

# Computers Composing Music: An Artistic Utilization of Hidden Markov Models for Music Composition

Lee Frankel-Goldwater, 2006

Advised by Dr. Chris Brown, Ph.D.

Department of Computer Science, University of Rochester

Natural systems are the source of inspiration for the human tendency to pursue creative endeavors. Music composition is a language for human expression and can therefore be utilized in conveying the expressive capabilities of other systems. Using a Hidden Markov Model (HMM) learning system, a computer can be taught to create music that is coherent and aesthetically sufficient, given the correct tools.

*"A Hidden Markov Model (HMM) is a statistical model where the system being modeled is assumed to be a Markov process [a finite, probabilistic system with state relationships] with unknown parameters, and the challenge is to determine the hidden parameters, from the observable parameters, based on this assumption. The extracted model parameters can then be used to perform further analysis, for example, for pattern recognition applications.*

*In a regular Markov model, the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. A hidden Markov model adds outputs: each state has a probability distribution over the possible output tokens. Therefore, looking at a sequence of tokens generated by an HMM does not directly indicate the sequence of states."*<sup>1</sup>

Such models have applications to speech recognition software, music transcription, and analysis of incomplete data because of this ability to approximate, as will be further described below. It is this capability of "fuzzy" pattern recognition that inspires the use of HMM's for this project.

The tools selected for this project include: twenty-two years of sun spot data as the natural system from which to creatively draw; a compositional framework for structure, pitch, dynamics, and rhythm to facilitate a human understanding of the system's expressiveness; the jMusic<sup>2</sup>, open source, music composition software; and an HMM learning system<sup>3</sup> with implementations of the Forward-Backward, Viterbi, and Baum-Welch algorithms. Details of these tools shall be discussed later. In composing a final piece of music, the attempt was made to impose as few creative restrictions on the system as possible. Through these tools, every aspect of the composition's generation can be repeated. In this way the robust analytical capabilities of the system are displayed via the piece and its generative procedures, thereby displaying an artificial intelligence's potential for music composition and perhaps larger creative projects.

## Artistic Motivations

The artistic bases of this project are as follows:

- Nature has discernable patterns within its manifestations
- Patterns are interpreted on a basis of past experience
- Music composition is driven by a need to express relationships

A clear and empirical observation is that natural systems exhibit certain repetitive and cyclic properties. Human beings interact with their environment on a constant basis, and one of the greatest skills we exhibit is the ability to recognize and learn from the patterns that occur within that environment. It is particularly interesting that we seem to be able to observe a pattern within nature and then apply the observations about that series of relationships to other, seemingly unrelated tasks.

Engineering, mathematics, psychology, and any science - martial arts, painting, music composition, and any artistic endeavor - appear to be exhibitions of the human capacity to interpret nature. Inquiries into the foundation of this ability are at the heart of some of the greatest pursuits of philosophical study. What is already apparent, however, is that humans utilize past experience to interpret newly recognized patterns and to build on their base of knowledge.

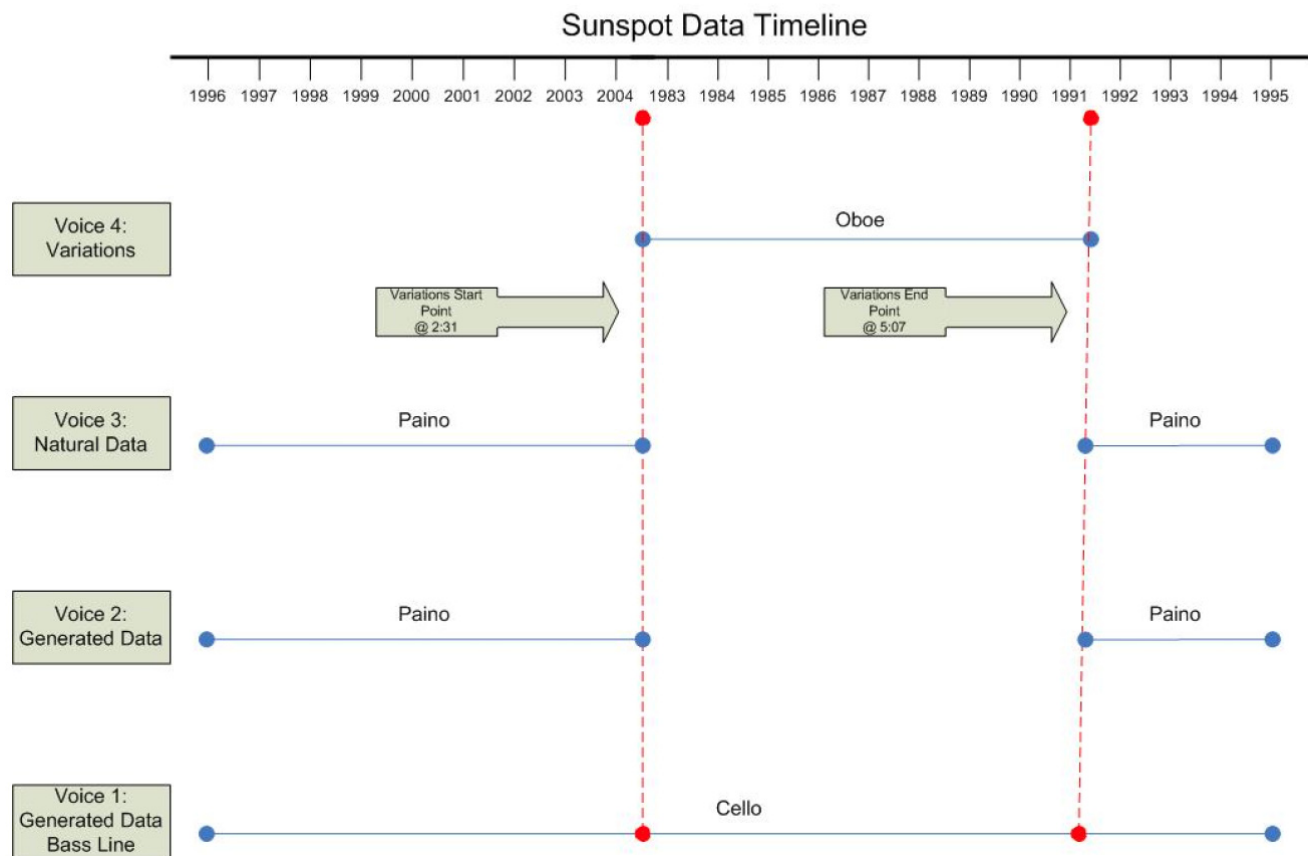
It is also interesting to observe the creative urge of the human condition. The need to express thoughts in some medium, whether it is through engineering or music composition, is a universal trait. One artistic assumption that was made during this project is that this human creative urge stems from a natural desire to express our interpretations of the patterns and relationships that we have recognized in nature. This does not appear to be an inappropriate leap in logic, because the fuel for creativity must come from nature in some capacity. Without learned concepts and patterns, we would have no basis for creation.

Therefore, it seems possible to imbue an artificial system with the tools necessary to execute a similar process of creation through the medium of music composition; utilizing the proper learning tools (Hidden Markov Models) and with a person to guide the creation and transmission of the product to listeners (jMusic and the composer).

## Tools

Two main software resources were used in the course of

# Overall Composition Structure



**Diagram A: Compositional Structure Flow Chart**

Represents the musical arrangement of the piece with respect to the sun spot data's year, the particular voice, and the instrument utilized to articulate that voice.

this project. The first is a Hidden Markov Model (HMM) software implementation of the Forward-Backward, Viterbi, and Baum-Welch algorithms. This is a somewhat ideal tool for the creation of HMM's that can be manipulated and reused to generate data from ambiguous observation sets. It can be easily installed on any UNIX system and generates output that can be easily stored as desired. Of the three algorithms mentioned, only the Baum-Welch implementation was utilized for this project, as other algorithms relate more to the internal workings of the HMM's and are therefore not of concern to the pure generation of artistic information.

There are two main executables of note from the package that were utilized in this project. The first is called 'esthmm,' which implements the Baum-Welch algorithm and is the HMM generator. It takes for input a number of states, a number of symbols, and an observation filename, which all correspond to the necessities for generating an HMM. It also takes for input an optional random number seed. This guarantees generation repeatability. Utilizing the Baum-Welch algorithm, an HMM is created that is an attempt by the algorithm to find the overall pattern of the data fed to it through the determination of weights in a state machine of a predetermined size (a finite series of nodes and fully connected paths) based on Markov probabilistic mathematics. This represents the foundation of the artistic process for the AI: getting data from nature and finding patterns.

The second executable is called 'genseq,' which generates

sequences from an HMM and represents the second step in the artistic process. By generating data from the HMM, an attempt is made to replicate the data that was fed to it. Depending on the nature of the data and the specifics of the HMM initialization, various degrees of accuracy are achieved. Generally, for sufficiently large datasets, true accuracy is not what is sought, due to memory constraints; rather, approximation is sought, and therefore allows for the variability in the creative process described in the Artistic Motivations section. This executable also allows a seed value in its input string to facilitate repeatability.

The second primary software tool is jMusic, an open source, java based, and algorithmic composition software package. As the author accurately describes, "jMusic is a project designed to provide composers and software developers with a library of compositional and audio processing tools. It provides a solid framework for computer-assisted composition in Java, and is also used for generative music, instrument building, interactive performance, and music analysis."<sup>4</sup> This is an extremely useful set of tools, more robust than necessary for this project, yet still intuitive to use and having every resource required.

The structure of jMusic is built around a note object that gets added to successively larger collections of notes within other object structures. These note objects represent a section of memory in the computer that corresponds to all of the features of a single note or rest. These sections of memory are then linked together with other similar sections to make

larger objects in memory that correspond to other musical structures, such as bars, voices, and whole compositions. The primary use for jMusic in this project was to manipulate the data compositionally once processing with the HMM tools was complete.

### Compositional Structure

The structure of this piece was designed to make evident to the listener the pattern recognizing and the expressive capabilities of the HMM system. The flowchart in Diagram A is a visual example of the structure. It is a four voice piece built around the data generated by the HMM system overlaid with the original observation data. For clarification, a musical voice is defined here as a single, horizontal line of music that bares some intended relationship to the notes that precede or proceed from any given note in that line. This data is a 22 year collection of sunspot data from 1983 to 2004 gathered from the World Sunspot Index.<sup>5</sup> Sunspots are a series of dark blotches that occur on the surface of the sun and are studied because of the magnetic phenomena surrounding them. Sunspot data was chosen because the spots seem to have a cyclic quality across a 22 year period. Note that the piece was rotated so that it begins in 1996 for aesthetic reasons.

Beginning in 1992, measurement techniques became more advanced, and so there is an increased amount of data starting

in this year. Voices 2 and 3 utilize an extra two columns of data provided in the years 1992 through 2004 for the purpose of varied rhythms and dynamics. The pitch material for all of the voices of the piece comes from the first column of data, which runs from 1983 through 2004 (See Diagram B below).

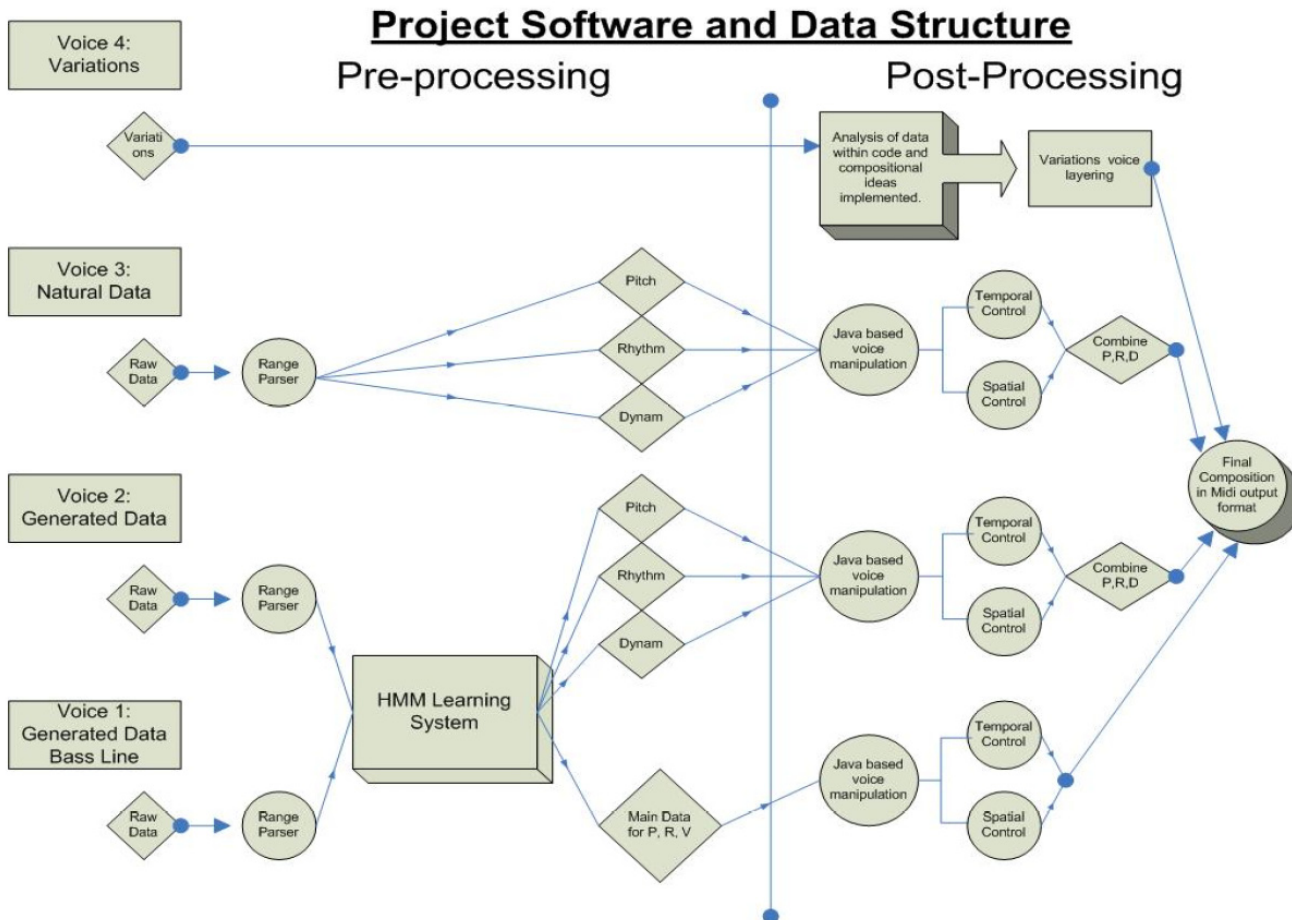
Voice 3 utilizes the original sunspot data as a source for pitch, rhythm, and dynamic material; voice 2 utilizes HMM generated data to determine the pitch, rhythm, and dynamics; the bass line, voice 1 uses HMM data for the pitch material, but the rhythm and dynamics are static. This choice was made so that the listener has some sense of meter, because the piece has specific meter. The variations voice, voice 4, is built over the section of the piece where the material for voices 2 and 3 is incomplete because of the year. While this voice is built off of the same pitch material as the line below it, it utilizes a more staggered structure that is apparent to the listener.

The tempo for the final form of the piece is 5000 beats per minute to facilitate a reasonable overall length across 22 years of data. The scale that was chosen to filter the pitch data through for the final form is pentatonic. It is believed that this scale allows the listener to further appreciate the aesthetic qualities of the piece. The chromatic version, however, facilitates a better sense of the cyclic motion naturally found within the data.

Most of the filtering was done linearly. A change in the data by one notch would affect the pitch, rhythm, or dynamics a

**Diagram B: Software Design Flow Chart**

The design procedure of the project followed this flow chart. There is an attempt in the creation of this chart to be general and robust enough to facilitate an understanding of the experimental procedure of the project. The line separating the Pre-Processing and Post-Processing sections represent the point at which "compositional" data has been collected from the HMM system and is then brought into the jMusic music processing package. The lone letters P, R, and D correspond to musical Pitch, Rhythm, and Dynamics (note volume) respectively.





single notch where appropriate. All layers of data were kept lined up by day so that, for example, the pitch data in voice 4 for April 3<sup>rd</sup>, 1995, will temporally line up with the same pitch data in the base line.

In these ways delineated above, the hypothesis is that the listener should be able to hear patterns found by the HMM system both melodically and harmonically.

### Program Structure

The program structure is well defined by diagram B shown above. A set of raw data from the sun spots was first parsed into the proper ranges for use in the music. Sunspot data had a range within a given column as high as 246 through 0. This is too great to be applied to the Western pitch system. The 'Range Parser' tool specified in the diagram changes these ranges in a balanced way to protect the cyclic structure of the data, while allowing for a reasonable range to work with compositionally.

Relative to its appropriate voice, the parsed data was then fed through the HMM learning system and saved; then, generation sequences were made based on the HMM's interpretation of the natural data. There was quite a rigorous stage of data manipulation in this section of the project that could have been more easily facilitated by an entirely Java based system.

The proper columns of data were then formatted for use in Java and passed to the proper voice controllers as specified in the previous section. Each voice was generated independently of the others. Temporal and spatial relationships were checked so that the patterns within data would remain intact and the artistic intent maintained. In this stage of the post-processing, the variations voice was compiled and collected with the other voices in preparation for exporting.

jMusic facilitates a convenient conversion of musical data into standard Midi format. Thus a completed piece of music is created. It is important to note for artistic and scientific purposes that all aspects of this project, including all generations and combinations, are repeatable using the methods specified in this paper and the seed value 0605.

### Conclusions

Most musically-related work with HMM's is in the areas of transcription and recognition, rather than composition. As this project shows, the potential for an HMM composition system is quite viable and applications for composition are available for expansion. The composer can choose any level of utility for the system, from thorough piece creation, as with this project, to simply using it as a source for compositional material. Future work in this area could be done on building more codependence between the voices and further elaboration on the variations section.

It is clear that an AI system has the potential to be highly expressive, given a sufficient base of experience and an advanced system with which to communicate with human beings.

### Acknowledgements

To my advisor Chris Brown for his leadership and valuable lessons in AI and group work. To Josh Mailman for his compositional advise. To Mark Bocco for allowing me to use the Electrical Engineering music facilities. To the various other University of Rochester and Internet resources that gave me assistance.

### References

1. Wikipedia, Hidden Markov Model entry <[http://en.wikipedia.org/wiki/Hidden\\_Markov\\_Model](http://en.wikipedia.org/wiki/Hidden_Markov_Model)>
2. Sorensen A. and Brown A. jMusic: Music Composition in Java. Source Code: <<http://jmusic.ci.qut.edu.au/>>
3. Kanungo T. Hidden Markov Model (HMM) Software: Implementation of Forward-Backward, Viterbi, and Baum-Welch algorithms. Source Code: <<http://www.cfar.umd.edu/~kanungo/software/software.html>>
4. jMusic Source and Documentation, <<http://jmusic.ci.qut.edu.au/>>
5. SIDC World Sunspot Index, <<http://sidc.oma.be/index.php3>>
6. Ames C. 1990. Artificial Intelligence and Musical Composition: How Can Intelligent Computers Help Compose Music? *The Age of Intelligent Machines*, Kurzweil R. MIT Press.
7. Biles, John A. (1994). GenJam: A Genetic Algorithm for Generating Jazz Solos. In *Proceedings of the 1994 International Computer Music Conference*, pp. 131-137. San Francisco: ICMA.
8. Biles J.A. 2002. GenJam in Transition: from Genetic Jammer to Generative Jammer. *Generative Art*, Rochester Institute of Technology.
9. Chong Y. 1996. Computer Generated Music Composition. Submitted to the MIT Department of Electrical Engineering and Computer Science for Graduate Thesis.
10. Dobrian C. 1993. Music and Artificial Intelligence. <<http://music.arts.uci.edu/dobrian/CD.music.ai.htm>>
11. McCormack J. 1996. Grammar Based Music Composition. *Complex Systems 96: From Local Interactions to Global Phenomena*, R. Stocker et. al., eds, ISO Press, Amsterdam.
12. Miranda E. R. 2002. Sounds of Artificial Life. Proceedings of the ACM SIGCHI International Conference - Creativity and Cognition 4, October 2002, Loughborough University ACM.
13. Rabiner L. and Juang B. 1986. An introduction to hidden Markov models. *IEEE ASSP Magazine*, pages 4-16.
14. Roads C. 1980a. Artificial Intelligence and Music. *Computer Music Journal*, Vol. 4, No. 2. CA: People's Computer Company: 13-25.
15. Thom B. 2000. BoB: an Interactive Improvisational Music Companion. Doctoral Thesis, Carnegie Mellon University.