5-20-2012

An Investigation of Melodic Musical Modeling Using Homogeneous and Non-Homogeneous Markov Chains

Eric Robert Sherman Buenger
Butler University

BUTLER UNIVERSITY HONORS PROGRAM

Honors Thesis Certification

Please type all information in this section:

Applicant: Eric Robert Sherman Buenger
(Name as it is to appear on diploma)

Thesis title: An Investigation of Melodic Musical Modeling Using Homogeneous and Non-Homogeneous Markov Chains

Intended date of commencement: 12 May 2012

Read, approved, and signed by:

Thesis adviser(s): Christopher Wilson — Date: 8 March 2012

Date

Reader(s): Rebecca C. Wahl — Date: 4/17/12

Wahl, R.

Date

Date

Certified by: Judith Harper Morrel, Director, Honors Program — Date: 12 July 2012

For Honors Program use:

Level of Honors conferred: University: Summa Cum Laude

Departmental: Actuarial Science with Highest Honors

University Honors Program

An Investigation of Melodic Musical Modeling Using Homogeneous and Non-Homogeneous Markov Chains

A Thesis

Presented to the Department of Mathematics and Actuarial Science

College of Liberal Arts and Sciences

and

The Honors Program

of

Butler University

In Partial Fulfillment

of the Requirements for Graduation Honors

Eric Robert Sherman Buenger

23 April 2012

The disciplines of mathematics and music have existed for centuries, as has the synthesis of these two subjects. In fact, Pythagoras, a name known to many mathematicians and dreaded by many algebra students, was one of the first people to investigate the relationship between math and music. In the 5th century BC, he was analyzing the ratios between musical notes and their string lengths [L]. He was the first to connect frequencies of sound to note pitch. Some people might be surprised that the most basic principle of sounds, and thus music, comes from the mathematics and physics of vibrations. Since Pythagoras, many different mathematicians and music theorists have made numerous connections between the two subjects.

Throughout history, the study of music by mathematical means has become a popular study topic. Music theorists have created numerous tuning schemes that differ from Pythagoras's tuning ratio. Others have turned to using mathematics to fuel their composition methods. One well known composition method is the use of twelve-tone rows, in which a sequence of which each of the twelve notes in Western music must be iterated at least once before any note can be repeated. Others yet have utilized mathematical generation of the notes they will compose. Joseph Haydn used Würfelspiel, German for a game of dice, to compose minuet trios. This method of composition "consisted of applying the outcome of throwing dice to choosing which of several possible musical motives would be selected from tables of precomposed musical figures" [L]. Today, with the sophistication of computers and other mathematical theories, composers can simply create a model that will produce music under certain parameters. The use of Markov chains as the model for musical composition has been a valuable tool for mathematically musical composers. Composition by means of Markov chains will be the focus of the rest of this discussion.

Background:

A *Markov chain* can be described as a collection of probabilities in the form of a matrix whose entries are the transitional probabilities from one *state* to another. A state can be defined as an event or occurrence that is being modeled. A Markov chain is an example of a *stochastic* model. The word stochastic is Greek in origin and means "random" or "chance" [T]. A stochastic model predicts a set of possible outcomes weighted by their likelihood or probabilities [T]. Table 1 is a general form of an *n*-state matrix. The transition probability $P_{i,j}$ appearing in the (row-i, column-j) entry of the matrix gives the relative likelihood of jumping from state i to state j, given that the current state is at state i [Am]. Markov chains can have a variety of *orders*. The order of a Markov chain tells "how much recent history is taken into account when determining the next state" [L]. For example, a third order Markov chain uses only the three most recent states in determining the conditional probabilities for the next transition. Throughout this paper, we consider only the first-order case. The lack of dependence on the sequence of events preceding the current state gives Markov chains a memoryless property.

Table 1

Source	Destination				
	State 1	State 2	State 3	…	State n
State 1	$P_{1,1}$	$P_{1,2}$	$P_{1,3}$	…	$P_{1,n}$
State 2	$P_{2,1}$	$P_{2,2}$	$P_{2,3}$	…	$P_{2,n}$
…	…	…	…	…	…
State n	$P_{n,1}$	$P_{n,2}$	$P_{n,3}$	…	$P_{n,n}$

In the field of actuarial science, actuaries make use a variety of probability models, including Markov chains. Two important types of Markov chains are *homogeneous* and *non-homogeneous* Markov chains. The main difference between the two versions is in respect to the aspect of time. When the model does not depend on the number n of states that have occurred, the chain is said to be homogeneous [Da]. The same model/matrix is applied for each single transition, whether it be the first transition, the last transition, or anywhere in between. On the other hand, a Markov chain whose transition probabilities are allowed to depend on the state at time n-1 *and additionally on the number of states that have occurred* is said to be *non-homogeneous*. In both cases (homogeneous and non-homogeneous), the Markov chain possesses *history independence*, that is, the probabilities for the n-th transition do not depend on which states occurred at times 1 through $n - 1$, though it could depend on the current (n-th) state. When referring to a non-homogeneous Markov chain matrix, the notation M_n will be used for the matrix of transition probabilities from time n to time $n + 1$, with n referring to the time the matrix is currently located. For homogeneous Markov chains, the notation M will be used. The lack of subscript reflects the fact that the same matrix will be used at each time interval.

These matrices can be used to model transitions in music from one tone to the next. For instance, if the melody of the song had the following note sequence: C→C→G→G→A→A→G (the first few notes of the melody from *Twinkle, Twinkle Little Star*). Its *first-order* transition matrix would resemble the matrix in Table 2. Recall that a first-order Markov process is defined so that the destination note is determined by probabilities that depend only on the one note preceding it. In the sequence, the note C occurs twice, with a C following one time and a G following the other time. The probability model would then define the transitional probability from a C to a C 50% of the time and a C to a G 50% of the time. Each row in a transition matrix must have the probabilities sum to 1. For a non-homogeneous Markov chain, the model could use one set of transition probabilities for the first grouping of measures (or phrase) and then use another for the next phrase, and so on.

Table 2

Source	Destination		
	C	G	A
C	1/2	1/2	0
G	0	1/2	1/2
A	0	1/2	1/2

Many mathematicians and music theorists have used Markov chains for random generation of music. Loy analyzed the melody from Stephen Foster's *Oh Suzanna* through the use of various orders of Markov processes [L]. Claude Alamkan used first-order Markov chain based on Beethoven's *Moonlight Sonata* while investigating similarities of compositional style [Al]. Michael Rubey investigated stochastic music generation, focusing on the generation aspect through computer programming [Ru].

While many students and researchers have worked on stochastic music generation, most tend to focus predominantly on the generation of the music, and all of them have used homogeneous Markov chains.

Thesis Statement:

As an actuarial science student, my observations have a different focus than the other composers. In the industry, actuaries aren't interested in a probability model for its own sake. Rather, they "want to use the model to analyze the ...impact of the events being modeled" [Da]. This analysis focuses equally on the generation of the model as well as the results of the model. While other researchers have investigated many topics in the field of musical generation through mathematical means, no one has yet explored non-homogeneous and homogeneous models simultaneously. This study compares melodic material generated from both homogeneous and non-homogeneous models in an attempt to determine which model leads to a more accurate representation of the given melody.

Assumptions:

There are a few assumptions that are necessary to state for this approach. There are limitless ways to describe an occurrence in music that goes beyond just the note being played. Some of those are dynamics, instrumentation, octave, and note duration. To narrow the focus of this study, the model only accounts for the pitch class of each note, that is, the pitch names (D, E, F#, et cetera), ignoring differences in octave. Also, the model analyzes only the melody of the selected songs, excluding all notes that comprise the harmony of the song. This significantly reduces the number of states permissible, as well as making the generated samples easier to use for comparison. When Loy did his generation only the pitch classes were synthesized; "the rhythms were copied from the original to aid comparison. This method carries a hint of the musical character of the original into the synthesized melody" [L]. This is the approach that is used in this study as well. This ensures that the Markov chain is limited to a finite number of states, specifically twelve states representing the twelve notes used in scales found in Western music. This model could be expanded so that the states are the possible combinations of both pitch and note duration. Rubey used this approach and had 42 unique states for his first-order probability matrix [Ru]. Since the focus of this project is more on the model and less on the methods of computer generation, the simplification of using only pitch classes makes generations of samples by hand much more obtainable. While the focus of this study is specifically first-order Markov chains, the study could be expanded to research higher order Markov chains (in which the chain keeps track of a bit more recent history) and their results. In regards to choosing what time the non-homogeneous Markov chains will begin and end, the use of musical phrases shall be used. For non-homogeneous chains, we will use one transition matrix to model each musical phrases from the original melody: the transition matrix M_n will be determined by the sequence of notes from the musical phrase that contained the *n*-th note in the original melody. We will additionally adopt the notation $M_{(j)}$ to describe the common one-step transition matrix for all the transition that occur within the *j*-th phrase. From the selected songs that shall be modeled, most of them can be divided in an obvious way into a few musical phrases that shall mark the limits of each non-homogeneous Markov chain.

The musical selections used for the formulation of the models come from a variety of sources. The pieces were chosen for their relatively well-known melodies. This will make informal aural comparisons of the generated samples back to the original piece easier due to the familiarity of the melody. Another primary reason for selecting these specific songs was for the ease of being able to determine where the musical phrases begin and end, which is important in forming the multi-state transition model. The songs also range a wide span of musical eras, from the Baroque period to the modern period. Full analysis was performed on three pieces, including Stephen Foster's *Camptown Races*, John Stafford Smith's melody that eventually became the melody of the *Star Spangled Banner*, and Aerosmith's *I Don't Want To Miss A Thing*. We also performed a more limited analysis on Johann Sebastian Bach's *Wachet auf, ruft uns die Stimme*.

It was necessary to make some modifications to the original melodies before constructing the Markov models. The Aerosmith song repeats the same note back to back many times throughout the song. This technique, common in modern popular music, is used in order to sing a lot of words in a short amount of time. In order to avoid giving extra weight to these repeated notes, an arrangement of this song was composed to eliminate that bias. Similarly, we focused on an abbreviated version of *Wachet auf* as the original melody is fairly repetitive. This truncated arrangement of that piece is roughly half as long as the original. Nothing aside from the length was altered. The other two songs required no alterations. All arrangements of the four songs that were used for the generation of the models can be found in Appendix A.

Methods:

The first step was to generate the transition probability matrices for each of the songs. Each song has two different sets of matrices, one homogeneous matrix and a series of non-homogeneous matrices. The construction of each matrix was a matter of counting the frequencies of one note passing to the next. Some songs will have a greater number of notes than others, which may lead to more variation in the generated samples. From the frequency table, it was only a matter of dividing the frequency of each possible transition away from the given note by the number of occurrences of transitions starting from that note to generate the probabilities needed to construct the transition matrices. For instance, there was one occurrence of a C→C transition in the *Twinkle, Twinkle Little Star* example shown above. Since there were two C's in the given sample, the probability of transitioning from C to C would be (1 occurrence of C→C) ÷ (2 transitions from C to any eligible note) = 0.5. The homogeneous transition matrix was generated by including all of the transitions and note occurrences in the song. The non-homogeneous transition matrices were generated by including only the note transitions and note occurrences that appeared during the corresponding selected musical phrase; that is, the entries in the matrix M_n are the transition probabilities computed using only the phrase of the original melody that contains the n^{th} note. All of these tables were entered into and stored in a Microsoft Excel spreadsheet to assist with generation and calculations. The model built in Excel could process up to a maximum of twelve states, corresponding with the twelve pitch classes most commonly used in Western music.

We now describe the mechanics of actually generating a melody using Microsoft Excel via Markov chain models. We first created cumulative probability tables, which were used in tandem with a sequence of randomly generated numbers in Excel to produce each of the song samples. This is best explained by example. If the transition probability matrix M_n looks like

Source	Destination		
	C	G	A
C	1/2	1/2	0
G	0	1/2	1/2
A	0	1/2	1/2

then the corresponding cumulative probability table (Table 3) would be as follows:

Table 3.

Source	Cumulative Probability		
	C	G	A
C	[0, .5)	[.5, 1)	0
G	0	[0, .5)	[.5, 1)
A	0	[0, .5)	[.5, 1)

The right endpoint of the (row *i* , column *j*) interval in the cumulative probability table is equal to the sum $P_{i,1} + P_{i,2} + \ldots + P_{i,j}$ of the first *j* transition probabilities from row *i* of the transition probability matrix M_n. Using Excel's ability to generate random numbers, a random sequence of decimal numbers between 0 and 1 were selected from a uniform distribution on [0, 1]. These numbers are used to select which note shall follow the current note: when the random number generated falls in a particular interval, the cumulative probability matrix determines what the next note shall be.

Let us describe the Excel formulas that were used to implement the procedure of generating sample melodies. The formula first looks at the current note, and finds that note's row in the cumulative probability table. Then, it compares the randomly generated number with the values in that note's row to find which interval the random number is contained. That interval then determines which note is returned by the function. This process is repeated until enough notes have been generated to match the number of notes in the original song. When determining which note to start the sample on, we decided that the first note of the original song would be the first note of all the samples generated from that particular song. This assumption assists in the processing the results of the samples, providing one aspect of consistency in an otherwise random scenario. It also allowed easier observation as to when, where, and how quickly the sample diverges from the original sequence.

The homogeneous and the non-homogeneous models have a few key differences. The homogeneous only has one of each of the three tables. Because all of the probabilities are pooled together, and there is no regard for differentiating how far the song has progressed, there is only one table that includes all of the transition probabilities. For each song, three sample songs were generated from the homogeneous model. For the non-homogeneous model, the exact number of transition probability tables depends on the number of phrases each song is divided into. For the purposes of this study, *Camptown Races* and *The Star-Spangled Banner* each have four musical phrases, *I Don't Want To Miss A Thing* has three phrases, and *Wachet auf* has five phrases. Each musical phrase comes with its own tally table, transition probability table, and cumulative table, each of which correspond with only the notes that appear within that phrase. Three complete samples were generated from each of the different transition probability tables. Within each model, there is an additional string of numbers that was used for testing that the model was functioning properly. Instead of random numbers, the testing numbers run from 0 to 1 incremented by 0.01. This was used to visually test the models to ensure that they were returning the proper new note name based for any given random number. All of the transition probability tables can be found in Appendix B.

Technical Problems:

There were a few problems in generating some of the non-homogeneous samples. Songs that are shorter in length have fewer opportunities for all of the states to communicate with one another, particularly when the song is being segmented into the different phrases. Here communicate means that "two states *i* and *j* that are accessible to each other" [Ro]. A good example of this occurring is the $M_{(2)}$, the second non-homogeneous matrix in the *Camptown Races* model. The phrase has an unusual cadence ending on the note F, the only occurrence of F within that phrase: this makes F a "dead-end" state. F is a dead-end because there is no note that F can transition to. This was a problem when generating the samples because through the normal process of the random transitions, any of the samples can transition to F before the very last note in the sample, which causes the model to halt. This "dead-end" concept is similar to that of recurrent, or absorbent, state. A state is said to be recurrent if the probability of transitioning from state *i* to state *i* is equal to 1. This means that once you've entered into state *i*, there is no leaving state *i*, thus the name absorbent. In the above example, if the note F had been a recurrent state, then every other note that follows in the piece would be an F as well. However, the model does not supply valid probabilities for transitions away from F. So instead of having the note F repeat over and over again, the model simply shut down and generated no more notes. Table 4 shows the $M_{(2)}$ matrix with the transition probabilities for F all being zero.

Table 4.

State	F	G	A	B♭	C	D
F	0	0	0	0	0	0
G	0.5	0	0	0.5	0	0
A	0	0.666667	0	0	0.333333	0
B♭	0	0	1	0	0	0
C	0	0	0.5	0	0.25	0.25
D	0	0	0	0	1	0

The approach that was determined best to circumvent this problem made use of *limiting probabilities*. Limiting probabilities are the probability that the "process will be in [a specific] state after a large number of transitions, and this value is independent of the initial [state]" [Ro]. When multiplying certain transition probability matrices by themselves many times over, the product converges to a matrix with identical rows whose entries are the limiting probabilities. We now describe the steps necessary to obtain a "replacement out-of-F transition row" for the transition matrix above.

1. Temporarily assume that state F is recurrent/absorbing (F→ F has transition probability equal to one) and then

2. Temporarily replace the other state's "transition-out" probabilities with those obtained by conditioning upon *not* ever entering state F from those states. In this case, only the "out-of-G" row was changed. This matrix is shown in Table 5.

Table 5.

State	F	G	A	B♭	C	D
F	1	0	0	0	0	0
G	0	0	0	1	0	0
A	0	0.666667	0	0	0.333333	0
B♭	0	0	1	0	0	0
C	0	0	0.5	0	0.25	0.25
D	0	0	0	0	1	0

3. Raise this matrix to a high power (the 7th power was sufficient for our application) to identify the common approximate limiting transition-out probabilities for all of the states except "out-of-F", that is, for all the non-absorbing states.

4. Replace the "out-of-F" row of the original M_2 with the common row of limiting probabilities obtained in Step 3, and use the original (unconditional) probabilities for the other states' rows. This new transition probability matrix (Table 6.) was then used for the generation process:

Table 6.

State	F	G	A	B♭	C	D
F	0	0.210534	0.315803	0.210505	0.210529	0.05263
G	0.5	0	0	0.5	0	0
A	0	0.666667	0	0	0.333333	0
B♭	0	0	1	0	0	0
C	0	0	0.5	0	0.25	0.25
D	0	0	0	0	1	0

This allowed the model to continue if it produced the note F anywhere within this phrase. The note to which F transitioned would be based upon the new limiting probabilities. Thus, these limiting probabilities are based on the assumption that *if* F could transition to another note, it would transition to the other notes based on the limiting probabilities shown in the first row of table 6. In *Camptown Races*, $M_{(2)}$ is equal to $M_{(4)}$, so limiting probabilities were needed for the generation of both of those phrases. The *Camptown Races* non-homogeneous model was the only model that necessitated the use of limiting probabilities. Very rarely would a song have an absorbing or a dead-end state on its own, since most music starts and ends on tonic. This regularity ensures there is the communication needed between states to avoid absorbing and dead-end states, particularly in the homogeneous model.

Another issue that appeared a few times during the sample generation process was transitioning between phrases. In three out of the twelve samples generated, the last note in one phrase would not appear in the following matrix. This problem understandably occurred only in the non-homogeneous samples. For instance, in Sample 1 of the non-homogeneous sample in *Camptown Races*, the $M_{(2)}$ phrase ends on B-flat, and B-flat does not appear at all in $M_{(3)}$. To avoid the model crashing and being unable to generate an entire song sample, the next note selected by the operator was based upon two criteria. The first note of the new phrase would be selected from the possible transitions the current note can transition to. Under the current phrase's transition matrix, the note with the highest transition probability would be selected to be the first note in the new phrase, given that that note exists in the new transition matrix. All three samples that required this treatment successfully used the highest transition probability note to continue the generation of the sample. While this error handling process does eliminate some of the randomness of the generation model, it was necessary for some process to be in place to allow the model to continue. In all subsequent calculations, the probabilities of these notes that were forced upon the model by this intervention were disregarded. Considering this study was the first to investigate non-homogeneous melodic musical generations, there were no prior examples to base troubleshooting the above two generation errors. If this study were to be expanded upon, one would want to consider other alternatives for these generation problems.

Analysis of Melodic Samples:

The first type of analysis done on the musical samples from the various models was a chi-square goodness-of-fit test on the samples' pitch class. The inspiration to utilize a chi-square goodness-of-fit test came from Soubhik Chakraborty et al [Ch], who used a chi-square goodness-of-fit test when evaluating frequencies of particular notes in the Ragas, melodic modes in classical Indian music. A chi-square goodness-of-fit test is used to fit a statistical model to observed data to see how well the model actually reflects the data. It tests the observed data, in this case the randomly generated samples of music, against the expected model, in our case the original piece of music. This analysis will evaluate the model's ability to generate a sample that resembles the original in various ways. This test was used on pitch class, to count and measure the frequency of each note occurring within the samples compared to the original. Some combinations of pitch class into a single category were necessary to ensure that each category had the minimum number of expected entries necessary to perform a chi-square goodness-of-fit test. The categories for the chi-square test were the various scale degrees appearing in the melodies, and possibly an additional "other" category for notes that are not part of the major scale. Another option for choosing which categories to combine was pairing notes that served similar purposes within the song, such as having similar harmonic functions. This type of pairing could be found in the *Camptown* Races model. The calculations for this analysis were performed within Microsoft Excel using the pitch class frequency count for both the original song and each of the six samples generated for each song, as well as the CHITEST function in Excel. This function returns the p-value for each test, which measures the degree to which the frequency of pitch classes in the sample matches the original melody. Table 7 shows the chi-square table for *Wachet auf*. All chi-square results can be found in Appendix C.

Table 7.

		Homogeneous			Non-Homogeneous		
Pitch Class	Original	Sample 1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3
E ♭	40	42	46	30	36	44	47
F	32	23	32	23	35	31	46
G	37	32	39	36	38	39	40
A ♭	21	23	26	25	28	20	17
B ♭	30	35	24	42	27	30	21
C	17	22	12	21	15	16	13
D	24	22	21	18	25	21	26
Other	10	12	11	16	7	10	1
$\chi 2$		**6.36799**	**5.34417**	**16.66136**	**4.518571**	**1.020801**	**20.26299**
P-value		**0.497497**	**0.61804**	**0.019715**	**0.71848**	**0.994486**	**0.005029**

As the P-values in Table 7 show, there is quite a bit of variety in the samples based on a chi-square goodness-of-fit test based on pitch class. The P-values range from numbers as large as 0.99 to very small numbers around 0.005. One definition of P-value is "the probability...of obtaining a test statistic value at least as contradictory to H_0 (null hypothesis) as the value that actually resulted. The smaller the P-value, the more contradictory is the data to H_0" [De]. The null hypothesis in our case is that the generated melody is similar to the original in terms of relative frequency of pitch class occurrences. To analyze the *Wachet auf* model, a chi-square value with seven degrees of freedom was used. Homogeneous Sample 3 and non-homogeneous Sample 3 have P-values less than 0.10, which indicates that these two generated melodies are significantly different statistically speaking from the original in terms of counts of pitch classes. Non-homogeneous Sample 2 has a P-value close to one, indicating that this sample very closely resembles the original in terms of the relative frequency of scale degrees appearing in the melody.

The exact same model resulted in melodies that were both incredibly similar to the original and incredibly divergent from the original. This variety of results can be seen in both the homogeneous and non-homogeneous cases for each of our songs. Of the twenty-four samples generated, six of the samples have a P-value less than 0.10 (P= 0.0414, 0.0067, 0.0001, 0.0313, 0.0197, 0.0050). The samples that are significantly dissimilar from the original appear in both the homogeneous and non-homogeneous models (two and four, respectively). Five out of the twenty-four samples had P-values greater than 0.90 (P= 0.9593, 0.9214, 0.9676, 0.9655, 0.9945), again rather evenly distributed between the two types (two for homogeneous and three for non-homogeneous). When it comes to generating samples that accurately resemble the original melody based on frequency of pitch class, both the homogeneous and non-homogeneous models were nearly equally good and equally bad at producing accurate results.

The next type of analysis done on the musical samples from the various models was a chi-square goodness-of-fit test on the samples' pitch intervals. In music, an interval is the distance between two notes' pitches. The frequency of the occurrence of certain intervals is one way composers impose a certain style within the piece. Measuring the frequency of intervals will help determine if the model produces samples that mimic the composer's style. The intervals analyzed within this study are all generic pitch intervals, such as "Third," "Fourth," "Fifth," etc. The four groupings for the categories for pitch intervals are measured on net change from the current note. The note could either identical, with a unison interval; up or down one, with the interval of a second; up or down two, with an interval of a third; or up or down three, with the interval of a fourth. Any intervals higher than a fourth are equivalent with a smaller interval measurement; for example, moving up a fifth is equivalent to moving down a fourth.

The frequency of each type of interval was counted and tabulated within Excel to perform the goodness-of-fit test. *Camptown Races* could not be included in this analysis because the original melody did not contain at least five of each type of interval. *Wachet auf* was also excluded from this type of analysis. The chi-square table for the

Aerosmith song is shown in Table 9. To analyze the *I Don't Want To Miss A Thing* model, a chi-square value with three degrees of freedom used. Non-homogeneous Sample 2 is another example of a poor fit; with a P-value of 0.031. However, the non-homogeneous model still produced melodies with very good fits as well. Homogeneous Sample 1 and Sample 3 both have P-values above 0.90 (P= 0.9290, 0.9970), indicating a good approximation of the original in regards to interval frequency.

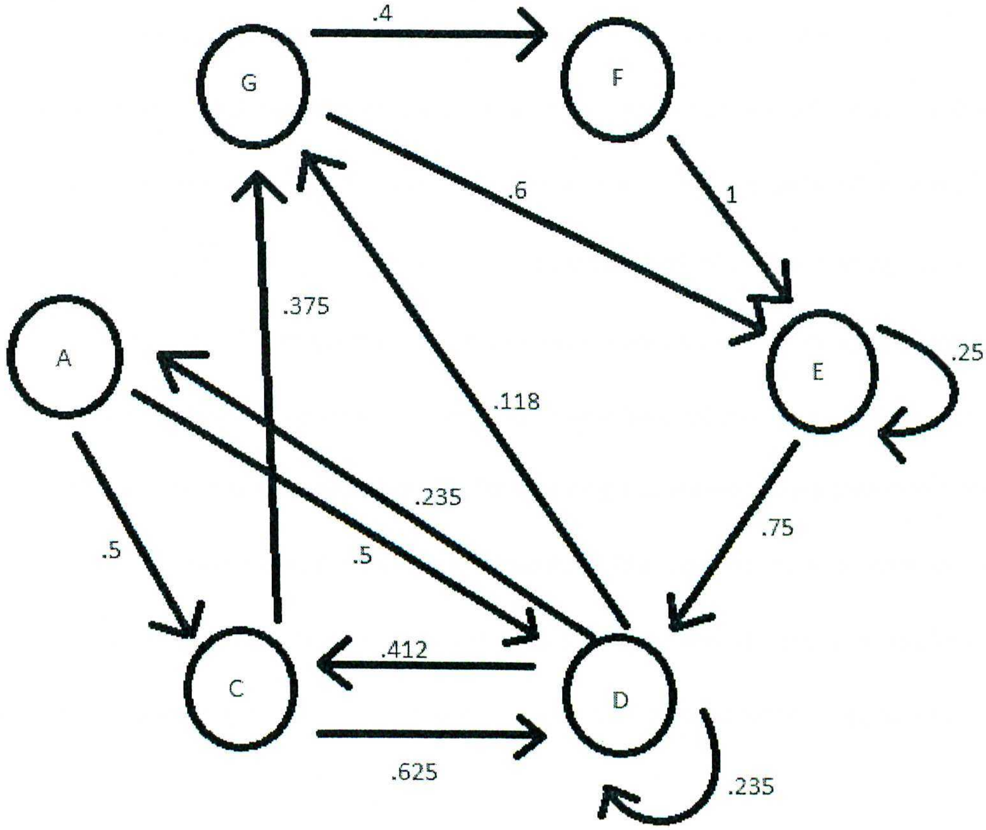
Table 9.

		Homogeneous			Non-Homogeneous		
Intervals	Original	Sample 1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3
Unison	30	31	35	29	34	35	33
Second	58	54	60	59	51	71	60
Third	39	41	30	39	48	28	41
Fourth	24	25	26	24	18	17	17
$\chi 2$		0.453426	3.145889	0.050575	4.955084	8.891357	2.513196
P-value		0.928998	0.369674	0.997021	0.1751158	0.030771	0.472911

In the twelve samples analyzed, the homogeneous model produces a greater frequency of more accurate results in terms of interval content. It didn't have any samples classified as bad fits, whereas the non-homogenous had one. The homogeneous model also produced two samples that were very good fits, while the non-homogeneous model didn't create any that could be called a very good fit. Because only twelve samples were gathered for the chi-square goodness-of-fit test for musical intervals, more samples could be generated to fully back the assertion that the homogeneous model is superior to the non-homogeneous model.

A reason why the homogeneous model seems to outperform the non-homogeneous model could be that in the non-homogeneous case it is much easier to be drawn into a repetitive loop that causes the sample to dramatically diverge from the original. This repetitive loop shares characteristics with a *degenerate cycle*. A degenerate cycle occurs when a state always returns back to itself. In the case of musical modeling, this would be a sequence of notes that cannot be escaped once they are entered. In the non-homogeneous Sample 2, there are an unusually high number of C→D transitions and D→C transitions. This causes the song to bounce between these two notes, increasing the frequency of second intervals disproportionately high. While the C→D and D→C transitions do not quite form the black hole of a degenerate cycle, they do create a strong pull towards themselves. Escaping that loop, while not impossible, is difficult for the model. Each of the non-homogeneous samples for *I Don't Want To Miss A Thing* have a high number of the C→D and D→C transitions, but Sample 2 has the most, which causes that sample to be very bad fit in both the pitch class and interval goodness-of-fit tests. Figure 1 shows a flowchart using the transition probabilities of $M_{(2)}$ matrix from the Aerosmith song. The high probability of transitioning to and from C and D as well as the pull back towards the notes C and D from all the other notes explains the possibility of seeing a higher than expected frequency of Cs and Ds.

Figure 1.



Geometric Mean Analysis:

Because probability is the basis for the generation of all of the musical samples, it is sensible to do some analysis of the probabilities related to each song. Each sample has an associated probability for each transition between notes that created the melody, so the Excel model kept track of each song's probability of generation. We need to calculate a normalized measurement that is able to compare songs of various lengths. For example, the forty-five notes of *Camptown Races* will generate a much larger product of transition probabilities than the 211 notes from the *Wachet auf* arrangement. Longer songs will have lower probabilities simply because more fractions have been multiplied together. We need an "average transition probability" that is independent from the number of notes in a song. In order to compare the song's probability on a relative and normalized basis, one must use the *geometric mean* of the probabilities. The geometric mean of *n* probabilities a_1 through a_n is defined as the n^{th} root of their product, $\sqrt[n]{a_1 \cdot a_2 \cdot \ldots \cdot a_n}$, which can be thought of as the average of a set of factors in a product. Through the use of the geometric mean, the probabilities of all the songs can be compared against one another, regardless of the number of notes in the song. We calculated the geometric mean for the original melody based on both the homogeneous Markov chain and the non-homogeneous Markov chains so as to have an appropriate basis for comparison. We also calculated the geometric means of each of the six samples for each song, using the specific model that generated that particular sample.

Table 10 shows the geometric means of each homogeneous sample for each of the songs analyzed, as well as the geometric mean for the original melody using the homogeneous model. The ratios of the geometric means comparing the sample to the original were calculated to show the magnitude of how much each sample differed from the original. Of the nine total samples gathered, five had geometric means greater than their respective original, and the other four had geometric means less than their respective original melody. Only one song, *The Star-Spangled Banner*, had all three samples with geometric means higher than the original. This leads to the conclusion that the original melody contains sequences of notes that would make this melody unique, in the sense that these sequences seem to be hard to generate using our random processes. The aural interpretation of this phenomenon is discussed below. When it comes to pitch class and interval frequency, Aerosmith's Sample 3 was one of the best fits. One might assume that because of quality of fit in those categories, the ratio of its geometric mean to the original's geometric mean be close to 1; however, this ratio ends being the furthest from 1 all of the samples tested. This might suggest that there is not much relationship between the chi-square goodness-of-fit tests and the geometric mean analysis. With more time, more samples could be generated to investigate how close the geometric averages of many such sample geometric means would be to the geometric mean of the original melody.

Table 10.

Geometric means: homogeneous model

Song	*Camptown Races*	Sample/ Original Ratio	*Star-Spangled Banner*	Sample/ Original Ratio	*I Don't Want To Miss A Thing*	Sample/ Original Ratio
Original	0.352378	1	0.274331	1	0.246231	1
Sample 1	0.393844	1.117677	0.279639	1.019350	0.251631	1.021930
Sample 2	0.350886	0.995766	0.289278	1.054484	0.231797	0.941382
Sample 3	0.337786	0.958592	0.296793	1.081881	0.226341	0.919221

Table 11 shows the geometric means of each non-homogeneous sample for each of the songs analyzed, as well as the geometric mean for the original melody using the non-homogeneous model. Also shown are the ratios of the geometric means of the samples to those of the original. Logically, all of the non-homogeneous geometric means are greater than their homogeneous counterparts simply because the probabilities used in the non-homogeneous models are much greater than those in the homogeneous models. Of the nine samples gathered from this method, two of the samples had geometric means were greater than their respective original, while seven were less than their respective original melody. The song that stands out here is the Aerosmith song, which was the only song to have all the geometric means fall below that of the original. An explanation for this could found in the nature of pop music. Contemporary pop music tends to have a great quantity of notes all within short intervals of each other, and these notes tend to be repeated frequently. This is because the musician strives to put more words into the melody in a short amount of time. This was expected ahead of time, and the arrangement used in this study tried to lessen the effect of these repeated notes by reducing repeating 16th notes into a single quarter note. However, this pop music effect still appeared within the geometric analysis as well as the previously mentioned goodness-of-fit testing. The frequency of these repeated notes and repetitive loops drove down the geometric means of the samples on a relative basis to the original by not allowing the melody to explore other sequences of notes that would have increased the geometric mean. *I Don't Want To Miss A Thing*'s non-homogeneous Sample 1 has the closest fit regarding geometric mean, but appears relatively average when analyzed through the pitch class and interval frequency testing methods. *I Don't Want To Miss A Thing*'s non-homogeneous Sample 2 was by far the worst performer in regards to the goodness-of-fit tests, but the geometric mean doesn't have a dramatically different geometric mean. These few test cases again show that the various methods of testing for accuracy are not strongly related to each other.

Table 11.

Geometric means: non-homogeneous model

Song	*Camptown Races*	Sample/ Original Ratio	*Star-Spangled Banner*	Sample/ Original Ratio	*I Don't Want To Miss A Thing*	Sample/ Original Ratio
Original	0.473151	1	0.392975	1	0.333049	1
Sample 1	0.541890	1.145278	0.381902	0.971823	0.332785	0.999206
Sample 2	0.464444	0.981596	0.378186	0.962369	0.319727	0.959999
Sample 3	0.436695	0.922949	0.405878	1.032836	0.324489	0.974297

In an attempt to find an upper limit for the geometric mean for each song, we investigated the "greedy" case. Here the term greedy is being used rather loosely, where the greedy case is defined as the melody that chooses the most likely next note at every opportunity. In the event that there were "multiple most likely" next notes, the path that results in the highest geometric mean is chosen. Calculating the truly highest geometric mean in the greedy case turned out to be incredibly difficult due to the unlimited number of paths and starting points. We placed a constraint upon the calculation, which stated that the greedy sample must start on the same note as the original. Because all of the music samples were generated in this method, the calculation maintained a sense of consistency. However, this number would not be the true “most likely” sequence of notes because there could be numerous paths that utilize higher probabilities than the sequence that was forced to start on a particular note. Moreover, by making a greedy choice at the first transition between pitches, the pitch sequence could be steered away from the starting point for a path that has a higher geometric mean of probabilities. Table 12 compares the geometric means of the original melody calculated from both the homogeneous and non-homogeneous methods to the conditional greedy scenario. For each of the three songs, the geometric means of the greedy case appears to act as an upper limit for the samples in both the homogeneous and non-homogeneous cases.

Table 12.

Song	*Camptown Races*	Sample/ Original Ratio	*Star-Spangled Banner*	Sample/ Original Ratio	*I Don't Want To Miss A Thing*	Sample/ Original Ratio
Original	0.352378	1	0.274331	1	0.246231	1
Greedy Homogeneous	0.562450	1.596158	0.368421	1.342980	0.391809	1.591226
Original Non Homogeneous	0.473151	1	0.392975	1	0.333049	1
Greedy Non-Homogeneous	0.655352	1.385079	0.604589	1.538494	0.449001	1.348154

Motive Analysis:

In music, motives are important phrases that are frequently found within the composition. The most famous example of a motive can be found in the first four notes of Beethoven's Symphony No. 5. While none of the songs analyzed in this study have an instantly recognizable motive when compared to Beethoven's, they do contain certain sequences of notes that most listeners would be able to identify and associate with the respective songs. We included these motives, along with other important sequences of notes, in the sequence analysis. The geometric means for all of the motives analyzed are displayed in Table 13. On all but a few occasions, the geometric means for the non-homogeneous models are greater than their respective homogeneous counterparts. This type of analysis is more subjective than the previous types of analysis due to the (mathematically) arbitrary selection of the motives. The motives selected varied in length, from four to fourteen notes long, with the average length being six notes. All of the motives can be found in Appendix D.

Table 13.

	Camptown Races		*Star-Spangled Banner*		*I Don't Want To Miss A Thing*	
	Homogeneous	Non-Homogeneous	Homogeneous	Non-Homogeneous	Homogeneous	Non-Homogeneous
Motive 1	0.33298	0.39685	0.21669	0.29710	0.29450	0.38385
Motive 2	0.36739	0.62996	0.39490	0.48836	0.33126	0.39447
Motive 3	0.49617	0.63894	0.34824	0.39203	0.28181	0.37927
Motive 4	0.21350	0.50000	0.20087	0.26085	0.24634	0.22690
Motive 5	0.29337	0.36889	0.12784	0.29755	0.20965	0.16708
Motive 6	-	-	0.27165	0.36559	0.21302	0.35335
Motive 7	-	-	0.38888	0.58989	0.18125	0.32439
Motive 8	-	-	0.26451	0.38606	0.21980	0.41714
Motive 9	-	-	0.21703	0.41017	0.28002	0.47690
Motive 10	-	-	0.20279	0.39936	0.36212	0.41474
Motive 11	-	-	-	-	0.21458	0.33817

Throughout the three songs analyzed, a full occurrence of a motive was only observed a total of seventeen times, seven times in *Camptown Races* and ten times in *I Don't Want To Miss A Thing*. The "Do-Da Do-Da" motive, or the A→G→A→G sequence in the first line in *Camptown Races* labeled Motive 2 in Table 13, occurred in its entirety the most often, once in two of the homogeneous models and twice in non-homogeneous Sample 1. The geometric mean for this motive isn't the highest by any means, but it still appeared the more frequently than any other motive. In the non-homogeneous model, this is understandable because once the sequence arrives at A, it has a strong tendency to bounce between G and A. The fact that only two pitches are used in this motive is also a factor, especially because of the high probabilities of transitioning between those two notes. The most interesting thing about the non-homogeneous Sample 1 event was that it occurred at the exact same place chronologically in the song as the original. Only one other motive, Motive 5 in *Camptown Races*'s non-homogeneous Sample 2, achieved this feat. No other occurrence of a motive aligned in this way. The non-homogeneous models were better at producing full length motives, generating ten of the seventeen, with the other seven generated by the homogeneous models. They were also better at placing those motives in the same spot as the original. This is unsurprising given the "limited" information contained in any one non-homogeneous transition matrix; however, the homogeneous models weren't too far behind. The advantage of being able to avoid degenerate/cyclic patterns might help the homogeneous models to produce recognizable motives.

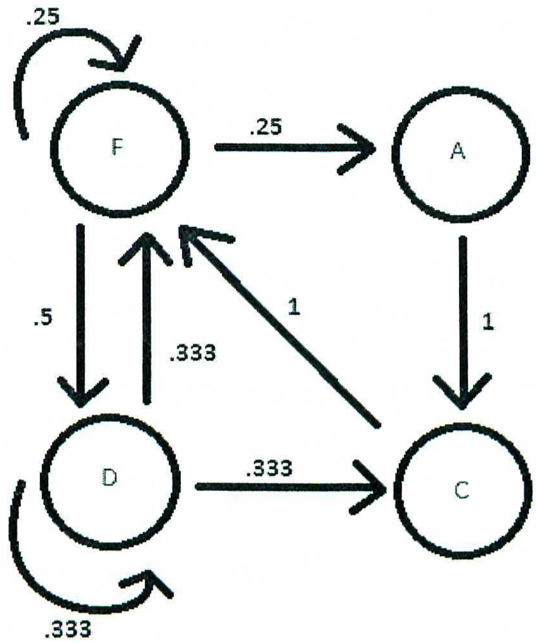
The non-homogeneous model was also better at producing partial motives, or sequences of notes that resemble the motive briefly before diverging from the true motive. The non-homogeneous models generated 212 partial motives, compared to the homogeneous models' ninety-four. However, these numbers are most likely skewed by some of the flaws of the non-homogeneous model. The non-homogeneous Sample 2 from *I Don't Want To Miss A Thing* again appears to be distorting the results. Motives 8, 10, and 11 all begin with either a C→D→C sequence or a D→C→D sequence. These were the only motives to have occurrences in the double digits, all within that one sample. Even after discarding the unique results of Sample 2's Motives 8, 10, and 11, the non-homogeneous models still produce a fair number more partial motives. There could be more subtle sequences of notes that are boosting the non-homogeneous models' ability to generate partial motives.

We also investigated whether the appearance of certain motives depended on a particular sequence of preceding notes. We calculated a retrospective count of the notes leading up to the motive. Roughly one-third of all the motives had matching preceding notes, with most motives having only one matching preceding note. Only three of the motives had a majority of their sequences with correct preceding notes: from *Camptown Races* Motives 3 and 5, and from *The Star-Spangled Banner* Motive 7. Along with having all but one of the sequences record preceding notes, they also have a higher frequency of longer strings of preceding notes. Motive 5 is found in the third line of phrase of *Camptown Races*, and all occurrences of this motive depend heavily on the preceding notes. They also only occur when the model is generating from $M_{(3)}$, the non-homogeneous transition probability matrix for the third phrase of *Camptown Races*. The high probabilities shown in the flowchart of transition probabilities for $M_{(3)}$ (Figure 2) exhibit why this motive is so dependent on the preceding sequence of notes. While some motives are dependent on preceding notes, these specific motives are really only generated because the model produced favorable conditions for them to be generated. When the model does produce these favorable conditions, the model generates very good representations of the original for that brief span of time.

could be more subtle sequences of notes that are boosting the non-homogeneous models' ability to generate partial motives.

We also investigated whether the appearance of certain motives depended on a particular sequence of preceding notes. We calculated a retrospective count of the notes leading up to the motive. Roughly one-third of all the motives had matching preceding notes, with most motives having only one matching preceding note. Only three of the motives had a majority of their sequences with correct preceding notes: from *Camptown Races* Motives 3 and 5, and from *The Star-Spangled Banner* Motive 7. Along with having all but one of the sequences record preceding notes, they also have a higher frequency of longer strings of preceding notes. Motive 5 is found in the third line of phrase of *Camptown Races*, and all occurrences of this motive depend heavily on the preceding notes. They also only occur when the model is generating from $M_{(3)}$, the non-homogeneous transition probability matrix for the third phrase of *Camptown Races*. The high probabilities shown in the flowchart of transition probabilities for $M_{(3)}$ (Figure 2) exhibit why this motive is so dependent on the preceding sequence of notes. While some motives are dependent on preceding notes, these specific motives are really only generated because the model produced favorable conditions for them to be generated. When the model does produce these favorable conditions, the model generates very good representations of the original for that brief span of time.

Figure 2.



Aural Analysis:

The aural analysis of each of the samples is a more subjective but still important part of the analysis of the generated melodies. We looked for a relationship between good fit songs according to the chi-square tests and how closely they resembled the original melody in terms an informal sense of what the melody sounds like. With *Camptown Races*, two particular samples stand out. The non-homogeneous Sample 2 song was conclusively the best sounding of all the *Camptown Races* samples. It generated a P-value of 0.9213 in the pitch class analysis. By comparison, non-homogeneous Sample 3 sounded the most random and furthest from the original. However, it generated the greatest P-value of all the *Camptown Races* samples with p= 0.9676.

Within *The Star-Spangled Banner*, non-homogeneous Sample 1 didn't sound very similar to the original. It deviated from the original very quickly and didn't seem to have any particular phrases that would enable one to identify the original song. The melody bounced between a few select notes. This is in contrast to homogenous Sample 3, which also didn't sound like the original and the notes appeared to wander aimlessly without any clear path or destination. Here, the non-homogeneous Sample 1 generated a very low P-value of 0.0414 in the pitch class analysis, while the homogenous Sample 3 produced a P-value of 0.8367. The results from both *Camptown Races* and *The Star-Spangled Banner* point to a basic principle of chi-square testing. With low P-values, we can hear the bad fits within the song. However, if the song has a high P-value, we can't determine if the song will sound like the original or not just based upon the P-value. The homogeneous samples from *The Star-Spangled Banner* model had a variety of results, with some having similar shaping to the original, while others only recorded one or two phrases within the whole sample that sounded comparable to the original. It appears that capturing those unique aspects of the original melody was indeed difficult for the model to do.

Within the many samples, there was quite a variety of deviations from the original melody. Some samples became dramatically repetitive and even degenerative. Others wandered from the original, but then produced a familiar phrase that would cause one to relate it back to the original melody. Overall, the best sounding samples came from the *Camptown Races* models. This is unsurprising due to the pentatonic nature of the song, relying on a five-note scale as opposed to the more common seven-note major scale. The least successful song in this type of analysis was *I Don't Want To Miss A Thing*. Again, this seems to be due to the repetitive nature of pop music, relying on duplicating notes in order to get more lyrics within a short span of time. Using Markov processes for melodic musical generation is best suited for genres that use simple melodies and phrases, such as folk tunes, hymns, and songs that are easy for an amateur singer to perform.

Concluding Remarks:

There are numerous ways to expand upon this type of mathematical musical analysis. One could alter the states of the models; instead of using pitch class, one could differentiate between octaves of notes. Also, instead of duplicating the original's rhythmic pattern, one could incorporate the individual rhythm paired with the note when defining the states of the Markov model. Also, instead of using only first-order Markov processes, one could create a higher-order Markov chain by defining the states in the Markov processes to be short sequences containing multiple notes. Each of these approaches would dramatically increase the number of possible states, and a more sophisticated computing method would be required. There could also be other alternative mathematical methods for analyzing the musical samples not used in this project. The motive analysis portion could also use refining. Instead of each motive needing a minimum of three correct consecutive notes to be counted as an iteration of a partial motive, each motive should have its own unique minimum to be chosen ahead of time. This would erase the error of overweighting certain partial motives that are repetitive in nature, as seen in motive analysis of the Aerosmith song.

In conclusion, the various types of analysis used on the homogeneous and the non-homogeneous Markov models did not reveal a clearly superior method of melodic musical generation. After performing a chi-square goodness-of-fit test on frequency of pitch class and musical interval, the results showed that both models could produce very accurate and very inaccurate samples. The geometric mean analysis of the probabilities of the samples generated also did not expose a better model, and the results did not clearly correspond with those of the goodness-of-fit tests. While the non-homogeneous model produced a higher frequency of accurate motives, the ability of the non-homogeneous model to create repetitive loops skewed the results of this analysis. Because of the generation challenges observed and the possibility of repetitive and degenerate cycles found in the non-homogeneous Markov model, the homogeneous model holds the slight edge in ability to produce accurate melodic musical samples.

References

[Al] Alamkan, Claude, William P. Birmingham, and Mary H. Simoni. "Stylistic Structures: An Initial Investigation of the Stochastic Generation of Tonal Music." Technical Report. University of Michigan, 1999.

[Am] Ames, Charles. "The Markov Process as a Compositional Model: A Survey and Tutorial." *Leonardo* 2nd ser. 22 (1989): 175-87. *JSTOR*. MIT Press.

[Ch] Chakraborty, Soubhik, Sandeep S. Solanki, Sayan Roy, Shivee Chauhan, Sanjaya S. Tripathy, and Kartik Mahto. "A Statistical Approach to Modeling Indian Classical Music Performance." 25 Oct. 2008. Unpublished. *Computing Research Repository*. <http://arxiv.org/abs/0809.3214>.

[Da] Daniel, James W. *Multi-state Transition Models for Actuarial Applications*. Society of Actuaries. Oct. 2007. <http://www.soa.org/files/pdf/edu-2008-spring-mlc-24-2nd.pdf>.

[De] Devore, Jay L., and Kenneth N. Berk. *Modern Mathematical Statistics with Applications*. 1st ed. Belmont, CA: Thomson Brooks/Cole, 2007.

[L] Loy, Gareth. *Musimathics: The Mathematical Foundations of Music*. Cambridge, MA: MIT, 2006.

[Ro] Ross, Sheldon Mark. *Introduction to Probability Models*. 10th ed. San Diego, CA: Academic Press, 2010.

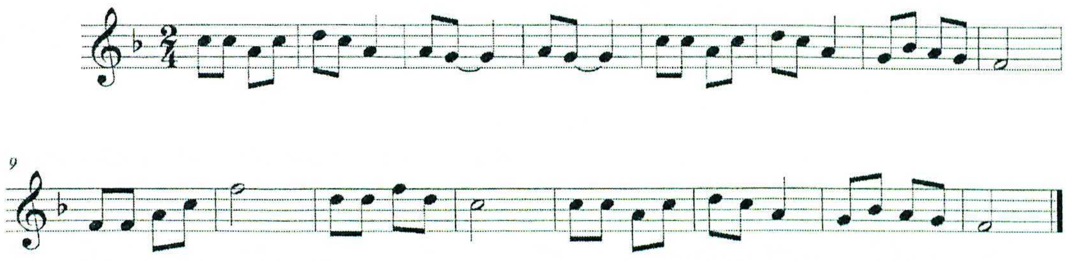
[Ru] Rubey, Michael, and Morris Marx. "Stochastic Generation of Music." MS Thesis. University of West Florida, 2009. <http://sites.google.com/site/stochasticmusic/>.

[T] Taylor, Howard M., and Samuel Karlin. *An Introduction to Stochastic Modeling*. Orlando: Academic Press, 1984.

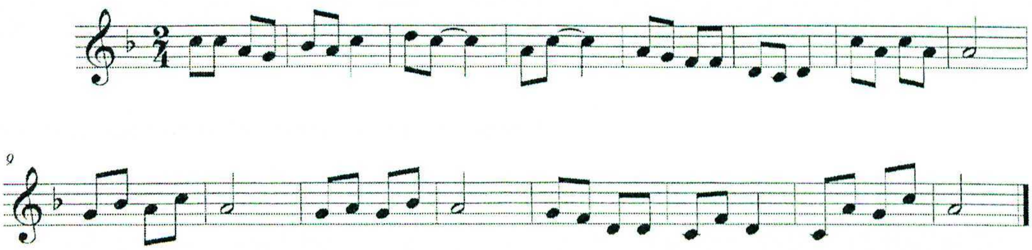
Appendix A

Musical Arrangements and Samples Generated

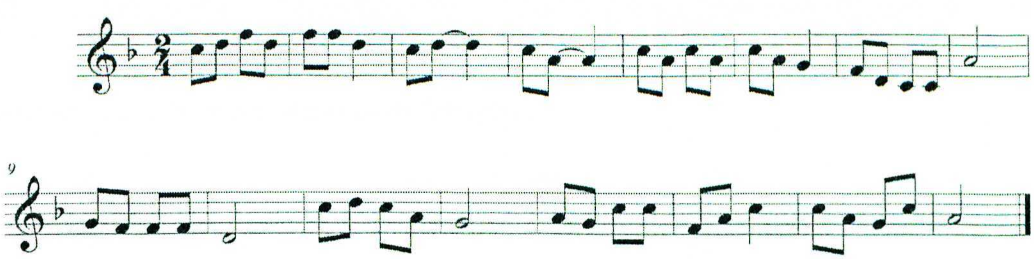
Camptown Races Original



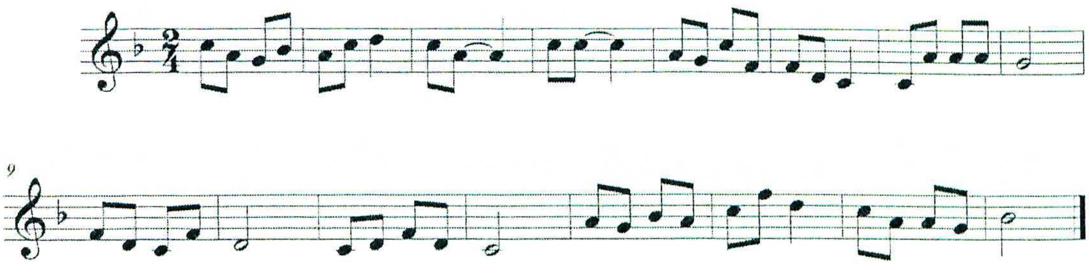
Camptown Races Homogeneous Sample 1



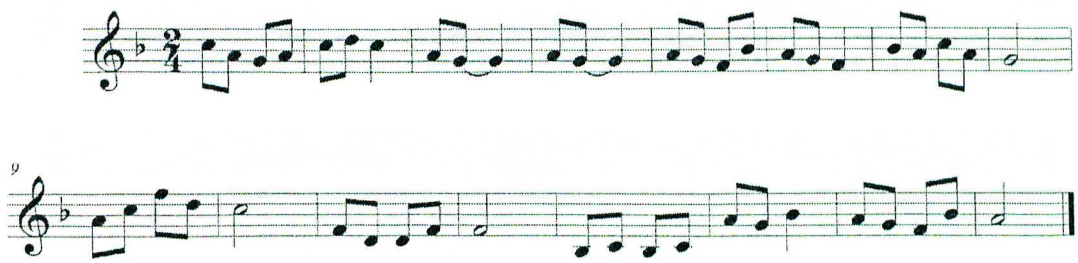
Camptown Races Homogeneous Sample 2



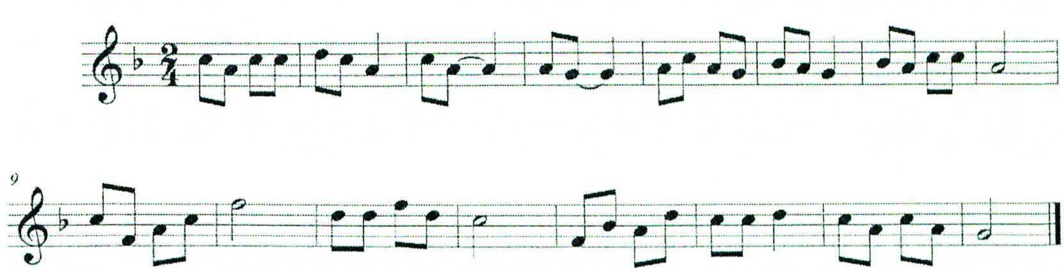
Camptown Races Homogeneous Sample 3



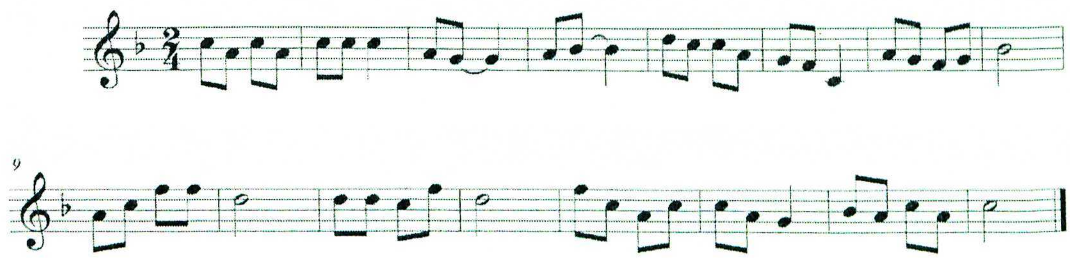
Camptown Races Non-Homogeneous Sample 1



Camptown Races Non-Homogeneous Sample 2



Camptown Races Non-Homogeneous Sample 3



The Star-Spangled Banner Original

The Star-Spangled Banner Homogeneous Sample 1

The Star-Spangled Banner Homogeneous Sample 2

The Star-Spangled Banner Homogeneous Sample 3

The Star-Spangled Banner Non-Homogeneous Sample 1

The Star-Spangled Banner Non-Homogeneous Sample 2

The Star-Spangled Banner Non-Homogeneous Sample 3

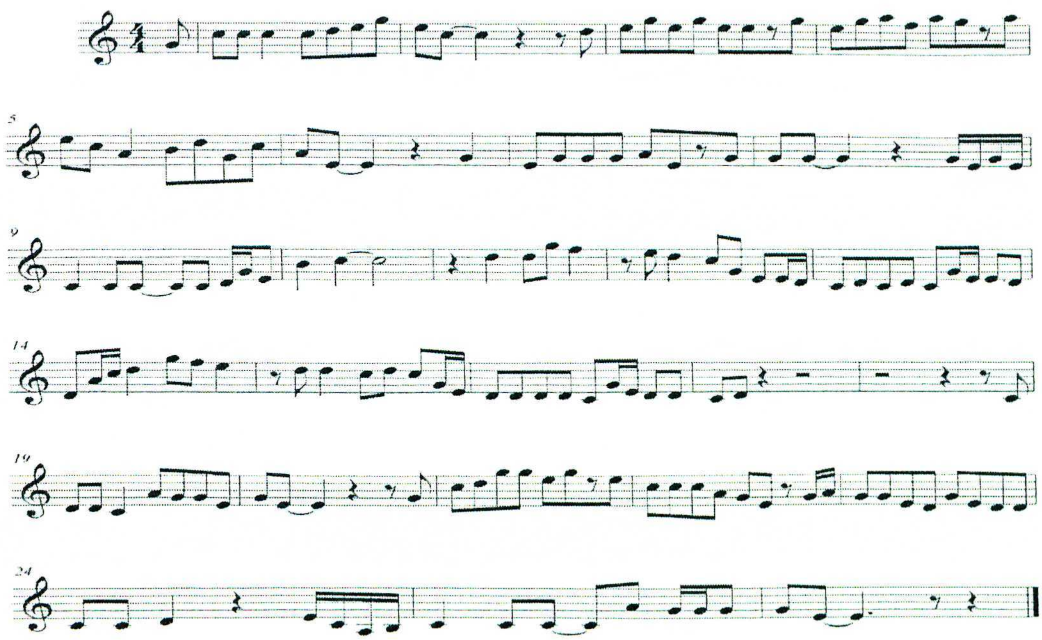
I Don't Want To Miss A Thing Original

I Don't Want To Miss A Thing Homogeneous Sample 1

I Don't Want To Miss A Thing Homogeneous Sample 2

I Don't Want To Miss A Thing Homogeneous Sample 3

I Don't Want To Miss A Thing Non-Homogeneous Sample 1



I Don't Want To Miss A Thing Non-Homogeneous Sample 2

I Don't Want To Miss A Thing Non-Homogeneous Sample 3

Wachet auf, ruft uns die Stimme Original

Wachet auf, ruft uns die Stimme Homogeneous Sample 1

Wachet auf, ruft uns die Stimme Homogeneous Sample 2

Wachet auf, ruft uns die Stimme Homogeneous Sample 3

Wachet auf, ruft uns die Stimme Non-Homogeneous Sample 1

Wachet auf, ruft uns die Stimme Non-Homogeneous Sample 2

Wachet auf, ruft uns die Stimme Homogeneous Sample 3

Appendix B

Probability Tables

Camptown Races

Homogeneous

State	F	G	A	B♭	C	D
F	0.4	0	0.2	0	0	0.4
G	0.333333	0	0.1666667	0.3333333	0.166667	0
A	0	0.545455	0.0909091	0	0.363636	0
B♭	0	0	1	0	0	0
C	0.071429	0	0.4285714	0	0.285714	0.214286
D	0.166667	0	0	0	0.666667	0.166667

Non-Homogeneous

$M_{(1)}$

State	F	G	A	B♭	C	D
F	0	0	0	0	0	0
G	0	0	1	0	0	0
A	0	0.5	0.25	0	0.25	0
B♭	0	0	0	0	0	0
C	0	0	0.5	0	0.25	0.25
D	0	0	0	0	1	0

$M_{(2)}$ and $M_{(4)}$

State	F	G	A	B♭	C	D
F	0	0.2105335	0.3158028	0.210505	0.210529	0.0526295
G	0.5	0	0	0.5	0	0
A	0	0.6666667	0	0	0.333333	0
B♭	0	0	1	0	0	0
C	0	0	0.5	0	0.25	0.25
D	0	0	0	0	1	0

$M_{(3)}$

State	F	G	A	B♭	C	D
F	0.25	0	0.25	0	0	0.5
G	0	0	0	0	0	0
A	0	0	0	0	1	0
B♭	0	0	0	0	0	0
C	1	0	0	0	0	0
D	0.333333	0	0	0	0.333333	0.333333

The Star-Spangled Banner

Homogeneous

State	C	D	E	F	F#	G	A	B
C	0.217391	0.086957	0.347826	0	0	0.130435	0	0.217391
D	0.636364	0	0.272727	0.090909	0	0	0	0
E	0.181818	0.318182	0.090909	0.181818	0.136364	0.090909	0	0
F	0	0.125	0.375	0.25	0	0.25	0	0
F#	0	0	0	0	0	1	0	0
G	0.263158	0	0.315789	0.052632	0	0.368421	0	0
A	0	0.166667	0	0	0	0	0.333333	0.5
B	0.375	0	0	0	0	0.125	0.5	0

Non-Homogeneous

$M_{(1)}$ and $M_{(2)}$

State	C	D	E	F	F#	G	A	B
C	0.1666667	0	0.5	0	0	0.166667	0	0.166667
D	1	0	0	0	0	0	0	0
E	0.3333333	0.3333333	0	0	0.166667	0.166667	0	0
F	0	0	0	0	0	0	0	0
F#	0	0	0	0	0	1	0	0
G	0.1666667	0	0.5	0	0	0.333333	0	0
A	0	0	0	0	0	0	0	1
B	0.5	0	0	0	0	0	0.5	0

$M_{(3)}$

State	C	D	E	F	F#	G	A	B
C	0	0	0.5	0	0	0	0	0.5
D	0.5	0	0.5	0	0	0	0	0
E	0	0.2857143	0.285714	0.285714	0.142857	0	0	0
F	0	0	0.4	0.4	0	0.2	0	0
F#	0	0	0	0	0	1	0	0
G	0	0	0	0.5	0	0.5	0	0
A	0	0	0	0	0	0	0	1
B	0.5	0	0	0	0	0	0.5	0

$M_{(4)}$

State	C	D	E	F	F#	G	A	B
C	0.428571	0.2857143	0	0	0	0	0	0.285714
D	0.4	0	0.4	0.2	0	0	0	0
E	0	0.3333333	0	0.666667	0	0	0	0
F	0	0.3333333	0.333333	0	0	0.333333	0	0
F#	0	0	0	0	0	0	0	0
G	0.75	0	0	0	0	0.25	0	0
A	0	0.3333333	0	0	0	0	0.66667	0
B	0	0	0	0	0	0.5	0.5	0

I Don't Want To Miss A Thing

Homogeneous

State	C	D	E	F	G	A	B
C	0.266667	0.366667	0	0.0333333	0.133333	0.2	0
D	0.419355	0.225806	0.09677419	0	0.129032	0.129032	0
E	0.10345	0.31035	0.1379310	0	0.34483	0	0.103458
F	0	0	0.6666667	0	0.33333	0	0
G	0.05556	0	0.4444444	0.055556	0.27778	0.16667	0
A	0.11765	0.11765	0.2352941	0	0.35294	0.05882	0.11765
B	0.4	0.6	0	0	0	0	0

Non-Homogeneous

$M_{(1)}$

State	C	D	E	F	G	A	B
C	0.375	0.25	0	0	0	0.375	0
D	0.4	0.2	0.2	0	0.2	0	0
E	0.083333	0.0833333	0.083333	0	0.583333	0	0.166667
F	0	0	0	0	0	0	0
G	0.052632	0	0.368421	0	0.315789	0.263158	0
A	0	0	0.333333	0	0.444444	0.111111	0.111111
B	0.333333	0.6666667	0	0	0	0	0

$M_{(2)}$

State	C	D	E	F	G	A	B
C	0	0.625	0	0	0.375	0	0
D	0.411765	0.235294	0	0	0.117647	0.2352941	0
E	0	0.75	0.25	0	0	0	0
F	0	0	1	0	0	0	0
G	0	0	0.6	0.4	0	0	0
A	0.5	0.5	0	0	0	0	0
B	0	0	0	0	0	0	0

$M_{(3)}$

State	C	D	E	F	G	A	B
C	0.384615	0.30769	0	0.07692	0	0.23077	0
D	0.5	0.25	0.125	0	0.125	0	0
E	0.222222	0.22222	0.11111	0	0.33333	0	0.11111
F	0	0	0	0	1	0	0
G	0.083333	0	0.5	0	0.33333	0.08333	0
A	0	0	0.25	0	0.5	0	0.25
B	0.5	0.5	0	0	0	0	0

Wachet auf, ruft uns die Stimme

Homogeneous

State	E♭	F	G	A♭	A	B♭	C	D♭	D
E♭	0.075	0.325	0.075	0.05	0.025	0.075	0.075	0.025	0.275
F	0.375	0	0.25	0.2813	0	0.0625	0	0	0.0313
G	0.1081	0.4595	0.0811	0.1081	0.0811	0.0811	0	0	0.0811
A♭	0	0.0476	0.8571	0	0	0	0.0476	0	0.0476
A	0	0	0.3333	0	0.1111	0.5556	0	0	0
B♭	0.2069	0.0345	0.069	0.2069	0.1034	0.2069	0.1724	0	0
C	0.0588	0	0	0	0.0588	0.4706	0	0	0.4118
D♭	0	0	0	0	0	0	1	0	0
D	0.625	0	0	0	0	0.0417	0.2917	0	0.0417

Non-Homogeneous

$M_{(1)}$

State	E♭	F	G	A♭	A	B♭	C	D♭	D
E♭	0	0.8	0	0	0	0.2	0	0	0
F	0	0	0.3333	0.6667	0	0	0	0	0
G	0.2	0.2	0.2	0	0	0.2	0	0	0.2
A♭	0	0	1	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0
B♭	0.5	0	0	0.5	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0
D♭	0	0	0	0	0	0	0	0	0
D	1	0	0	0	0	0	0	0	0

$M_{(2)}$

State	E♭	F	G	A♭	A	B♭	C	D♭	D
E♭	0.1429	0.1429	0.0714	0	0.0714	0	0.0714	0	0.5
F	0.4444	0	0.1111	0.1111	0	0.2222	0	0	0.1111
G	0	0.6667	0	0	0.3333	0	0	0	0
A♭	0	0	1	0	0	0	0	0	0
A	0	0	0.3333	0	0.1111	0.5556	0	0	0
B♭	0.125	0.0625	0.125	0.0625	0.1875	0.1875	0.25	0	0
C	0.1	0	0	0	0.1	0.5	0	0	0.3
D♭	0	0	0	0	0	0	0	0	0
D	0.4545	0	0	0	0	0.0909	0.4545	0	0

$M_{(3)}$

State	E♭	F	G	A♭	A	B♭	C	D♭	D
E♭	0	1	0	0	0	0	0	0	0
F	0	0	0.3333	0.6667	0	0	0	0	0
G	0.2	0.2	0.2	0	0	0.2	0	0	0.2
A♭	0	0	1	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0
B♭	0.5	0	0	0.5	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0
D♭	0	0	0	0	0	0	0	0	0
D	1	0	0	0	0	0	0	0	0

$M_{(4)}$

State	E♭	F	G	A♭	A	B♭	C	D♭	D
E♭	0	0.4286	0.1429	0.1429	0	0	0.1429	0	0.1429
F	0.5714	0	0.1429	0.2857	0	0	0	0	0
G	0	0.6667	0	0.3333	0	0	0	0	0
A♭	0	0	0.8	0	0	0	0.2	0	0
A	0	0	0	0	0	0	0	0	0
B♭	0	0	0	0.3333	0	0.3333	0.3333	0	0
C	0	0	0	0	0	0.25	0	0	0.75
D♭	0	0	0	0	0	0	0	0	0
D	0.75	0	0	0	0	0	0.25	0	0

$M_{(5)}$

State	E♭	F	G	A♭	A	B♭	C	D♭	D
E♭	0.1	0.2	0.1	0.1	0	0	0.1	0.1	0.3
F	0.5714	0	0.4286	0	0	0	0	0	0
G	0.1429	0.5714	0	0.2857	0	0	0	0	0
A♭	0	0	0.75	0	0	0	0	0	0.25
A	0	0	0	0	0	0	0	0	0
B♭	0.3333	0	0	0.3333	0	0.3333	0	0	0
C	0	0	0	0	0	0.6667	0	0	0.3333
D♭	0	0	0	0	0	0	1	0	0
D	0.6667	0	0	0	0	0	0.1667	0	0.1667

Appendix C

Chi-Square Goodness-of-Fit Results

Pitch Class Frequency testing

Camptown Races

Pitch Class	Original	Sample 1	Sample 2	Sample 3	Non Sample 1	Non Sample 2	Non Sample 3
F	5	4	8	6	7	4	6
G + B♭	8	10	5	8	12	7	7
A	11	13	10	11	12	13	11
C	15	12	15	13	8	16	16
D	6	6	7	7	6	5	5
$\chi 2$		1.663636	3.182576	0.633333	6.157576	0.92197	0.558333
P-Value		**0.797312**	**0.52775**	**0.959289**	**0.187685**	**0.921388**	**0.967579**

The Star-Spangled Banner

Pitch Class	Original	Sample 1	Sample 2	Sample 3	Non Sample 1	Non Sample 2	Non Sample 3
C	23	31	25	26	31	19	31
D	11	13	9	8	15	10	11
E	22	19	24	23	22	24	19
F	8	3	9	6	4	13	5
G	20	18	15	18	9	18	16
A	6	5	8	6	6	8	11
B + F#	11	12	11	14	14	9	8
$\chi 2$		7.137912	2.761034	2.773123	13.10534	3.141864	27.52131
P-Value		**0.308281**	**0.838186**	**0.836736**	**0.041393**	**0.503018**	**0.12044**

I Don't Want To Miss A Thing

Pitch Class	Original	Sample 1	Sample 2	Sample 3	Non Sample 1	Non Sample 2	Non Sample 3
C	30	29	43	28	33	42	29
D	31	28	32	30	30	48	46
E	29	32	18	28	31	18	30
G	37	41	26	39	41	22	27
A	17	11	21	20	12	17	11
F + F# + B	8	11	12	7	5	5	9
$\chi 2$		4.30908	16.04945	0.962594	3.49821	145.2977	12.27123
P-Value		**0.505827**	**0.006704**	**0.965526**	**0.623659**	**0.000111**	**0.031254**

Wachet auf, ruft uns die Stimme

Pitch Class	Original	Sample 1	Sample 2	Sample 3	Non Sample 1	Non Sample 2	Non Sample 3
E♭	40	42	46	30	36	44	47
F	32	23	32	23	35	31	46
G	37	32	39	36	38	39	40
A♭	21	23	26	25	28	20	17
B♭	30	35	24	42	27	30	21
C	17	22	12	21	15	16	13
D	24	22	21	18	25	21	26
A + D♭	10	12	11	16	7	10	1
$\chi 2$		6.36799	5.344173	16.66136	4.518571	1.020801	20.26299
P-Value		**0.497497**	**0.61804**	**0.019715**	**0.71848**	**0.994486**	**0.005029**

Interval Frequency testing

The Star-Spangled Banner

Interval	Original	Sample 1	Sample 2	Sample 3	Non Sample 1	Non Sample 2	Non Sample 3
Unison	18	19	15	18	16	18	27
Second	52	55	50	48	59	48	44
Third	21	18	27	26	17	26	19
Fourth	9	8	8	8	8	8	10
$\chi 2$		0.768315	2.40232	1.60928	2.0375458	1.60928	6.032357
P value		0.85703	0.493203	0.657288	0.5646506	0.657288	0.110047

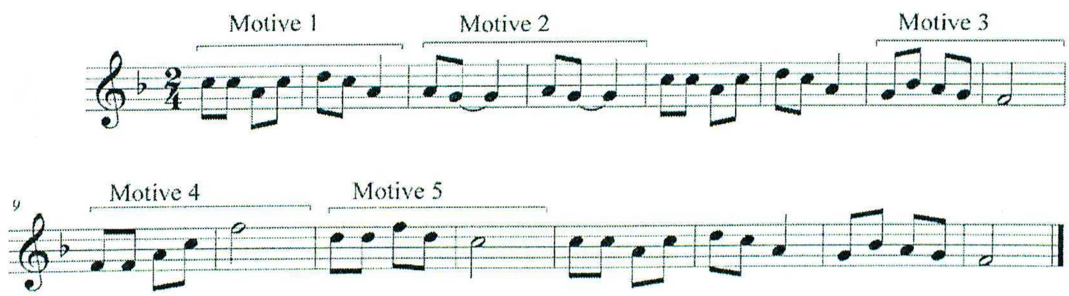
I Don't Want To Miss A Thing

Interval	Original	Sample 1	Sample 2	Sample 3	Non Sample 1	Non Sample 2	Non Sample 3
Unison	30	31	35	29	34	35	33
Second	58	54	60	59	51	71	60
Third	39	41	30	39	48	28	41
Fourth	24	25	26	24	18	17	17
$\chi 2$		0.453426	3.145889	0.050575	4.955084	8.891357	2.513196
P value		0.928998	0.369674	0.997021	0.1751158	0.030771	0.472911

Appendix D

Motive list

Camptown Races



The Star-Spangled Banner