Markov Chains
As an Aid To Computer Assisted Composition

Emilios Cambouropoulos

Abstract

In this article we introduce some fundamental features of Markov Chains and their use as a mathematical means of encoding temporal or sequential information in computer-assisted composition. Simple examples are presented at every stage, as well as a more extended application on 16th century motet melody construction.

Introduction

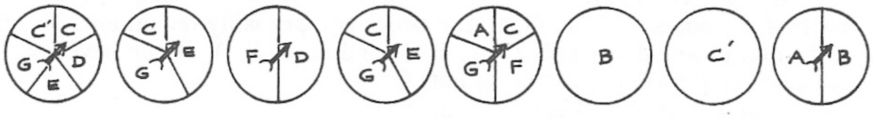
In aleatory composition with computers two main classes of random processes have been used by composers: random processes with independent observations and processes in which the current outcome is influenced by previous results.

A simple way to generate a sequence of randomly distributed notes is by constructing a spinner that is divided into a number of sectors each labelled with one note (e.g. 8 sectors with the 8 notes of the C major scale above middle C). By recording the outcome of each spin, a random, totaly uncorrelated melodic sequence may be produced:

A more sophisticated melodic surface can be generated if the next outcome depends on the current pitch value. For this purpose we should construct 8 spinners, each used only after a specific note has occured. For example:

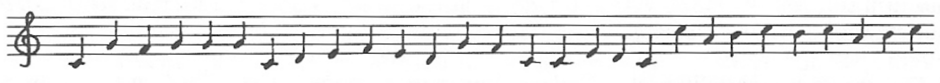
current note: C D E F G A B C'

next note:



In this case, if we suppose that the first note is a C, we could have melodic sequences as the ones shown below:

s1



s2

One can discern an obvious control over the melodic progression in contrast to the first random melodic sequence. The second mechanism described above is said to be a Markov chain.

Definition of Markov Chains

A probability process whose outcomes are functions of time (stochastic process) is said to be a finite Markov Chain[1] if a sequence of outcomes satisfy the following two properties:

a. Each outcome belongs to a discrete finite event space $\{e_1, e_2, e_3 \dots e_m\}$

b. The next outcome depends only on the current event and not on any other previous outcomes.

If the outcome of the n^{th} trial is e_i, then we say that the system is in event e_i at the n^{th} step (or at time n). For each pair of events (e_i, e_j) the probability[2] p_{ij}, that e_j occurs immediately after e_i occurs, is defined. The transition probabilities p_{ij} can be arranged in an mxm matrix :

$$P = \begin{vmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{vmatrix}$$

If the system is in event e_i then the i^{th} row $(p_{i1}, p_{i2} \dots p_{im})$ of the matrix P, presents the probabilities of all the possible events of the next trial (probability vector). Each probability takes a value within the interval $0<p_{ij}<1$ and the sum of the probabilities of each row is equal to one, i.e. $\sum_{j=0}^{m} p_{ij}=1$.

The transition probabilities of a Markov chain depend on the events i,j and on time t. In most applications (everywhere in this article) transition probabilities do not depend on time (p_{ij}) and in this case we speak of a Markov chain with stationary transition probabilities or of a homogeneous Markov chain.

In our 8-spinner example, the event space is the C major scale i.e. {C, D, E, F, G, A, B, C'}. Each spinner represents all the transition probabilities for the next event (probability vector) after one specific note of the event space. For example, after C each of the notes C, D, E, G, C' can occur equiprobably (equal arcs on spinner) with transition probabilities p=0.2 (i.e. 20%) and notes F, A, B can never occur, p=0. So we have the probability vector for note C:

p_C= (0.2, 0.2, 0.2, 0, 0.2, 0, 0, 0.2)

Similarly we get the probability vector for note D:

p_D=(0.2, 0, 0.4, 0, 0.4, 0, 0, 0) etc.

---

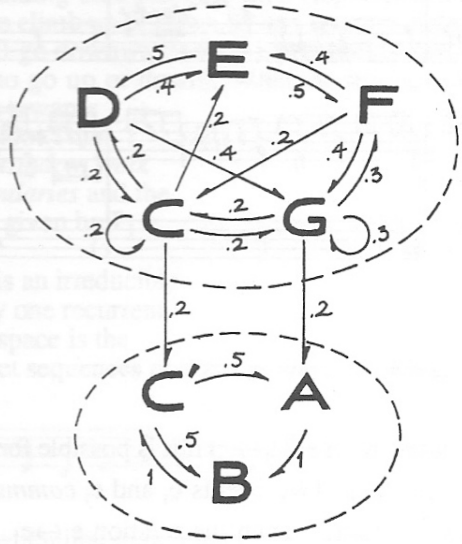
[1] For the purposes of this article we will consider Markov Chains with discrete time parameter i.e. time will take positive integer values (t=1,2,3...). Instead of time values we can speak about sequence of steps (n=1,2,3...) of the process.

[2] The transition probabilities of a Markov chain depend on the events i,j and on time t. In most applications (everywhere in this article) transition probabilities do not depend on time (p_{ij}) and in this case we speak of a Markov chain with stationary transition probabilities or of a homogeneous Markov chain.

The probability vectors for each note can be arranged in a transition matrix:

present note \ next note	C	D	E	F	G	A	B	C'
C	.20	.20	.20	0	.20	0	0	.20
D	.20	0	.40	0	.40	0	0	0
E	0	.50	0	.50	0	0	0	0
F	.20	0	.40	0	.40	0	0	0
G	.20	0	0	.30	.30	20	0	0
A	0	0	0	0	0	0	1	0
B	0	0	0	0	0	0	0	1
C'	0	0	0	0	0	.50	.50	0

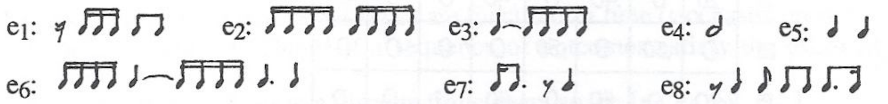
The transition probabilities of a Markov chain can be represented by a transition diagram (or kinematic diagram):



Event Space

Event spaces of different kinds have been defined by composers that use stochastic processes for musical composition. For the introductory example of this article our event space consisted of the notes of the C major scale. An event space may be made up of any fundamental compositional set of ordered elements such as pitches, durations, intervals, chords, motives, dynamic levels, timbre qualities, sound complexes, natural sounds, sound synthesis elements etc. On the other hand, the events of an event space may be defined not as single static entities but they may themselves be dynamic processes (e.g. stochastic processes like Markov chains with different event spaces each). The events may depend on sets of parameters each manipulated by various computer techniques. This way one can organise musical material in more than one levels.

In the computer program "Synthetis" we have implemented Markov chains for the rhythmic progression of a composition. The user inputs a set of rhythmic patterns for each part separately which are processed in a Markovian manner and defines other parameters for melodic and harmonic progression manipulated by other techniques. For example, in the beginning of the following three part composition, the rhythm event space for the soprano was defined as indicated below:



Properties of Markov Chains

An event e_j is said to be *reachable* from e_i ($e_i \rightarrow e_j$) if it is possible for the chain to reach e_j from e_i in a finite number of steps. Two events e_i and e_j *communicate* ($e_i \leftrightarrow e_j$) if they are both reachable from eachother. Then the relation $e_i \leftrightarrow e_j$ is an equivalence relation:

a) $e_i \leftrightarrow e_j$ for each state e

b) if $e_i \leftrightarrow e_j$ then $e_j \leftrightarrow e_i$ and

c) if $e_i \leftrightarrow e_j$ and $e_j \leftrightarrow e_k$ then $e_i \leftrightarrow e_k$.

The events of a Markov chain can be partitioned into *equivalence classes* of communicating events. One event may be reachable from an event of a different equivalence class but the two can not communicate.

In the Markov chain example presented previously we see that e.g. notes A and C' communicate (A↔C') because A→B→C' and C'→A whereas A is reachable from G but not the other way round (i.e. they don't communicate) because G→A but from A we can never reach G again. Events (e.g. note A) that once reached are certain to occur again are said to be ***recurrent*** and the equivalence class they belong to recurrent class. A recurrent class with only one event is said to be an *absorbing event.* Events that have a possibility of not occuring again (not recurrent) are said to be ***transient*** and the equivalence class they belong to transient class. Every Markov chain consists of recurrent (at least one) and transient classes (optional).

Once a process enters a recurrent class it can never leave it whereas once a process exits a transient class it can never re-enter. A Markov Chain that contains only one recurrent class (not periodic) and probably some transient classes it is called *ergodic*. If it contains one and only one recurrent class is called *irreducible.*

The Markov chain we have illustrated previously is an ergodic Markov chain that consists of one reccurent and one transient class. We can rewrite the recurrent class in a separate matrix (it is made up from the last three columns of the last three rows of the original matrix) which alone describes an irreducible Markov chain:

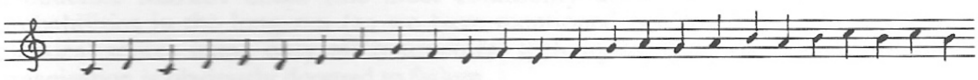
$$P = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ .5 & .5 & 0 \end{vmatrix}$$

A classic Markov Chain example (Random Walk)

Consider a man standing on a staircase on a step between the first step e_1 and the 8th step e_8 being able to climb up or down by one step at a time. If he has a probability p to go upwards, he will go downwards with a probability q=1-p (if p=0.5 then q=0.5 and it is equiprobable to go up or down). When he reaches the top or the bottom of the staircase, he is sent towards the opposite direction. This case is said to be a *random walk with reflecting boundaries* and the transition matrix is given by P_1.

$$P_1 = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & 0 & 0 & 0 \\ 0 & 0 & q & 0 & p & 0 & 0 & 0 \\ 0 & 0 & 0 & q & 0 & p & 0 & 0 \\ 0 & 0 & 0 & 0 & q & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{vmatrix}$$

This random walk is an irreducible Markov chain (only one recurrent class). If the event space is the C major scale we get sequences of notes as shown below:



Consider the case where the man remains permanently at the top or the bottom of the staircase when he reaches there. This process is a *random walk with absorbing bourdaries* and its transition matrix is given by P_2.

$$P_2 = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & 0 & 0 & 0 \\ 0 & 0 & q & 0 & p & 0 & 0 & 0 \\ 0 & 0 & 0 & q & 0 & p & 0 & 0 \\ 0 & 0 & 0 & 0 & q & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

This Markov chain is non-ergodic as it consists of two absorbing events e_1, e_8 (recurrent classes) and one transient class ($e_2 \ldots e_7$).

Random walks can be used for musical applications in which small gradual changes over the musical material are required (e.g. dynamic levels, sonic structure parameters etc.)

Higher Transition Probabilities and Stationary Distributions

The one-step transition from e_i to e_j is given by the probability p_{ij}. The transition from e_i to e_j in n-steps ($e_i \rightarrow e_{k1} \rightarrow e_{k2} \rightarrow e_{k3} \ldots \rightarrow e_{kn-1} \rightarrow e_j$) is denoted by the probability $p_{ij}^{(n)}$. If P is the transition matrix of a Markov chain then the n-step transition matrix is given by the n^{th} power of P, i.e. $P^{(n)}=P^n$.

Example: For the 3x3 irreducible matrix P for notes {A, B, C'} what is the probability to receive a B note after an A that occured 5 stages earlier, $P_{AB}^{(5)}$?

$$P = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ .5 & .5 & 0 \end{vmatrix} \qquad \text{Hence}^3,\ P^{(2)} = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ .5 & .5 & 0 \end{vmatrix} \cdot \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ .5 & .5 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 \\ .5 & .5 & 0 \\ 0 & .5 & .5 \end{vmatrix}$$

$$P^{(4)} = P^{(2)} \cdot P^{(2)} = \begin{vmatrix} 0 & .5 & .5 \\ .25 & .25 & .5 \\ .25 & .5 & .25 \end{vmatrix} \qquad P^{(5)} = P \cdot P^{(4)} = \begin{vmatrix} .25 & .25 & .50 \\ .25 & .50 & .25 \\ .125 & .375 & .50 \end{vmatrix}$$

So we get $P_{AB}^{(5)} = 0.25$ that is 25%.

We note that in matrix $P^{(5)}$ all entries are positive. Therefore a transition can be made between any two events in 5 steps, so this Markov chain is proved to be irreducible.

It can be shown that for an irreducible Markov chain, in the long run, the probability of occurence for any event e_j is stationary and is given by the component t_j on the unique fixed probability vector t of matrix P.

Example: For matrix $P = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ .5 & .5 & 0 \end{vmatrix}$ the unique fixed vector is found to be t = (.2, .4, .4) that means that in the long run note A will occur 20% and B and C' 40% of the time each.

---

3 Matrix multiplication for mxm matrices A and B:

$$\begin{vmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{im} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mm} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \cdot\cdot & b_{1j} & \cdot\cdot & b_{1m} \\ \cdots & & \cdots & & \cdots \\ \cdots & & \cdots & & \cdots \\ \cdots & & \cdots & & \cdots \\ b_{m1} & \cdot\cdot & b_{mj} & \cdot\cdot & b_{mm} \end{vmatrix} = \begin{vmatrix} c_{11} & \cdots & c_{1m} \\ \cdots & \cdots & \cdots \\ \cdots & c_{ij} & \cdots \\ \cdots & \cdots & \cdots \\ c_{m1} & \cdots & c_{mm} \end{vmatrix}$$

where $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{im}b_{mj} = \sum_{k=1}^{m} a_{ik}b_{kj}$

i.e. the ij-entry is obtained by multiplying the elements of the i^{th} row of matrice A with the corresponding elements of the j^{th} column of B and then adding.

"Memory" of Markov Chains and Higher Order Processes

According to the definition of Markov chains the current event depends explicitly only on the previous one. On the other hand, we note that there is a weaker indirect dependence on more than just the previous event. If event e_i is followed by e_j with probability p_{ij} and e_j by e_k with probability p_{jk} then the probability that e_i will be followed by e_j and then by e_k is p_{ijk} and is given by the relationship $p_{ijk}=p_{ij} \cdot p_{jk}$. For example on the 8-spinner matrix note E is followed by F with probability 0.5 (i.e. 50%) and F followed by C with probability 0.2 (i.e. 20%). The probability then of having the whole sequence $F \rightarrow G \rightarrow C$ is $p_{FGC}=0.5 \cdot 0.2=0.1$ (i.e.10%).

It is therefore obvious that the current event depends on earlier outcomes as well. This influence strongly decreases the more distant a previous event is.

It is possible to define higher order processes in which the current event depends explicitly on more than one of the previous events. For example, in a 2nd-order process, the current event depends on its two previous events. Its transition probabilities can be represented by a 3-dimensional matrix (on one axis the current event, on the other the last and on the third axis the second-to-last event). Higher order processes result in structures with greater coherence and regularity of succession. As in traditionally composed scores a current event usually depends on previous events in a stronger manner than a simple Markov chain designates, it is often practical or even necessary to implement higher order processes.

Justification and Criticism on the use of Markov Chains

The use of Markov chains in computer-assisted composition has been criticized by composers as producing musical texts that are too regular in the long run and that there is no phrase structure and hierarchical control on the musical material. Classical musical scores (e.g. by Bach, Beethoven etc.) have been statistically analysed and then reproduced using Markov chains or higher order processes. The music produced by such stochastic methods seems at first similar to the original score but in the long run there is no actual relationship to the initially analysed musical material.

These observations are rather obvious as, by definition, a Markov chain is a unilevel probability process (no hierarchical multilevel control) and, more specifically, an irreducible Markov chain (the most commonly used) has stationary transition probabilities in the long run (as we have mentioned before). Trying to reproduce classical scores using only microstructural techniques, ignoring higher structural organisation is obviously doomed to failure.

Composing with the computer requires a very accurate set of controlling mechanisms on many levels which contain a great amount of data (much of this data is considered for granted by a human composer and is used in classical score writing with out even been realized). Markov chains are a convenient way to reduce information that concerns the succession of musical material (in some level) into a formalized mathematical process that can easily be implemented on the computer. They can successfully be used to control the temporal or sequential flow of events from the microstructural level up to the overall formal level. More than one levels can be simultaneously manipulated by using Markov chains within Markov chains and so on. Alterations on the evolution of musical progression can be achieved by using transient classes till a recurrent class is finally reached.

If used in combination with other methods (generate and test methods, deterministic methods, grammars, transition networks, other stochastic techniques etc.) Markov chains can prove to be a very useful tool into computer composition.

An Application for 16th century Melody Construction

We will now present a practical application of Markov chains into the synthesis of melodic surfaces according to the style of Palestrina. We chose the style of early counterpoint (motets) as the microstructural features of composition of this era dominate over higher level organisation (in opposition to the multileveled structures of classical tonality).

The fundamental rules of melodic construction can be summarised as follows:

1. Rhythmic structure
   a. Metres in use: 4/2 and 3/2 (less common)
   b. Basic time-values[4] in use: 𝅗𝅥. , 𝅗𝅥 , 𝅗𝅥. , 𝅗𝅥 , ♩
      (In this application we have omitted 𝅜 , ♪). The dotted values are always a time-value tied with its half and not the opposite (i.e. 𝅝. = 𝅝⁀𝅗𝅥 never 𝅗𝅥⁀𝅝)
   c.Succession of time values:
      i. Quarter notes come in pairs in the place of half-note beats
      ii. Never two quarter notes alone in a stressed half-note

2. Melodic structure
   a. Notes are selected from one of the authentic modes (in this application without key-signature)
   b. Allowed intervals: 2nd , 3rd , 4th , 5th , minor 6th ↑, 8th
      No augmented or diminished intervals allowed
   c. Succession of intervals:
      i. The minor 6th and 8th are preceded and followed by a 2nd in the opposite direction
      ii. Never two or more consecutive melodic leaps in the same direction
      iii. During an ascending succession of notes a melodic leap (if there is one) will always be at the beginning. Exception: inverse cambiata
      iv. During a descending succession of notes a melodic leap (if there is one) will always be at the end
      v. A syncopation is always followed by a 2nd descending interval
      vi. Never an ascending melodic leap after a stressed quarter-note
      vii. Never a descending melodic leap after a weak quarter-note
      viii An ascending succession of notes can never end on a weak quarter note
   d. Accidentals
      The note B can be replaced by bB (musica ficta) as long as melodic intervals before and after conform with rule 2b
   e. Overall melody
      i. Starts and ends on the tonic degree (for one part compositions)
      ii. The second-to-last note (leading note) is preceded and followed by a 2nd interval and is sharpened for the Dorian, Mixolydian and Aeolian modes (in the Phrygian mode the second-to-last note is the supertonic)
      iii. Only one peak for each phrase.

---

4 Time-values have been seperated into two categories depending on whether they occur on a strong or on a weak half-note beat, This is necessary as we can make distinctions as the following:

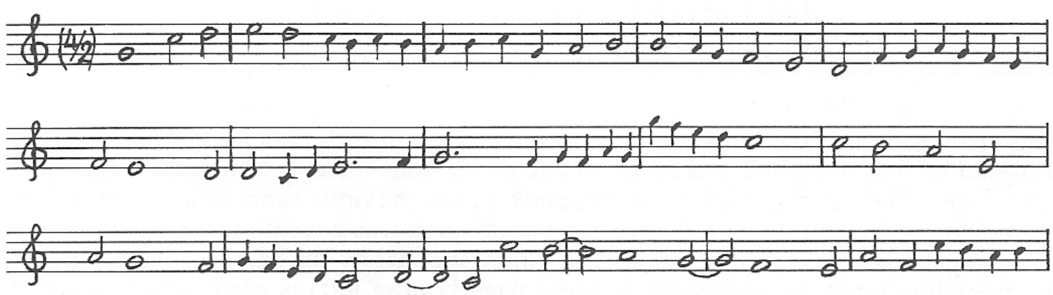
1. a whole-note on a weak beat is a syncopation
2. a dotted whole-note on a strong beat is a syncopation
3. a dotted whole-note is not allowed on a weak beat (𝅝. = 𝅗𝅥⁀𝅝). We also make the distinction between strong and weak quarter-notes.

Most of these rules along with statistical analyses concerning the succession of time-values and melodic intervals (from Palestrina motet melodies) can be represented by a Markov chain and some transition tables. We implemented a Markov chain for the succession of time-values and depending on the previous time-value we constructed transition tables for the selection of the pitch of the current note. Probability values $p=1$ indicate compulsory progression of the mechanism according to the rules and for $0<p<1$ we have probabilities taken from the statistical analysis of a small sample of Palestrina melodies.

Of course, there are quite a few matters that have to be taken care of by different algorithmic approaches. It is obvious that things such as the start or final cadence or the peak of the melody need different techniques. Or the balance of the melody in the middle range of the particular voice range needs special boundary treatment (e.g. by incrementing probabilities for opposite direction melodic leaps etc.).

But even details of the microstructure are not covered by the Markov chain (1st order process). For example, the cambiata or inverse cambiata figure needs at least a 3rd order transition process (it consists of 4 notes) or some special deterministic method that will interrupt the flow of the Markov chain. Or, in order to avoid having two quarter notes alone in a stressed half-note beat, we have forced the Markov chain to give another consecutive pair of quarter notes. But this might not be necessary as other quarter-notes may have preceded the strong beat pair. Again this problem is automatically solved by a 2^{nd} order process or some other algorithmic method.

We can see below a melodic surface generated by the mechanism described above:



The following matrix describes the time-value Markov chain:

		current time-value										
		strong beat						weak beat				
previous time-value	strong beat	0	0	0	0	0	0	,10	,05	,75	,10	0
		,05	,40	,10	,40	,05	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	1
		0	0	0	0	0	0	,23	,13	,50	,14	0
		0	0	0	0	0	1	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	1	0
	weak beat	0	0	0	0	0	0	,10	,05	,75	,10	0
		0	0	0	0	0	1	0	0	0	0	0
		,05	,40	,10	,40	,05	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	1
		,05	,05	,05	,40	,45	0	0	0	0	0	0

This chain generates the sequence of time-values for a melody. For each time-value the program selects one pitch according to one of the four interval transition tables (see next page). The interval table used, depends on the previous time-value (depicted in rhombus). For example, if the previous time-value was a weak whole-note (i.e. we have a syncopation), for the current time-value selected by the Markov chain process the transition table for syncopations (next page bottom left) is used, which actually forces the melody to descend by a 2nd interval. Or, if the previous time-value was a stressed quarter note (♩), for the current time-value the corresponding table (upper right) is used.

The four transition tables below are used for the selection of the current melodic interval (gives current pitch), depending on the time-value of the previous note. The bottom left table for syncopations can be replaced by a simple routine that gives a descending 2[nd] interval. (In the tables: ↑ for ascending intervals, ↓ for descending).

		↑				↓			
		2	3	45	68	2	3	45	8
↑	2	.50	0	0	0	.15	.15	.15	.05
	3	.60	0	0	0	.25	.05	.10	0
	45	.60	0	0	0	.25	.10	.05	0
	68	0	0	0	0	1	0	0	0
↓	2	.20	.05	.07	.05	.55	.03	.03	.02
	3	.60	.10	.30	0	0	0	0	0
	45	.40	.20	.40	0	0	0	0	0
	8	1	0	0	0	0	0	0	0

		↑				↓			
		2	3	45	68	2	3	45	8
↑	2	.60	0	0	0	.30	.03	.04	.03
	3	.50	0	0	0	.47	.01	.02	0
	45	.50	0	0	0	.47	.01	.02	0
	68	0	0	0	0	1	0	0	0
↓	2	.22	0	0	0	.75	.01	.01	.01
	3	1	0	0	0	0	0	0	0
	45	1	0	0	0	0	0	0	0
	8	1	0	0	0	0	0	0	0

		↑				↓			
		2	3	45	68	2	3	45	8
↑	2	0	0	0	0	1	0	0	0
	3	0	0	0	0	1	0	0	0
	45	0	0	0	0	1	0	0	0
	68	0	0	0	0	1	0	0	0
↓	2	0	0	0	0	1	0	0	0
	3	0	0	0	0	1	0	0	0
	45	0	0	0	0	1	0	0	0
	8	0	0	0	0	1	0	0	0

		↑				↓			
		2	3	45	68	2	3	45	8
↑	2	1	0	0	0	0	0	0	0
	3	.50	0	0	0	.50	0	0	0
	45	.50	0	0	0	.50	0	0	0
	68	0	0	0	0	1	0	0	0
↓	2	.12	.01	.01	.01	.85	0	0	0
	3	.97	.01	.02	0	0	0	0	0
	45	.97	.01	.02	0	0	0	0	0
	8	1	0	0	0	0	0	0	0

Conclusion

In this article we investigated the use of Markov chains as a mathematical means of encoding information relating to temporal or sequential progression of the musical material. Such mathematical processes can prove to be effective tools, used under the guidance of musical theory, to formulate computer models for classical or contemporary compositional purposes.

References

Mathematics and Markov chains

Allen, A.O., (1990), *Probability, Statistics and Queuing Theory with Computer Applications*, 2nd edition, Academic Press Inc., Boston

Grimmet,G.R. and Stirzaker,D.R.,(1982), *Probability and Random Processes*, Oxford University Press, New York

Kolhas, J., (1982), *Stochastic Methods of Operation Research*, Cambridge University Press, Cambridge

Lipschutz, S., (1974), *Probability*, McGraw Hill Book Company, New York

Ross, S.M.,(1963), *Stochastic Processes*, John Wiley, New York

Music and Markov chains

Dodge, C. & Jerse, T.A., (1985), *Computer Music*, Schirmer Books, N. York

Hiller, L., (1984),"The composer and the Computer", *Abakus*, I,4, 1984:9-31

Jones, K., (1981),"Compositional Applications of Stochastic Processes", *Computer Music Journal*, Vol. 5, No 2, Summer, 81

Xenakis, I., (1971), *Formalized Music*, Indiana University Press, Bloomington

Counterpoint

Jeppesen, K., (1950), *Counterpoint-The Polyphonic Vocal Style of the Sixteenth Century*, Williams and Norgate Ltd, London