The Melody Triangle - Pattern and Predictability in Music

Henrik Ekeus
Queen Mary University of London
Media and Arts Technology
School of Electronic Engineering and Computer Science
hekeus@eecs.qmul.ac.uk

Samer Abdallah
Queen Mary University of London
Center for Digital Music
School of Electronic Engineering and Computer Science
samer.abdallah@eecs.qmul.ac.uk

Mark D. Plumbley
Queen Mary University of London
Center for Digital Music
School of Electronic Engineering and Computer Science
mark.plumbley@eecs.qmul.ac.uk

Peter W. McOwan
Queen Mary University of London
Computer Vision Group
School of Electronic Engineering and Computer Science
Peter.McOwan@eecs.qmul.ac.uk

ABSTRACT

The Melody Triangle is an exploratory interface for the discovery of melodic content, where the input—positions within a triangle—directly map to information theoretic measures associated with the output. The measures are the entropy rate, redundancy and *predictive information rate*[1] of the random process used to generate the sequence of notes. These are all related to the *predictability* of the the sequence and as such address the notions of expectation and surprise in the perception of music.

We describe some of the relevant ideas from information dynamics, how the Melody Triangle is defined in terms of these, and describe two physical incarnations of the Melody Triangle. The first is a multi-user installation where collaboration in a performative setting provides a playful yet informative way to explore expectation and surprise in music. The second is a screen based interface where the Melody Triangle becomes a compositional tool for the generation of musical textures; the user's control at the abstract level of randomness and predictability. Finally we outline a pilot study where the screen-based interface was used under experimental conditions to determine how the three measures of predictive information rate, entropy and redundancy might relate to musical preference.

Keywords

Information dynamics, Markov chains, Collaborative performance, Aleatoric composition, Information theory

1. INFORMATION DYNAMICS

Music involves patterns in time. When listening to music we continually build and re-evaluate expectations of what is to come next. Composers commonly, consciously or not, play with this process by setting up expectations which may, or may not be fulfilled, manipulating the expectations of the listener and inducing surprise or not as the music progresses [2, 4]. Central to this is the idea that music is not a static object, presented as a whole, but as a phenomenon that 'unfolds' and is experienced *in time*.

Information dynamics[1] considers several different kinds of predictability in musical patterns, how these might be quantified using the tools of information theory, and how they shape or affect the listening experience. Central to this is the idea that listeners maintain a dynamically evolving statistical model that enables them to make predictions about how a piece of music will continue. They do this using both the immediate context of the piece as well as using previous musical experience. As the music unfolds, listeners continually revise their model; in other words, they revise their own, subjective probabilistic belief state. These changes in probabilistic beliefs can be associated with quantities of information; these are the focus of information dynamics.

2. THE MELODY TRIANGLE

The Melody Triangle enables the discovery of melodic content matching a set of information theoretic criteria. Positions within the triangle correspond with pairs of values of entropy rate and redundancy. The relationship with the predictive information rate is not explicitly controlled as this would require a three-dimensional interface, but an implicit relationship emerges, which is described in section 2.2. The physical interface to the Triangle has so far been realised in two forms: as an interactive installation and as a screen based interface.

Given the information coordinates corresponding to a point in the triangle, we select from a pre-built library of random processes, choosing one whose entropy rate and redundancy match the desired values. The implementations discussed in this paper use first order Markov chains as the content generator, since it is easy to compute the theoretically exact values of entropy rate, redundancy and predictive information rate given the transition matrix of the Markov

NIME'12, May 21 – 23, 2012, University of Michigan, Ann Arbor.

chain. However, in principle, any generative system could be used to create the library of sequences, given an appropriate probabilistic listener model supporting the estimation of entropy rate and redundancy.

The Markov chain based implementation generates streams of symbols in the abstract; the alphabet of symbols is then mapped to a set of distinct sounds, such as pitched notes in a scale or a set of percussive sounds. Further by layering these streams intricate musical textures can be created. The selection of notes or sounds is arbitrary, as long as they are all distinguishable. Indeed, the symbols could be mapped to even non sonic outputs such as visible shapes, colours, or movements.

Any sequence of symbols can be analysed and information theoretic measures estimated from it. The novelty of the Melody Triangle lies in that we reverse this mapping: given desired values for these measures, as determined from the user interface, we return a stream of symbols with the desired properties. In the next section we describe the three information theoretic measures that we use.

2.1 Sequential information measures

The *entropy rate* of a random process is a basic measure of its randomness or unpredictablity. Consider the viewpoint of an observer at a certain time, and split the sequence into an infinite *past*, as single symbol in the *present*, and the infinite *future*. The entropy rate is a conditional entropy; informally:

$$\text{EntropyRate} = H(\text{Present}|\text{Past}), \qquad (1)$$

that is, it represents our average uncertainty about the present symbol *given* that we have observed everything before it. Processes with zero entropy rate can be predicted perfectly given enough of the preceeding context.

The *redundancy* of the a process, in the sense we are using the term here, is a measure of how much the predictability of the process depends on knowing the preceeding context. It is the difference between the entropy of a single element of the sequence in isolation (imagine chosing a note from a musical score at random with your eyes closed and then trying to guess the note) and its entropy after taking into account the preceeding context:

$$\text{Redundancy} = H(\text{Present}) - H(\text{Present}|\text{Past}). \qquad (2)$$

If the previous symbols reduce our uncertainty about present symbol a great deal, then the redundancy is high. For example, if we know that a sequence consists of a repeating cycle such as $\ldots b, c, d, a, b, c, d, a \ldots$, but we don't know which was the first symbol, then the redundancy is high, as $H(\text{Present})$ is high (because we have no idea about the present symbol in isolation, but $H(\text{Present}|\text{Past})$ is zero, because knowing the previous symbol immediately tells us what the present symbol is.

The *predictive information rate* (PIR) brings in our uncertainty about the future. It is a measure of how much each symbol reduces our uncertainty about the future as it is observed, *given* that we have observed the past:

$$\text{PIR} = H(\text{Future}|\text{Past}) - H(\text{Future}|\text{Present}, \text{Past}). \qquad (3)$$

It is a measure of the *new* information in each symbol. Notice that if the past completely determines both the present and the future (as in the cyclic pattern above) the PIR is zero, since the present symbol brings no new information. However, if the symbols in a sequence are generated completely independently, e.g. by rolling a die for each one, then again, the present symbol provides no information about the future and the PIR is zero. However, there do exist processes that have high predictive information rates

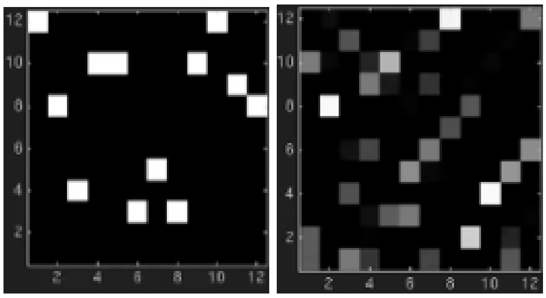

Figure 1: Two transition matrixes. The shade of white represents the probabilities of transition from one symbol to the next (black=0, white=1). The current symbol is along the bottom, and in this case there are twelve possibilities (mapped to a chromatic scale). The left hand matrix has no uncertainty; it represents a periodic pattern. The right hand matrix contains unpredictability but nonetheless is not completely without perceivable structure, it is of a higher entropy rate.

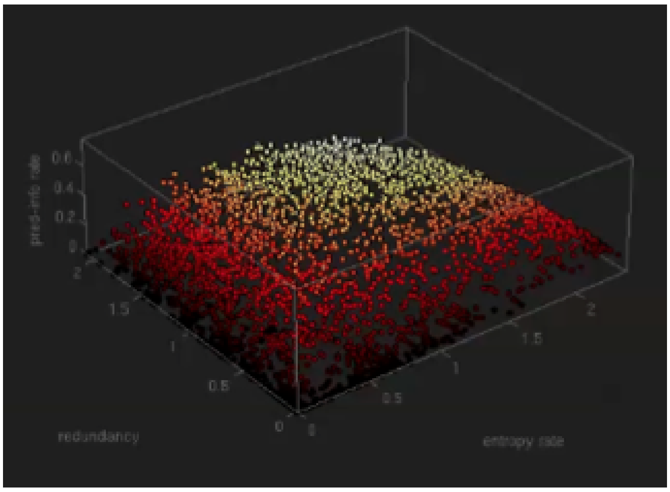

Figure 2: The population of transition matrixes distributed along three axes of redundancy, entropy rate and predictive information rate. Note how the distribution makes a curved triangle-like plane floating in 3d space.

as compared with their entropy rates: within the class of Markov chains, these are neither the periodic nor the sequentially uncorrellated ones. Rather they tend to yield sequences that have certain recognisable patterns or motifs, but which occur at irregular times. A certain symbol might tell us about which one of the characteristic patterns will appear next. Each symbol tell a us little bit about the future; in order to make good predictions, the listener must continually pay attention, building up expectations on the basis of each new observation.

2.2 Populating the triangle

Before the Melody Triangle can used, it has to be 'populated' with possible parameter values for the melody generators. These are then plotted in a 3d statistical space of redundancy, entropy rate and predictive information rate. In our case we generated thousands of transition matrixes, representing first-order Markov chains, by a random sampling method. In figure 2 we see a representation of how these matrixes are distributed in the 3d statistical space;

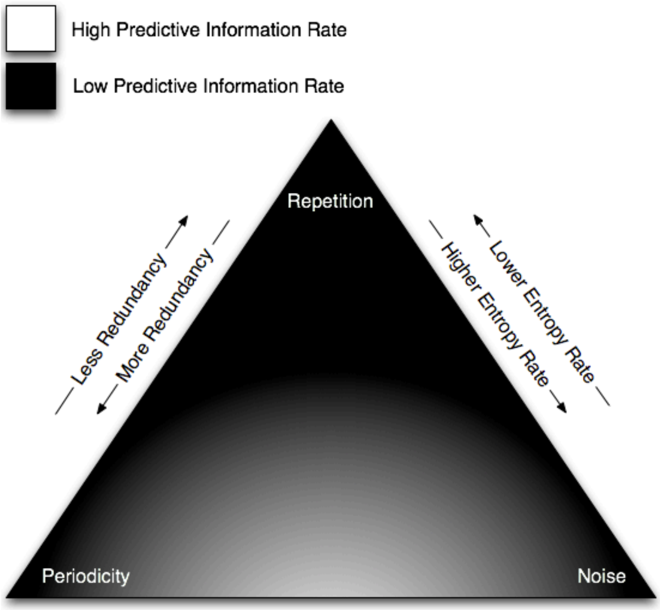

Figure 3: The Melody Triangle

each one of these points corresponds to a transition matrix.

When we look at the distribution of transition matrixes plotted in this space, we see that it forms an arch shape that is fairly thin. It thus becomes a reasonable approximation to pretend that it is just a sheet in two dimensions; and so we stretch out this curved arc into a flat triangle. It is this triangular sheet that is our 'Melody Triangle' and forms the interface by which the system is controlled.

When the Melody Triangle is used, regardless of whether it is as a screen based system, or as an interactive installation, it involves a mapping to this statistical space. When the user, through the interface, selects a position within the triangle, the corresponding transition matrix is returned. Figure 3 shows how the triangle maps to different measures of redundancy, entropy rate and predictive information rate.

Each corner corresponds to three different extremes of predictability and unpredictability, which could be loosely characterised as 'periodicity', 'noise' and 'repetition'. Melodies from the 'noise' corner have no discernible pattern; they have high entropy rate, low predictive information rate and low redundancy. These melodies are essentially totally random. A melody along the 'periodicity' to 'repetition' edge are all deterministic loops that get shorter as we approach the 'repetition' corner, until it becomes just one repeating note. It is the areas in between the extreems that provide the more 'interesting' melodies. That is, those that have some level of unpredictability, but are not completely random. Or, conversely, that are predictable, but not entirely so. This triangular space allows for an intuitive exploration of expectation and surprise in temporal sequences based on a simple model of how one might guess the next event given the previous one.

3. USER INTERFACES

Any number of interfaces could be developed for the Melody Triangle[1]. We have developed two; a standard screen based interface where a user moves tokens with a mouse in and around a triangle on screen, and a multi-user interactive installation where a Kinect[2] camera tracks individuals in a space and maps their positions in the space to the triangle.

[1]The Melody Triangle was developed in Prolog and MatLab. It can be controlled with OpenSoundControl messages, and thus is independent of any specific interface implementation.

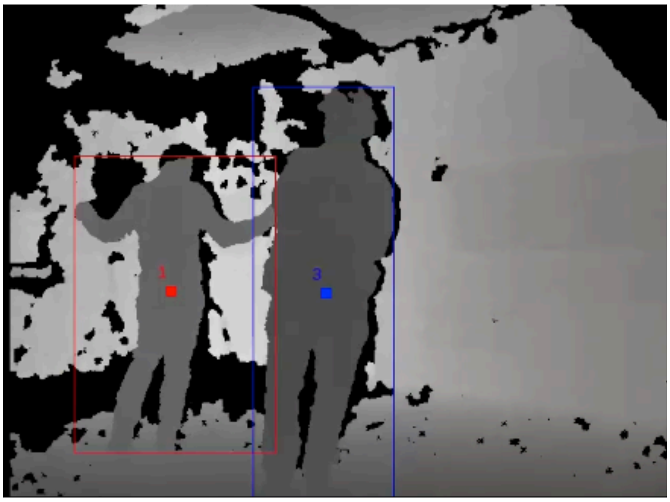

Figure 4: The depth map as seen by the Kinect, and the bounding box outlines the blobs detected by OpenNI.

3.1 The Multi-User Installation

As a Kinect camera overlooks a space, its range naturally forms a triangle. As visitors/users comes into the range of the camera, they start generating a melody, the statistical properties of this melody determined by the mapping of physical space to statistical space as discussed above. Thus by exploring the physical space the participant changes the predictability of the generated melodic content. When multiple people are in the space they can cooperate to create interweaving melodies, forming intricate polyphonic textures.

The streams of symbols are mapped to MIDI and then played with software instruments in Logic. The tracking system was capable of detecting gestures, and these were mapped to different musical effects such as tempo changes, periodicity changes (going to the off-beat), instrument/register changes and volume (see Figure 1).

3.1.1 Tracking and Control

Tracking and control was done using the OpenNI libraries' API[3] and high level middle-ware for tracking with Kinect. This provided reliable blob tracking of humanoid forms in 2d space. By triangulating this to the Kinect's depth map it became possible to get reliable coordinate of visitors' positions in the space.

By detecting the bounding box of the 2d blobs of individuals in the space, and then normalising these based on the distance of the depth map it became possible to work out if an individual had an arm stretched out or if they were crouching. With this it was possible to define a series of gestures for controlling the system without the use of any controllers(see table 1). Thus for instance by sticking out one's left arm quickly, the melody doubles in tempo. By pulling one's left arm in at the same time as sticking the right arm out the melody would shift onto the offbeat. Sending out both arms would change the instrument being 'played'.

3.1.2 Observations

[2]http://www.xbox.com/en-GB/Kinect
[3]http://OpenNi.org/

Table 1: Gestures and their resulting effect

left arm	right arm	meaning
out	static	double tempo
in	static	halve tempo
static	out	triple tempo
static	in	one-third tempo
out	in	shift to off-beat
out	out	change instrument
in	in	reset tempo

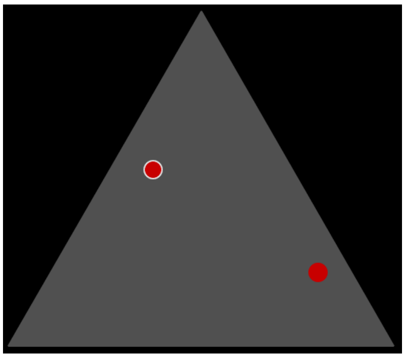

Figure 5: Screen shot of the screen based interface for the Melody Triangle

Although visitors would need an initial bit of training they would then quickly be able to collaboratively design musical textures. For example, one person could lay down a predictable repeating bass line by keeping themselves to the periodicity/repetition side of the room, while a companion can generate a freer melodic line by being nearer the 'noise' part of the space.

The collaborative nature of this installation is an area that merits attention. By not having one user be able to control the whole narrative, the participants would communicate verbally and direct each other in the goals of learning to use the system and finding interesting musical textures. This collaboration added an element of playfulness and enjoyment that was clearly apparent.

As an artefact this installation is an exploratory prototype and occupies an ambiguous role in terms of purpose; it is in a nebulous middle ground between instrument, art installation and technical demonstration. It is clear however, that as a vehicle for communicating ideas related to the expectation, pattern and predictability in music to the public, it is very effective.

3.2 The Screen Based Interface

The Melody Triangle can also be explored with a standard keyboard and mouse interface. A triangle is drawn on the screen, screen space thus mapped to the statistical space of the Melody Triangle. A number of round tokens, each representing a melody can be dragged in and around the triangle. When a token is dragged into the triangle, the system will start generating the sequence of notes with statistical properties that correspond to its position in the triangle.

Additionally there are a number of keyboard controls. These include controls for changing the overall tempo, for enabling and disabling individual voices, changing registers, going to off-beats and changing the speed of individual voices. The system gives visual feedback to indicate when a token has locked on to a new melody, and contains a buffer zone for allowing tokens to be pushed right to the edges of the triangle without falling out.

In this mode, the Melody Triangle can be used as a kind of composition assistant for the generation of interesting musical textures and melodies. However unlike other computer aided composition tools or programming environments, here the composer engages with music on the high and abstract level of expectation, randomness and predictability.

4. MUSICAL PREFERENCE AND INFORMATION DYNAMICS STUDY

We carried out a preliminary study that sought to identify any correlation between aesthetic preference and the information theoretical measures of the Melody Triangle. In this study participants were asked to use the screen based interface but it was simplified so that all they could do was move tokens around. To help discount visual biases, the axes of the triangle would be randomly rearranged for each participant.

The study was divided in to two parts, the first investigated musical preference with respect to single melodies at different tempos. In the second part of the study, a background melody is playing and the participants are asked to find a second melody that 'works well' with the background melody. For each participant this was done four times, each with a different background melody from four different areas of the Melody Triangle. For all parts of the study the participants were asked to 'mark', by pressing the space bar, whenever they liked what they were hearing.

After the study the participants were surveyed with the Goldsmiths Musical Sophistication Index[3] to elicit their prior musical experience.

4.1 Results

Note to the reviewers -

Due to time constraints we have been unable complete the analysis of the data captured during the study. The results would be outlined here and will include scatter plots of where in the triangle the subjects 'marked' and a heat map indicating where they tended to linger. These will be ready for final version if the paper is accepted.

4.1.1 Participant Feedback

After the experiments the participants were asked to comment on what they thought was happening when they moved the token around the triangle. They all correctly identified the repetitiveness of the melodies from the low redundancy, low entropy and low information rate area of the triangle, that is the 'repetition' corner (see fig 3). They all identified the randomness of the high entropy rate corner of the triangle. Similarly most participants identified the 'looping' inherent in the periodic melodies along the 'repetition' to 'periodicity' edge.

However the descriptions of the intermediate areas and lower edge, that is areas with greater predictive information rate, varied. Some felt that those areas were quite random and they couldn't really distinguish these from the 'noise' area of the triangle, others found that these were more 'interesting' and 'melodic' and reported enjoying dwelling there. Although there seemed to be a positive correlation between these positive ascriptions and the musical sophistication index, the sample size of this pilot study was too small to claim statistical significance.

5. FURTHER WORK

The Melody Triangle has so far only been used with first-order Markov chains for generating content. This mean that the melodies generated don't have any long term structure or form and hence don't seem to 'go anywhere'. As such the system in its current form is better suited to creating textures and short phrases as oppose to composing overarching musical structures.

We are currently investigating how higher-order Markov models can be mapped to information theoretic measures and adapting the Melody Triangle to those models. This would generate higher level patterns and provide more long-term structures. Further more sophisticated listener models[5][6] could be used for computing information measures for more conventional or ecologically valid music.

As it stands, the streams of symbols generated are only mapped to note values. However they could just as well be applied to any other musical property, such as intervals, chords, dynamics, timbres, structures and key changes. The possibilities for the Melody Triangle to be compositional guide in these other domains remains to be investigated.

The Melody Triangle in its current form however forms an ideal tool for investigations into musical preference and their relationship to the information dynamics models, and as such more detailed studies under wider experimental conditions and with more participants will be carried out.

6. ACKNOWLEDGMENTS

This work is supported by EPSRC Doctoral Training Centre EP/G03723X/1 (HE), GR/S82213/01 and EP/E045235/1(SA), an EPSRC Leadership Fellowship, EP/G007144/1 (MDP) and EPSRC IDyOM2 EP/H013059/1. Thanks to Louie McCallum and Davie Smith from QMUL EECS for Kinect programming support.

7. REFERENCES

[1] S. Abdallah and M. Plumbley. Information dynamics: patterns of expectation and surprise in the perception of music. *Connection Science*, 21(2):89–117, 2009.

[2] L. Meyer. Music, the arts, and ideas: Patterns and Predictions in Twentieth-Century Culture. University of Chicago Press, Chicago, 1967.

[3] D. Müllensiefen, B. Gingras, L. Stewart, and J. Musil. *Goldsmiths Musical Sophistication Index (Gold-MSI): Technical Report and Documentation v0. 9.* Goldsmiths, University of London, London, 2011.

[4] E. Narmour. *Beyond Schenkerism.* the need for alternatives in music analysis. Univ of Chicago Press, 1977.

[5] M. T. Pearce. *The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition.* PhD thesis, City University, London, Jan. 2005.

[6] K. Potter, G. Wiggins, and M. T. Pearce. An Objective Basis for Music Theory: Information-Dynamic Analysis of Minimalist Music. *Musicae Scientiae*, 11(2):295–322, Feb. 2007.