Abstract

Anitoo is a group of Max patches that extracts pitch, duration and intensity values from a MIDI note sequence, performs some transformations and provides statistical information of those values. In main sum is to allow for testing hypotheses on musical structure in a simple and direct way, so it can be described as a first step toward the development of tools to help define more comprehensive musical models. It is part of a framework in progress which analyzes musical structure resulting from the interaction of sound, perceptual and social systems. Some theoretical issues are discussed and a brief description of the patches is given.

1. Introduction

This paper reports some results of a framework in progress that approaches musical structure as an interaction of three interdependent systems: sound, perceptual and social. These are dynamically interlinked and their interaction provides a transformation of the state of each system [Keller and Silva, 1995]. The sound system is defined as a time series - variations of a chosen variable as a function of time - that is, the unfolding of a sequence of acoustical signals [Foon et al., 1990]. Information provided by the sound system is processed by a perceptual system whose dynamics are described by its states - or representations - which are modified by processes - or operations. Global constraints - established by the social system - set the range of possible perceptual states and thus the probability that the whole musical process will take place in a specific way.

In previous works we have defined musical information as a measure of the elements' range and rate of transformation in the sound system [Keller and Silva, 1995]. Periodicity and entropy are the two modeling forces that act on it. Three levels of observation are defined for this system: sound, syntax and morphological [Todd and Ley, 1991; Syman, 1995; De Foli et al., 1991; Geорguescu and Geорguescu, 1990].

The task of building a theoretical structure that could account for musical phenomena either in traditional or contemporary forms of organization in western music brought up the need of some quantification tools [Bootsz, 1992; Demster and Brown, 1990]. When the concept of dynamical system is applied to music, the characteristics of the system should be inferred from musical data. Before such system has been done, any observation about the behavior of a musical system or any prediction on the results of a specific transformation might be too sufficiently supported [Artallo, 1994, Keller and Silva, 1995].

Max, a graphical programming language for Macintosh, appeared as the best environment to develop some tools for this purpose.

Here we will be discussing part of the theoretical background behind Anitoo - although we refer the reader to the paper by Keller and Silva [1995] for a complete description of the framework. A short explanation on Anitoo is included. Finally, some implications and perspectives of development are discussed.

2. Musical Process

When music is considered as a time series the transformation of a variable in relation to time can be observed. This allows us to look for regularities in the sound, syntax or morphology of a musical signal. If a musical piece is understood as a system in equilibrium, its structure would be a time-invariant representation embodying all states. Their time history represents the dynamic process of the system. A plot of this process in a three-dimensional graph shows the evolution of the system as a trajectory in the phase-space sustained by x, y, z axes. This space equals the state of the system [Foon et al., 1990].

From this perspective a process can be defined as a finite group of relationships in transformation that establish a perceptual unit [Temney and Polansky, 1980]. Variables in big scale musical processes approximate a 1/2 correlation. That is halfway between unpredictable dynamics, as in white noise, and very high correlation not providing enough information to catch a listener's attention [Schumacker and Gilden, 1993].

The stability of a musical system is proportional to its entropy. Almost completely entropic systems are the most stable ones. So strong processes - highly energetic - are needed to disrupt their equilibrium - as it happens in stochastic music or random noise where foreign events can be introduced without great modification of the system's behavior. On the other hand a system with high periodicity will suffer a
strong impact from small perturbations [Keller and Silva, 1995].

3. Information filters
Correlation has been defined from two different perspectives: statistics and signal processing. From the statistical point of view, the distance of the values of X to its mean is calculated to find the standard deviation of X. Here we are dealing with the probability of occurrence for each value of X. A comparison of the standard deviation of X and Y provides a measure of their correlation [Tabachnick and Fidell, 1989]. Cross-correlation can be thought of as a similarity test between two waveforms and as a way to find out their common frequency components [Ramirez, 1985].

Using these tools from statistics and signal processing, we can measure the correlation of musical structures at the level of their micro and macro-structure. This way we can establish the behavior of a variable and get a measure of its information content - defined by the range and rate of transformation of its elements [Knopoff and Hutchinson, 1981; Lutfi, 1992; Green, 1988].

Once this analysis has been done, musical data can easily be organized by processes that act directly on the range and rate of transformation of each variable. Thus, we can define a musical structure by using filters to control musical information.

4. Parameters
The development of extra-musical systems - mathematical, physical, biological - which could accomplish any valid musical structure - new or already existing - is bound up by the difficulty of defining a direct correlation between extra-musical and musical variables [Wustner, 1995]. The difference between physical variation of parameters and the corresponding perceptual outcome should be taken into account [Astello, 1994; De Poli et al., 1991; Masiero and Cowan, 1993].

Generally values from a control system are directly matched to variation of a musical variable within a predefined absolute range [Hullick, 1992; Gogins, 1990]. This procedure does not consider the influence of context and interaction observed in musical systems. In fact, much experimental work has shown that pitch, intensity, duration and timbre are continually interacting [Melara and Marks, 1990]. Furthermore, micro-variations in timbre in each sound event serve as important clues to identify instruments, and it is likely that they help construct a mental image of musical structure [Clynes, 1995; Troncoli and Keller, 1995].

5. Anitoo - Analysis Tools
Anitoo's patches are divided in three functional groups: the sequence itself where notes are recorded and played back, the analysis-playback group that parses pitch, velocity and duration displaying each parameter on a graph (table), and the statistics group that calculates mean and standard deviation for the values provided by the notes. The difference between the two groups is that the second is an irreversible process not allowing to recover the original data [Bennet, 1988]. Therefore, transformations of musical parameters are handled by the first group and analytical computations by the second.

Once parameters have been parsed, the basic framework can be enhanced by introducing new mechanisms of organization and transformation of musical data.

5.1 Materials
Anitoo was developed on a PowerPC 7100A/4, using Max version 2.52. The external object LitterState by Peter Castine was used in some patches. The midi keyboard was a Roland JD 800. All these resources belong to the Laboratory of Electroacoustic Music of the University of Bari.

5.2 Pre-processing
Anitoo does not make any previous assumption regarding the size or range of the midi sequence. Therefore some preliminary calculations have to be done before all parameters can be parsed and displayed correctly. This is why it does not work in real time.

To calculate duration of notes and to display them on a 127 scale (as the table object allows), we first have to find out what the longest value of the sequence is, set it as equal to 127 and calculate all the other values as a fraction of 127. Once this has been done, all values can be directly sent to be able to be displayed.

We noticed some difficulty to get values of durations between note-on and note-off and total duration for each midi note - from note-on to the next note-on - when reading SMF with 3do. Thus, we set up the task of developing a timing mechanism inside Max, to calculate these durations from note-on and note-off messages. Nevertheless we should state that its precision is limited to 5 ms., which is the shortest interval measured by the clocker object.

5.3 Pitch, Intensity, Duration and Articulation
Pitch values are sent to a table to be shown on a graph. Thus, a melodic context equivalent to the succession of pitch values is displayed on the screen.
This contour information serves as a quick visual reference to establish overall comparisons among melodies. Some procedure was implemented for intensity and duration, although a conceptual differentiation between duration and articulation was established. Time intervals from note-on to note-off are used as a measure of note duration, and articulation is given by note-on note-off intervals. This means that a legato sequence and a staccato sequence can be represented by the same duration values.

5.4 Statistics

The simplest way to observe the distribution of values in a sequence is by means of the Histo object. It keeps track of the number of times a value is repeated. When we display its output on a table, we have a rough view of how the values are distributed and which are the most frequent values for the parameter observed.

LettinStats object [Peter Castine] provides a measure of the mean and standard deviation of all values given. This is an alternative way to compare the information content of two sequences, centering on overall parameters' variation rather than focusing on each individual value.

6. Summary

Amto is a group of Max patches that extracts pitch, duration and intensity values from a monophonic midi sequence, performs some transformations and provides statistical information of these values. One main aim while developing it was to allow for testing hypotheses on musical structure in a simple and direct way, so it can be described as a first step toward the development of tools to help defining musical models.

This work forms part of a broader theoretical framework that can be described as a system that brings together physical, perceptual and social information to musical structure. The sound system is divided in three levels: sound, syntax and morphology. Entropy and periodicity are used to define the stability and behavior of the sound system. Fusion and parsing, as high-level processes, control the structure of perceptual mechanisms. Constraints defined by the social system influence the musical process for a specific environment.

Many limitations can be pointed out regarding the use of midi protocol and automatic musical production; therefore, the tools provided serve only as a starting point toward more investigations. The analysis of real music is suggested as a strong methodology to establish comparisons with algorithmic production and to develop comprehensive and precise musical models.

As works in interpretation analysis have shown, micro-structure plays an important role in musical organization [Clayes, 1995; Todd, 1992]. This micro level should be combined with observation of the dynamic of events and macro-structure to provide a complete description of musical phenomena.

Acknowledgments

Thanks to Dr. Karlheinz Essl for sending LettinStats. Thanks to Prof. Cecídio Silva, Ana Lucía and Nhoit for invaluable help and food. CNPq-UnB provided partial financial support.

References


