The pentatonic base scale 'mohana' can give rise to four other perceptually equivalent pentatonic scales by shifting the base note accordingly.

the rankings of tones are equal in both ascending and descending segments for the two scales.

Figure 2 shows how a base scale can lead to perceptually equivalent scales. We will illustrate the concept of perceptual equivalence through examples.

Example 3: Consider the symmetric pentatonic base scale 'mohana' whose ascending notes are given by $\mathcal{R}^{(a)} = \{s, r_2, g_3, p, d_2, s^{(1)}\}$. The rankings of the notes are given by $\theta(\mathcal{R}^{(a)}) = \{1, 3, 5, 8, 10, 13\}$. Let $\phi(\mathcal{R}, k)$ denote the function that shifts the scale $\mathcal{R}$ by k units in the increasing/decreasing direction depending upon the sign of k. Shifting the base scale by two tones, we get $\phi(\mathcal{R}^{(a)}, 2) = \{g_3, p, d_2, s^{(1)}, r_2^{(1)}, g_3^{(1)}\}$. The rankings of the notes for the shifted scale are given by $\theta(\phi(\mathcal{R}^{(a)})) = \{5, 8, 10, 13, 15, 17\}$. The first order finite difference of the shifted scale is given by $\nabla(\theta(\phi(\mathcal{R}^{(a)}, 2))) = \{3, 2, 3, 2, 2\}$.

We consider a completely different symmetric pentatonic scale called 'hindola' whose ascending scale is given by $\mathcal{S}^{(a)} = \{s, g_2, m_1, d_1, n_2, s^{(1)}\}$. Computing the first order finite differences of the rank for this scale, we get, $\nabla(\theta(\mathcal{S}^{(a)})) = \{3, 2, 3, 2, 2\}$. The above results hold for descending portion of the scale as well. Thus, scale 'mohana' is perceptually equivalent to 'hindola', i.e., $\mathcal{R} \sim \mathcal{S}$. From 2 one can observe other equivalent scales for 'mohana'.

Procedure to obtain a perceptually equivalent class

We outline a general procedure for finding the set of scales perceptually equivalent to a base scale.

- Inputs: Scale $\mathcal{R} = (\mathcal{R}^{(a)}, \mathcal{R}^{(d)})$, $n^{(a)} = |\mathcal{R}^{(a)}|$, $n^{(d)} = |\mathcal{R}^{(d)}|$

1) Obtain the first order finite differences $\nabla(\theta(\phi(\mathcal{R}^{(a)}, i)))$ of the rank for each shift $2 \leq i \leq n^{(a)}$.
2) Set the starting index for the new scale to always 1 corresponding to the base tone s.
3) Obtain the new tone indices $\mathbf{v}$ by adding the finite differences from step 1 as
$\mathbf{v} = 1 + \nabla(\theta(\phi(\mathcal{R}^{(a)}, i)))$.
These indices form the rankings for the new tones.
4) Using Table I, find the new tone labels $\mathcal{S}^{(a)}$ corresponding to the ranking indices in $\mathbf{v}$ for the shift i. In case of a ranking tie, it is fine to choose any tone that corresponds to the ranking since the ranking is not bijective but based on aesthetics.
5) Store the new scale $\mathcal{S}^{(a)}$ for every $\phi(\mathcal{R}^{(a)}, i)$.
6) Repeat the above steps 1-5 over all tones in the descending scale to obtain $\mathcal{S}^{(d)}$ corresponding to $\phi(\mathcal{R}^{(d)}, i)$, $2 \leq i \leq n^{(d)}$.
7) Form the perceptually equivalent scales $\mathcal{S}$ by concate-

nating the pairs of ascending and descending segments appropriately.

IV. Musical Automaton: Modeling and Results

Modeling music signals is a challenging task. As mentioned earlier, there are several features within music that one needs to consider to make synthesized music on par with human rendition. With an appropriate source model, one can interpret a scale constrained musical composition as an instance of the underlying random process. To build a musical automaton, we considered two different scales 'mohana' and 'kalyani'. The first scale is pentatonic as described earlier. The second scale is a symmetric seven toned major scale comprising of notes $\mathcal{R}^{(a)} = \{s, r_2, g_3, m_2, p, d_2, n_3, s^{(1)}\}$ during ascent. Several lengthy musical compositions from the south-Indian genre pertaining to these scales were used for learning the transition probabilities. All the compositions had variations spanning two octaves as well as dyadic duration in the emission of the notes. The scale 'mohana' had totally 11 states spanning over two octaves from $p^{(-1)}$ to $p^{(1)}$, while the scale 'kalyani' had 15 states spanning over two octaves. Only pure tones without any overtones were considered during the modeling of the states. We restricted to 3 different layers for handling emission durations $1, \frac{1}{2}$ and $\frac{1}{4}$. We evaluated the steady state distribution of the states using the standard eigenvalue equation [9]. The maxentropic rates for the two scales are estimated to be 3.42 and 3.63 bits/unit-time respectively based on the model. This result is not surprising since the seven toned major is expected to produce more combinatorial music sequences than the pentatonic scale.

We also experimented the concept of perceptual scale shifting for the scale 'mohana'. The accompanying audio files demonstrate the performance of the automaton. Incorporating many instruments will lead to a thicker texture. Also, having overtones will significantly enhance the quality of music. These enhancements can be easily incorporated within our model with almost a ten fold increase in the number of states in each layer of the underlying constrained graph.

V. Conclusion and future directions

Algorithmic music synthesis is an interesting area where music theory meets symbolic dynamics. Building appropriate music source models has potential applications in music creation, music content quantification, music clip analysis and musical transcription (music-to-text conversion). Using the concept of layered Markov chains, we can translate the lexicographic constraints to aesthetically pleasing constrained transitions. Professional musicians have perfected the art of rendering music totally exalting a listener's experience. Our eventual goal is to model a musician's creativity and enact a virtual concert performance.

Acknowledgement

The authors would like to thank Praveen Nandakumar for his help in developing the software for this application.

References

[1] L. B. Meyer, "Emotion and meaning in music," Chicago Univ. Press., 1956.
[2] Dannenberg, "Music representation issues, techniques, and systems" *Computer. Music. Journal*, vol. 17, pp. 20-30, 1993.
[3] Benward and Saker, "Music: In theory and practice," McGraw-Hill Humanities/Social, vol. 1, pp. 47, 2003.
[4] D. Tymoczko, "The geometry of musical chords," *Science*, vol. 313, no. 5783, pp. 72-74, 2006.
[5] D. J. Levitin, "This is your brain on music: The science of a human obsession," Penguin Group Publishers, 2007.
[6] R. W. Hall, "Geometrical music theory," *Science*, vol. 320, no. 5874, pp. 328-329, 2008.
[7] Bhagyalekshmy, "Ragas in carnatic music," CBH Publishers India, ed. 7, 2006.
[8] R. H. M. Bosanquet, "On the Hindoo division of the octave, with additions to the theory of higher orders," *Proc. Royal. Soc. London*, vol. 26, pp. 372-384, 1877.
[9] D. Lind and B. Marcus, "An introduction to symbolic dynamics and coding," Cambridge Univ. Press., 1995.
[10] K. A. S. Immink, P. H. Siegel, and J. K. Wolf, "Codes for digital recorders," *IEEE. Trans. Inform. Theory.*, vol. 44, pp. 2260-2299, 1998.
[11] R. J. Baxter, "Exactly solved models in statistical mechanics," London: Academic. Press. Inc., 1982.
[12] M. V. Matthews, "The digital computer as a musical instrument," *Science*, vol. 142, pp. 553-557, 1963.
[13] J.-C. Risset, "An introductory catalog of computer-synthesized sounds," Murray-Hill, New Jersey: Bell Laboratories, 1969.
[14] J.-C. Risset and M. V. Mathews, "Analysis of music instrument tones," *Physics Today*, vol. 22, no. 2, pp. 23-30, 1969.
[15] P. Beyls, "The musical universe of cellular automata," *Proc. of the 1989 Intl. Comp. Music Conf.*, pp. 34-41, 1989.
[16] A. Venkataraman, "The ancient Katapayadi formula and the modern hashing technique," *IEEE. Annals. Computing.*, vol. 19, pp. 59-64, 1997.