

Cognitive Music Modelling: an Information Dynamics Approach

Samer A. Abdallah, Henrik Ekeus, Peter Foster
Andrew Robertson and Mark D. Plumbley
Centre for Digital Music
Queen Mary University of London
Mile End Road, London E1 4NS

Abstract—We describe an information-theoretic approach to the analysis of music and other sequential data, which emphasises the predictive aspects of perception, and the dynamic process of forming and modifying expectations about an unfolding stream of data, characterising these using the tools of information theory: entropies, mutual informations, and related quantities. After reviewing the theoretical foundations, we discuss a few emerging areas of application, including musicological analysis, real-time beat-tracking analysis, and the generation of musical materials as a cognitively-informed compositional aid.

I. INTRODUCTION

The relationship between Shannon’s [1] information theory and music and art in general has been the subject of some interest since the 1950s [2]–[6]. The general thesis is that perceptible qualities and subjective states like uncertainty, surprise, complexity, tension, and interestingness are closely related to information-theoretic quantities like entropy, relative entropy, and mutual information.

Music is also an inherently dynamic process, where listeners build up expectations about what is to happen next, which may be fulfilled immediately, after some delay, or modified as the music unfolds. In this paper, we explore this “Information Dynamics” view of music, discussing the theory behind it and some emerging applications.

A. Expectation and surprise in music

The idea that the musical experience is strongly shaped by the generation and playing out of strong and weak expectations was put forward by, amongst others, music theorists L. B. Meyer [5] and Narmour [7], but was recognised much earlier; for example, it was elegantly put by Hanslick [8] in the nineteenth century:

“The most important factor in the mental process which accompanies the act of listening to music, and which converts it to a source of pleasure, is . . . the intellectual satisfaction which the listener derives from continually following and anticipating the composer’s intentions—now, to see his expectations fulfilled, and now, to find himself agreeably mistaken.

An essential aspect of this is that music is experienced as a phenomenon that unfolds in time, rather than being apprehended as a static object presented in its entirety. Meyer argued that the experience depends on how we change and revise

our conceptions *as events happen*, on how expectation and prediction interact with occurrence, and that, to a large degree, the way to understand the effect of music is to focus on this ‘kinetics’ of expectation and surprise.

Prediction and expectation are essentially probabilistic concepts and can be treated mathematically using probability theory. We suppose that when we listen to music, expectations are created on the basis of our familiarity with various styles of music and our ability to detect and learn statistical regularities in the music as they emerge. There is experimental evidence that human listeners are able to internalise statistical knowledge about musical structure, e.g. [9], and also that statistical models can form an effective basis for computational analysis of music, e.g. [10]–[12].

B. Information dynamic approach

Our working hypothesis is that, as an intelligent, predictive agent (to which we will refer as ‘it’) listens to a piece of music, it maintains a dynamically evolving probabilistic belief state that enables it to make predictions about how the piece will continue, relying on both its previous experience of music and the emerging themes of the piece. As events unfold, it revises this belief state, which includes predictive distributions over possible future events. These can be characterised in terms of a handful of information theoretic-measures such as entropy and relative entropy, what Berlyne [13] called ‘collative variables’, since they are to do with *patterns* of occurrence, rather than the details of which specific things occur, and developed the ideas of ‘information aesthetics’ in an experimental setting. By tracing the evolution of these measures, we obtain a representation which captures much of the significant structure of the music.

One consequence of this approach is that regardless of the details of the sensory input or even which sensory modality is being processed, the resulting analysis is in terms of the same units: quantities of information (bits) and rates of information flow (bits per second). The information theoretic concepts in terms of which the analysis is framed are universal to all sorts of data. Together, these suggest that an information dynamic analysis captures a high level of *abstraction*, and could be used to make structural comparisons between different temporal media, such as music, film, animation, and dance.

Another consequence is that the information dynamic approach gives us a principled way to address the notion of *subjectivity*, since the analysis is dependent on the probability model the observer starts off with, which may depend on prior experience or other factors, and which may change over time. Thus, inter-subject variability and variation in subjects' responses over time are fundamental to the theory.

II. THEORETICAL REVIEW

A. Entropy and information

Let X denote some variable whose value is initially unknown to our hypothetical observer. We will treat X mathematically as a random variable, with a value to be drawn from some set \mathcal{X} and a probability distribution representing the observer's beliefs about the true value of X . In this case, the observer's uncertainty about X can be quantified as the entropy of the random variable $H(X)$. For a discrete variable with probability mass function $p : \mathcal{X} \rightarrow [0, 1]$, this is

$$H(X) = \sum_{x \in \mathcal{X}} -p(x) \log p(x), \quad (1)$$

The negative-log-probability $\ell(x) = -\log p(x)$ of a particular value x can usefully be thought of as the *surprisingness* of the value x should it be observed, and hence the entropy is the expectation of the surprisingness, $E \ell(X)$.

Now suppose that the observer receives some new data \mathcal{D} that causes a revision of its beliefs about X . The *information* in this new data *about* X can be quantified as the relative entropy or Kullback-Leibler (KL) divergence between the prior and posterior distributions $p(x)$ and $p(x|\mathcal{D})$ respectively:

$$\mathcal{I}_{\mathcal{D} \rightarrow X} = D(p_{X|\mathcal{D}} \| p_X) = \sum_{x \in \mathcal{X}} p(x|\mathcal{D}) \log \frac{p(x|\mathcal{D})}{p(x)}. \quad (2)$$

When there are multiple variables X_1, X_2 etc. which the observer believes to be dependent, then the observation of one may change its beliefs and hence yield information about the others. The joint and conditional entropies as described in any textbook on information theory (e.g. [14]) then quantify the observer's expected uncertainty about groups of variables given the values of others. In particular, the *mutual information* $I(X_1; X_2)$ is both the expected information in an observation of X_2 about X_1 and the expected reduction in uncertainty about X_1 after observing X_2 :

$$I(X_1; X_2) = H(X_1) - H(X_1|X_2), \quad (3)$$

where $H(X_1|X_2) = H(X_1, X_2) - H(X_2)$ is the conditional entropy of X_1 given X_2 . A little algebra shows that $I(X_1; X_2) = I(X_2; X_1)$ and so the mutual information is symmetric in its arguments. A conditional form of the mutual information can be formulated analogously:

$$I(X_1; X_2|X_3) = H(X_1|X_3) - H(X_1|X_2, X_3). \quad (4)$$

These relationships between the various entropies and mutual informations are conveniently visualised in *information diagrams* or *I-diagrams* [15] such as the one in fig. 1.

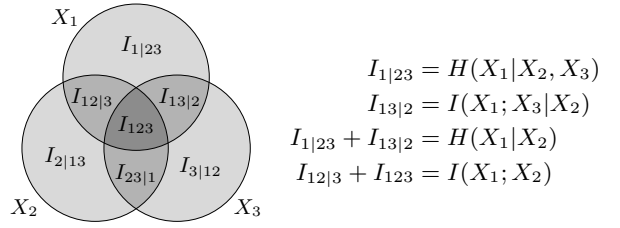


Fig. 1. I-diagram of entropies and mutual informations for three random variables X_1, X_2 and X_3 . The areas of the three circles represent $H(X_1)$, $H(X_2)$ and $H(X_3)$ respectively. The total shaded area is the joint entropy $H(X_1, X_2, X_3)$. The central area I_{123} is the co-information [16]. Some other information measures are indicated in the legend.

B. Surprise and information in sequences

Suppose that $(\dots, X_{-1}, X_0, X_1, \dots)$ is a sequence of random variables, infinite in both directions, and that μ is the associated probability measure over all realisations of the sequence. In the following, μ will simply serve as a label for the process. We can identify a number of information-theoretic measures meaningful in the context of a sequential observation of the sequence, during which, at any time t , the sequence can be divided into a 'present' X_t , a 'past' $\bar{X}_t \equiv (\dots, X_{t-2}, X_{t-1})$, and a 'future' $\bar{X}_t \equiv (X_{t+1}, X_{t+2}, \dots)$. We will write the actually observed value of X_t as x_t , and the sequence of observations up to but not including x_t as \bar{x}_t .

The in-context surprisingness of the observation $X_t = x_t$ depends on both x_t and the context \bar{x}_t :

$$\ell_t = -\log p(x_t | \bar{x}_t). \quad (5)$$

However, before X_t is observed, the observer can compute the *expected* surprisingness as a measure of its uncertainty about X_t ; this may be written as an entropy $H(X_t | \bar{X}_t = \bar{x}_t)$, but note that this is conditional on the *event* $\bar{X}_t = \bar{x}_t$, not the *variables* \bar{X}_t as in the conventional conditional entropy.

The surprisingness ℓ_t and expected surprisingness $H(X_t | \bar{X}_t = \bar{x}_t)$ can be understood as *subjective* information dynamic measures, since they are based on the observer's probability model in the context of the actually observed sequence \bar{x}_t . They characterise what it is like to be 'in the observer's shoes'. If we view the observer as a purely passive or reactive agent, this would probably be sufficient, but for active agents such as humans or animals, it is often necessary to *anticipate* future events in order, for example, to plan the most effective course of action. It makes sense for such observers to be concerned about the predictive probability distribution over future events, $p(\bar{x}_t | \bar{x}_t)$. When an observation $X_t = x_t$ is made in this context, the *instantaneous predictive information* (IPI) \mathcal{I}_t at time t is the information in the event $X_t = x_t$ about the entire future of the sequence \bar{X}_t , given the observed past $\bar{X}_t = \bar{x}_t$. Referring to the definition of information (2), this is the KL divergence between prior and posterior distributions over possible futures, which written out

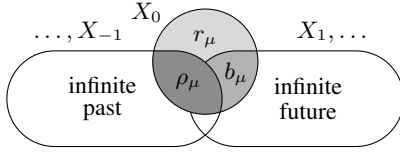


Fig. 2. I-diagram illustrating several information measures in stationary random processes. Each circle or oval represents a random variable or sequence of random variables relative to time $t = 0$. Overlapped areas correspond to various mutual informations. The circle represents the ‘present’. Its total area is $H(X_0) = \rho_\mu + r_\mu + b_\mu$, where ρ_μ is the multi-information rate, r_μ is the residual entropy rate, and b_μ is the predictive information rate. The entropy rate is $h_\mu = r_\mu + b_\mu$.

in full, is

$$\mathcal{I}_t = \sum_{\vec{x}_t \in \mathcal{X}^*} p(\vec{x}_t | x_t, \overleftarrow{x}_t) \log \frac{p(\vec{x}_t | x_t, \overleftarrow{x}_t)}{p(\vec{x}_t | \overleftarrow{x}_t)}, \quad (6)$$

where the sum is to be taken over the set of infinite sequences \mathcal{X}^* . Note that it is quite possible for an event to be surprising but not informative in a predictive sense. As with the surprisingness, the observer can compute its *expected* IPI at time t , which reduces to a mutual information $I(X_t; \vec{X}_t | \overleftarrow{X}_t = \overleftarrow{x}_t)$ conditioned on the observed past. This could be used, for example, as an estimate of attentional resources which should be directed at this stream of data, which may be in competition with other sensory streams.

C. Information measures for stationary random processes

If we step back, out of the observer’s shoes as it were, and consider the random process $(\dots, X_{-1}, X_0, X_1, \dots)$ as a statistical ensemble of possible realisations, and furthermore assume that it is stationary, then it becomes possible to define a number of information-theoretic measures, closely related to those described above, but which characterise the process as a whole, rather than on a moment-by-moment basis. Some of these, such as the entropy rate, are well-known, but others are only recently being investigated. (In the following, the assumption of stationarity means that the measures defined below are independent of t .)

The *entropy rate* of the process is the entropy of the ‘present’ X_t given the ‘past’:

$$h_\mu = H(X_t | \overleftarrow{X}_t). \quad (7)$$

The entropy rate is a measure of the overall surprisingness or unpredictability of the process, and gives an indication of the average level of surprise and uncertainty that would be experienced by an observer computing the measures of § II-B on a sequence sampled from the process.

The *multi-information rate* ρ_μ [17] is the mutual information between the ‘past’ and the ‘present’:

$$\rho_\mu = I(\overleftarrow{X}_t; X_t) = H(X_t) - h_\mu. \quad (8)$$

It is a measure of how much the preceeding context of an observation helps in predicting or reducing the suprisingness of the current observation.

The *excess entropy* [18] is the mutual information between the entire ‘past’ and the entire ‘future’ plus ‘present’:

$$E = I(\overleftarrow{X}_t; X_t, \vec{X}_t). \quad (9)$$

Both the excess entropy and the multi-information rate can be thought of as measures of *redundancy*, quantifying the extent to which the same information is to be found in all parts of the sequence.

The *predictive information rate* (or PIR) [19] is the mutual information between the ‘present’ and the ‘future’ given the ‘past’:

$$b_\mu = I(X_t; \vec{X}_t | \overleftarrow{X}_t) = H(\vec{X}_t | \overleftarrow{X}_t) - H(\vec{X}_t | X_t, \overleftarrow{X}_t), \quad (10)$$

which can be read as the average reduction in uncertainty about the future on learning X_t , given the past. Due to the symmetry of the mutual information, it can also be written as

$$b_\mu = H(X_t | \overleftarrow{X}_t) - H(X_t | \overleftarrow{X}_t, \vec{X}_t) = h_\mu - r_\mu, \quad (11)$$

where $r_\mu = H(X_t | \overleftarrow{X}_t, \vec{X}_t)$, is the *residual* [20], or *erasure* [21] entropy rate. The PIR gives an indication of the average IPI that would be experienced by an observer processing a sequence sampled from this process. The relationship between these various measures are illustrated in Fig. 2; see James et al [22] for further discussion.

D. First and higher order Markov chains

In [19] we derived expressions for all the information measures described in § II-B for ergodic first order Markov chains (i.e. that have a unique stationary distribution). We also showed that the PIR can be expressed simply in terms of entropy rates: if we let a denote the $K \times K$ transition matrix of a Markov chain over an alphabet $\{1, \dots, K\}$, such that $a_{ij} = \Pr(X_t = i | X_{t-1} = j)$, and let $h : \mathbb{R}^{K \times K} \rightarrow \mathbb{R}$ be the entropy rate function such that $h(a)$ is the entropy rate of a Markov chain with transition matrix a , then the PIR is

$$b_\mu = h(a^2) - h(a), \quad (12)$$

where a^2 is the transition matrix of the Markov chain obtained by jumping two steps at a time along the original chain.

Second and higher order Markov chains can be treated in a similar way by transforming to a first order representation of the high order Markov chain. With an N th order model, this is done by forming a new alphabet of size K^N consisting of all possible N -tuples of symbols from the base alphabet. An observation \hat{x}_t in this new model encodes a block of N observations $(x_{t+1}, \dots, x_{t+N})$ from the base model. The new Markov of chain is parameterised by a sparse $K^N \times K^N$ transition matrix \hat{a} , in terms of which the PIR is

$$h_\mu = h(\hat{a}), \quad b_\mu = h(\hat{a}^{N+1}) - Nh(\hat{a}), \quad (13)$$

where \hat{a}^{N+1} is the $(N+1)$ th power of the first order transition matrix. Other information measures can also be computed for the high-order Markov chain, including the multi-information rate ρ_μ and the excess entropy E . (These are identical for first

order Markov chains, but for order N chains, E can be up to N times larger than ρ_μ .)

In our experiments with visualising and sonifying sequences sampled from first order Markov chains [19], we found that the measures h_μ , ρ_μ and b_μ correspond to perceptible characteristics, and that the transition matrices maximising or minimising each of these quantities are quite distinct. High entropy rates are associated with completely uncorrelated sequences with no recognisable temporal structure (and low ρ_μ and b_μ). High values of ρ_μ are associated with long periodic cycles (and low h_μ and b_μ). High values of b_μ are associated with intermediate values of ρ_μ and h_μ , and recognisable, but not completely predictable, temporal structures. These relationships are visible in fig. 7 in § IV, where we pick up this thread again, with an application of information dynamics in a compositional aid.

III. INFORMATION DYNAMICS IN ANALYSIS

A. Musicological Analysis

In [19], we analysed two pieces of music in the minimalist style by Philip Glass: *Two Pages* (1969) and *Gradus* (1968). The analysis was done using a first-order Markov chain model, with the enhancement that the transition matrix of the model was allowed to evolve dynamically as the notes were processed, and was tracked (in a Bayesian way) as a *distribution* over possible transition matrices, rather than a point estimate. Some results are summarised in fig. 3: the upper four plots show the dynamically evolving subjective information measures as described in § II-B, computed using a point estimate of the current transition matrix; the fifth plot (the ‘model information rate’) shows the information in each observation about the transition matrix. In [23], we showed that this ‘model information rate’ is actually a component of the true IPI when the transition matrix is being learned online, and was neglected when we computed the IPI from the transition matrix as if it were a constant.

The peaks of the surprisingness and both components of the IPI show good correspondence with structure of the piece both as marked in the score and as analysed by musicologist Keith Potter, who was asked to mark the six ‘most surprising moments’ of the piece (shown as asterisks in the fifth plot). In contrast, the analyses shown in the lower two plots of fig. 3, obtained using two rule-based music segmentation algorithms, while clearly *reflecting* the structure of the piece, do not *segment* the piece, with no tendency to peaking of the boundary strength function at the boundaries in the piece.

The complete analysis of *Gradus* can be found in [19], but fig. 4 illustrates the result of a metrical analysis: the piece was divided into bars of 32, 64 and 128 notes. In each case, the average surprisingness and IPI for the first, second, third etc. notes in each bar were computed. The plots show that the first note of each bar is, on average, significantly more surprising and informative than the others, up to the 64-note level, where as at the 128-note, level, the dominant periodicity appears to remain at 64 notes.

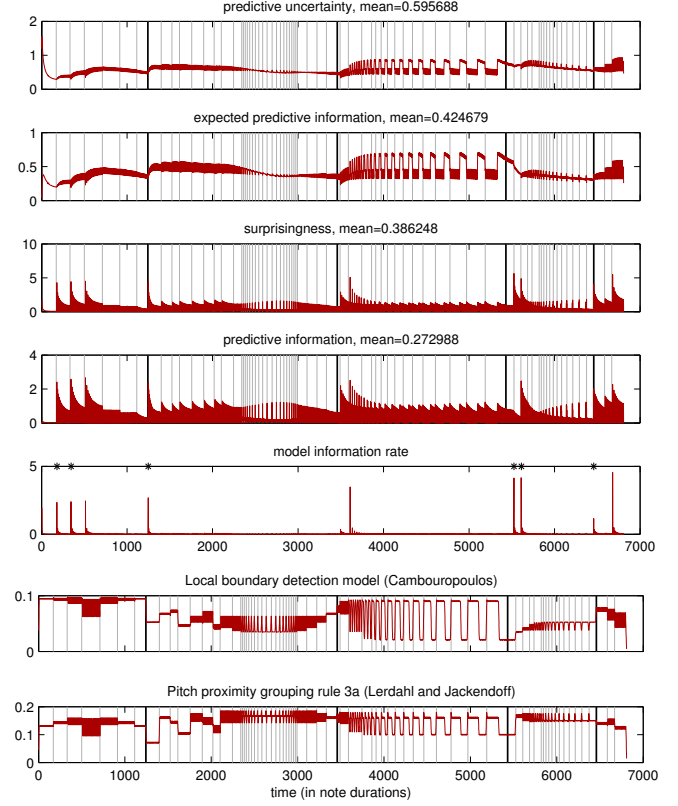


Fig. 3. Analysis of *Two Pages*. The thick vertical lines are the part boundaries as indicated in the score by the composer. The thin grey lines indicate changes in the melodic ‘figures’ of which the piece is constructed. In the ‘model information rate’ panel, the black asterisks mark the six most surprising moments selected by Keith Potter. The bottom two panels show two rule-based boundary strength analyses. All information measures are in nats. [24].

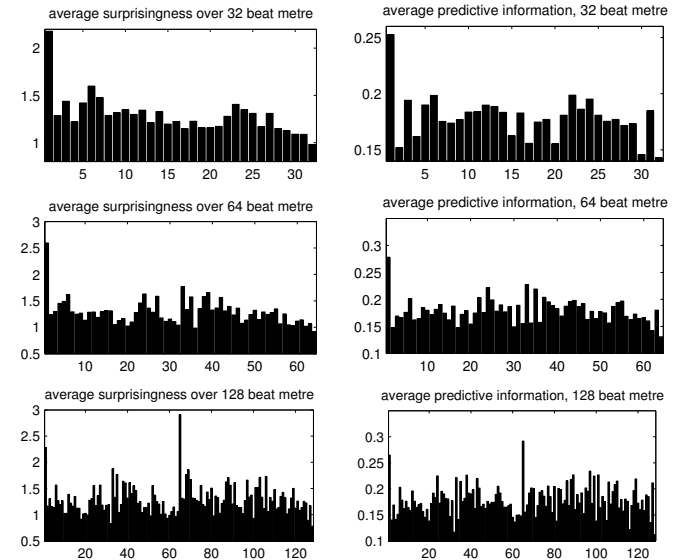


Fig. 4. Metrical analysis by computing average surprisingness and IPI of notes at different periodicities (i.e. hypothetical bar lengths) and phases (i.e. positions within a bar).

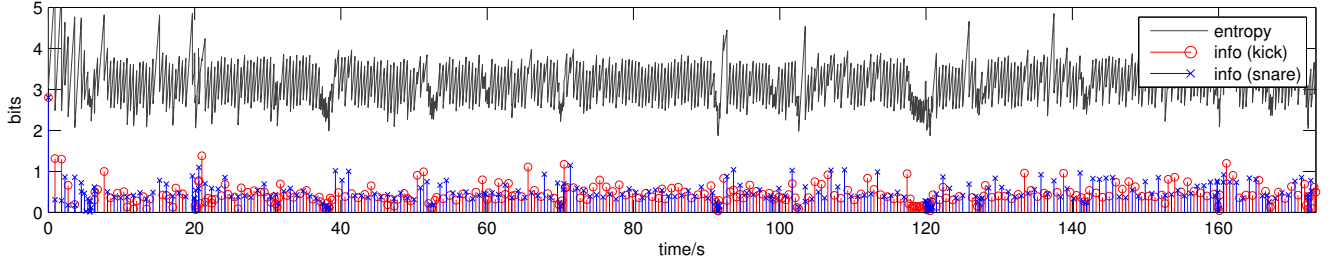


Fig. 5. Information dynamic analysis derived from audio recordings of drumming, obtained by applying a Bayesian beat tracking system to the sequence of detected kick and snare drum events. The grey line show the system’s varying level of uncertainty (entropy) about the tempo and phase of the beat grid, while the stem plot shows the amount of information in each drum event about the beat grid. The entropy drops instantaneously at each event and rises gradually between events.

B. Real-valued signals and audio analysis

Using analogous definitions based on the differential entropy [14], the methods outlined in § II-B and § II-C can be reformulated for random variables taking values in a continuous domain and thus be applied to expressive parameters of music such as dynamics, timing and timbre, which are readily quantified on a continuous scale. Dubnov [17] considers the class of stationary Gaussian processes, for which the entropy rate may be obtained analytically from the power spectral density function $S(\omega)$ of the signal, and found that the multi-information rate can be expressed as

$$\rho_\mu = \frac{1}{2} \left(\log \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) d\omega - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log S(\omega) d\omega \right). \quad (14)$$

Dubnov also notes that $e^{-2\rho_\mu}$ is equivalent to the well-known *spectral flatness measure*, and hence, Gaussian processes with maximal multi-information rate are those with maximally non-flat spectra, which are those dominated by a single frequency component.

We have found (to appear in forthcoming work) that the predictive information for autoregressive Gaussian processes can be expressed as

$$b_\mu = \frac{1}{2} \left(\log \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{S(\omega)} d\omega - \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \frac{1}{S(\omega)} d\omega \right), \quad (15)$$

suggesting a sort of duality between b_μ and ρ_μ which is consistent with the duality between multi-information and predictive information rates we discuss in [25]. A consideration of the residual or erasure entropy rate [21] suggests that this expression applies to Gaussian processes in general but this is yet to be confirmed rigorously.

Analysis shows that in stationary autoregressive processes of a given finite order, ρ_μ is unbounded, while for moving average process of a given order, b_μ is unbounded. This is a result of the physically unattainable infinite precision observations which the theoretical analysis assumes; adding more realistic limitations on the amount of information that can be extracted from one measurement is the one of the aims of our ongoing work in this area.

C. Beat Tracking

A probabilistic method for drum tracking was presented by Robertson [26]. The system infers a beat grid (a sequence of approximately regular beat times) given audio inputs from a live drummer, for the purpose of synchronising a music sequencer with the drummer. The times of kick and snare drum events are obtained using dedicated microphones for each drum and a percussive onset detector [27]. These event times are then sent to the beat tracker, which maintains a belief state in the form of distributions over the tempo and phase of the beat grid. Every time an event is received, these distributions are updated with respect to a probabilistic model which accounts both for tempo and phase variations and the emission of drum events at musically plausible times relative to the beat grid.

The use of a probabilistic belief state means we can compute entropies representing the system’s uncertainty about the beat grid, and quantify the amount of information in each event about the beat grid as the KL divergence between prior and posterior distributions. Though this is not strictly the instantaneous predictive information (IPI) as described in § II-B (the information gained is not directly about future event times), we can treat it as a proxy for the IPI, in the manner of the ‘model information rate’ described in § III-A, which has a similar status.

We carried out the analysis on 16 recordings; an example is shown in fig. 5. There we can see variations in the entropy in the upper graph and the information in each drum event in the lower stem plot. At certain points in time, unusually large amounts of information arrive; these may be related to fills and other rhythmic irregularities, which are often followed by an emphatic return to a steady beat at the beginning of the next bar—this is something we are currently investigating. We also analysed the pattern of information flow on a cyclic metre, much as in fig. 4. All the recordings we analysed are audibly in 4/4 metre, but we found no evidence of a general tendency for greater amounts of information to arrive at metrically strong beats, which suggests that the rhythmic accuracy of the drummers does not vary systematically across each bar. It is possible that metrical information existing in the pattern of kick and snare events might emerge in an analysis using a model that attempts to predict the time and type of the next drum event, rather than just inferring the beat grid as the current model does.

IV. INFORMATION DYNAMICS AS COMPOSITIONAL AID

The use of stochastic processes in music composition has been widespread for decades—for instance Iannis Xenakis applied probabilistic mathematical models to the creation of musical materials [28]. While such processes can drive the *generative* phase of the creative process, information dynamics can serve as a novel framework for a *selective* phase, by providing a set of criteria to be used in judging which of the generated materials are of value. This alternation of generative and selective phases as been noted before [29]. Information-dynamic criteria can also be used as *constraints* on the generative processes, for example, by specifying a certain temporal profile of suprisingness and uncertainty the composer wishes to induce in the listener as the piece unfolds.

In particular, the behaviour of the predictive information rate (PIR) defined in § II-C make it interesting from a compositional point of view. The definition of the PIR is such that it is low both for extremely regular processes, such as constant or periodic sequences, *and* low for extremely random processes, where each symbol is chosen independently of the others, in a kind of ‘white noise’. In the former case, the pattern, once established, is completely predictable and therefore there is no *new* information in subsequent observations. In the latter case, the randomness and independence of all elements of the sequence means that, though potentially surprising, each observation carries no information about the ones to come.

Processes with high PIR maintain a certain kind of balance between predictability and unpredictability in such a way that the observer must continually pay attention to each new observation as it occurs in order to make the best possible predictions about the evolution of the sequence. This balance between predictability and unpredictability is reminiscent of the inverted ‘U’ shape of the Wundt curve (see fig. 6), which summarises the observations of Wundt [30] that stimuli are most pleasing at intermediate levels of novelty or disorder, where there is a balance between ‘order’ and ‘chaos’.

Using the methods of § II-D, we found [19] a similar shape when plotting entropy rate against PIR—this is visible in the upper envelope of the plot in fig. 7, which is a 3-D scatter plot of three of the information measures discussed in § II-C for several thousand first-order Markov chain transition matrices generated by a random sampling method. The coordinates of the ‘information space’ are entropy rate (h_μ), redundancy (ρ_μ), and predictive information rate (b_μ). The points along the ‘redundancy’ axis correspond to periodic Markov chains. Those along the ‘entropy’ axis produce uncorrelated sequences with no temporal structure. Processes with high PIR are to be found at intermediate levels of entropy and redundancy.

A. The Melody Triangle

These observations led us to construct the ‘Melody Triangle’, a graphical interface for the discovery of melodic materials, where the input—positions within a triangle—directly map to information theoretic properties of the output. The triangle is populated with first order Markov chain transition matrices as illustrated in fig. 7. The distribution of transition matrices in

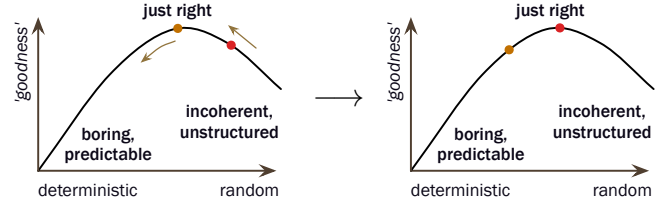


Fig. 6. The Wundt curve relating randomness/complexity with perceived value. Repeated exposure sometimes results in a move to the left along the curve [13].

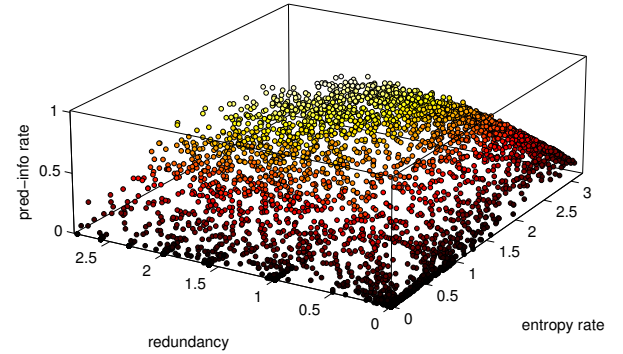


Fig. 7. The population of transition matrices in the 3D space of entropy rate (h_μ), redundancy (ρ_μ) and PIR (b_μ), all in bits. The concentrations of points along the redundancy axis correspond to Markov chains which are roughly periodic with periods of 2 (redundancy 1 bit), 3, 4, etc. all the way to period 7 (redundancy 2.8 bits). The colour of each point represents its PIR—note that the highest values are found at intermediate entropy and redundancy, and that the distribution as a whole makes a curved triangle. Although not visible in this plot, it is largely hollow in the middle.

this space forms a relatively thin curved sheet. Thus, it is a reasonable simplification to project out the third dimension (the PIR) and present an interface that is just two dimensional. The right-angled triangle is rotated, reflected and stretched to form an equilateral triangle with the $h_\mu = 0, \rho_\mu = 0$ vertex at the top, the ‘redundancy’ axis down the left-hand side, and the ‘entropy rate’ axis down the right, as shown in fig. 8. This is our ‘Melody Triangle’ and forms the interface by which the system is controlled. The user selects a point within the triangle, this is mapped into the information space and the nearest transition matrix is used to generate a sequence of values which are then sonified either as pitched notes or percussive sounds. By choosing the position within the triangle, the user can control the output at the level of its ‘collative’ properties, with access to the variety of patterns as described above and in § II-D. Though the interface is 2D, the third dimension (PIR) is implicitly present, as transition matrices retrieved from along the centre line of the triangle will tend to have higher PIR. We hypothesise that, under the appropriate conditions, these will be perceived as more ‘interesting’ or ‘melodic.’

The Melody Triangle exists in two incarnations: a screen-based interface where a user moves tokens in and around a triangle on screen, and a multi-user interactive installation where a Kinect camera tracks individuals in a space and maps their

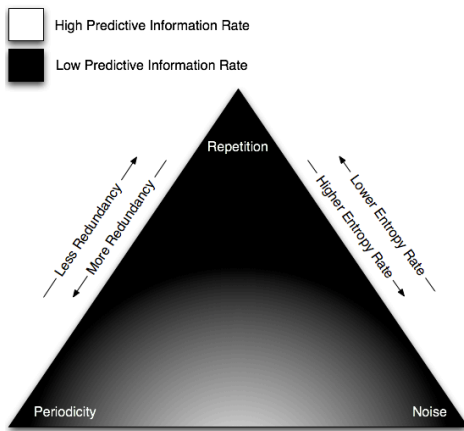


Fig. 8. The Melody Triangle

positions in physical space to the triangle. In the latter each visitor that enters the installation generates a melody and can collaborate with their co-visitors to generate musical textures. This makes the interaction physically engaging and (as our experience with visitors both young and old has demonstrated) more playful. The screen based interface can serve as a compositional tool. A number of tokens, each representing a sonification stream or ‘voice’, can be dragged in and around the triangle. For each token, a sequence of symbols is sampled using the corresponding transition matrix, which are then mapped to notes of a scale or percussive sounds¹. Keyboard commands give control over other musical parameters such as pitch register and inter-onset interval. The system is capable of generating quite intricate musical textures when multiple tokens are in the triangle, but unlike other computer aided composition tools or programming environments, the composer exercises control at the abstract level of information-dynamic properties.

B. User trials with the Melody Triangle

We are currently in the process of using the screen-based Melody Triangle user interface to investigate the relationship between the information-dynamic characteristics of sonified Markov chains and subjective musical preference. We carried out a pilot study with six participants, who were asked to use a simplified form of the user interface (a single controllable token, and no rhythmic, registral or timbral controls) under two conditions: one where a single sequence was sonified under user control, and another where an additional sequence was sonified in a different register, as if generated by a fixed invisible token in one of four regions of the triangle. In addition, subjects were asked to press a key if they ‘liked’ what they were hearing.

We recorded subjects’ behaviour as well as points which they marked with a key press. Some results for two of the subjects are shown in fig. 9. Though we have not been able to detect any systematic across-subjects preference for any particular region of the triangle, subjects do seem to exhibit

¹The sampled sequence could easily be mapped to other musical processes, possibly over different time scales, such as chords, dynamics and timbres. It would also be possible to map the symbols to visual or other outputs.

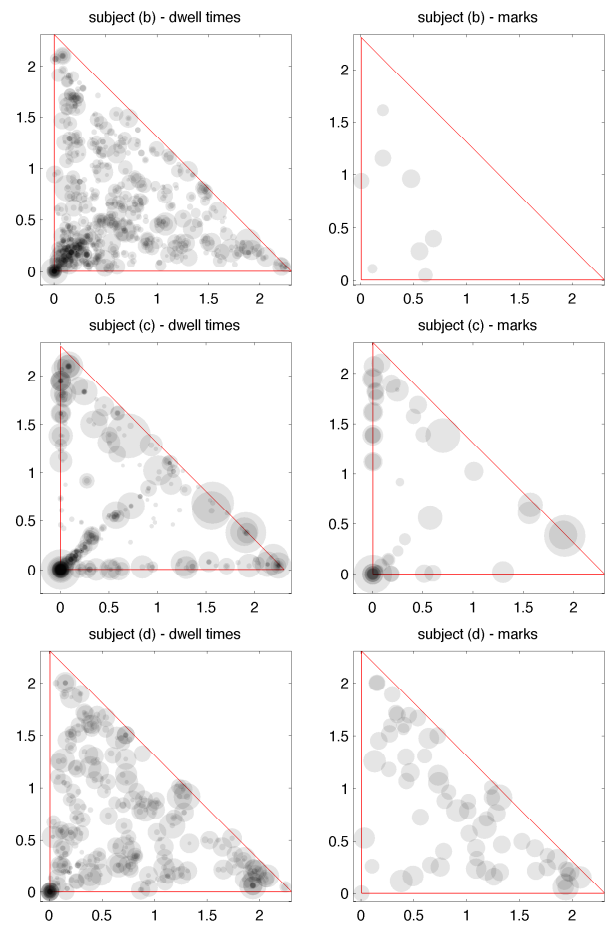


Fig. 9. Dwell times and mark positions from user trials with the on-screen Melody Triangle interface, for three subjects. The left-hand column shows the positions in a 2D information space (entropy rate vs multi-information rate in bits) where each spent their time; the area of each circle is proportional to the time spent there. The right-hand column shows point which subjects ‘liked’; the area of the circles here is proportional to the duration spent at that point before the point was marked.

distinct kinds of exploratory behaviour. Our initial hypothesis, that subjects would linger longer in regions of the triangle that produced aesthetically preferable sequences, and that this would tend to be towards the centre line of the triangle for all subjects, was not confirmed. However, it is possible that the design of the experiment encouraged an initial exploration of the space (sometimes very systematic, as for subject c) aimed at *understanding* how the system works, rather than finding musical patterns. It is also possible that the system encourages users to create musically interesting output by *moving the token*, rather than finding a particular spot in the triangle which produces a musically interesting sequence by itself.

Comments collected from the subjects suggest that the information-dynamic characteristics of the patterns were readily apparent to most: several noticed the main organisation of the triangle, with repetitive notes at the top, cyclic patterns along one edge, and unpredictable notes towards the opposite corner. Some described their systematic exploration of the space. Two

felt that the right side was ‘more controllable’ than the left (a consequence of their ability to return to a particular distinctive pattern and recognise it as one heard previously). Two reported that they became bored towards the end, but another felt there wasn’t enough time to ‘hear out’ the patterns properly. One subject did not ‘enjoy’ the patterns in the lower region, but another said the lower central regions were more ‘melodic’ and ‘interesting’.

We plan to continue the trials with a slightly less restricted user interface in order to make the experience more enjoyable and thereby give subjects longer to use the interface; this may allow them to get beyond the initial exploratory phase and give a clearer picture of their aesthetic preferences. In addition, we plan to conduct a study under more restrictive conditions, where subjects will have no control over the patterns other than to signal (a) which of two alternatives they prefer in a forced choice paradigm, and (b) when they are bored of listening to a given sequence.

V. CONCLUSIONS

We have looked at several emerging areas of application of the methods and ideas of information dynamics to various problems in music analysis, perception and cognition, including musicological analysis of symbolic music, audio analysis, rhythm processing and compositional and creative tasks. The approach has proved successful in musicological analysis, and though our initial data on rhythm processing and aesthetic preference are inconclusive, there is still plenty of work to be done in this area: where-ever there are probabilistic models, information dynamics can shed light on their behaviour.

ACKNOWLEDGMENTS

This work is supported by EPSRC Doctoral Training Centre EP/G03723X/1 (HE), GR/S82213/01 and EP/E045235/1(SA), an EPSRC DTA Studentship (PF), an RAEng/EPSRC Research Fellowship 10216/88 (AR), an EPSRC Leadership Fellowship, EP/G007144/1 (MDP) and EPSRC IDyOM2 EP/H013059/1. This work is partly funded by the CoSound project, funded by the Danish Agency for Science, Technology and Innovation. Thanks also Marcus Pearce for providing the two rule-based analyses of *Two Pages*.

REFERENCES

- [1] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.
- [2] J. E. Youngblood, “Style as information,” *Journal of Music Theory*, vol. 2, pp. 24–35, 1958.
- [3] E. Coons and D. Kraehenbuehl, “Information as a measure of structure in music,” *Journal of Music Theory*, vol. 2, no. 2, pp. 127–161, 1958.
- [4] A. Moles, *Information Theory and Esthetic Perception*. University of Illinois Press, 1966.
- [5] L. B. Meyer, *Music, the arts and ideas: Patterns and Predictions in Twentieth-century culture*. University of Chicago Press, 1967.
- [6] J. E. Cohen, “Information theory and music,” *Behavioral Science*, vol. 7, no. 2, pp. 137–163, 1962.
- [7] E. Narmour, *Beyond Schenkerism*. University of Chicago Press, 1977.
- [8] E. Hanslick, *On the musically beautiful: A contribution towards the revision of the aesthetics of music*. Indianapolis, IN: Hackett, 1854/1986.
- [9] J. R. Saffran, E. K. Johnson, R. N. Aslin, and E. L. Newport, “Statistical learning of tone sequences by human infants and adults,” *Cognition*, vol. 70, no. 1, pp. 27–52, 1999.

- [10] D. Conklin and I. H. Witten, “Multiple viewpoint systems for music prediction,” *Journal of New Music Research*, vol. 24, no. 1, pp. 51–73, 1995.
- [11] D. Ponsford, G. A. Wiggins, and C. S. Mellish, “Statistical learning of harmonic movement,” *Journal of New Music Research*, vol. 28, no. 2, pp. 150–177, 1999, also available as Research Paper 874, from the Division of Informatics, University of Edinburgh.
- [12] M. T. Pearce, “The construction and evaluation of statistical models of melodic structure in music perception and composition,” Ph.D. dissertation, Department of Computing, City University, London, 2005.
- [13] D. E. Berlyne, *Aesthetics and Psychobiology*. New York: Appleton Century Crofts, 1971.
- [14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley and Sons, 1991.
- [15] R. Yeung, “A new outlook on Shannon’s information measures,” *Information Theory, IEEE Transactions on*, vol. 37, no. 3, pp. 466–474, 1991.
- [16] W. McGill, “Multivariate information transmission,” *Information Theory, IRE Professional Group on*, vol. 4, no. 4, pp. 93–111, 1954.
- [17] S. Dubnov, “Generalization of spectral flatness measure for non-gaussian linear processes,” *Signal Processing Letters, IEEE*, vol. 11, no. 8, pp. 698–701, 2004.
- [18] J. Crutchfield and N. Packard, “Symbolic dynamics of noisy chaos,” *Physica D: Nonlinear Phenomena*, vol. 7, no. 1-3, pp. 201–223, 1983.
- [19] S. A. Abdallah and M. D. Plumbley, “Information dynamics: Patterns of expectation and surprise in the perception of music,” *Connection Science*, vol. 21, no. 2, pp. 89–117, 2009.
- [20] —, “Predictive information, multi-information and binding information,” Queen Mary University of London, Tech. Rep. C4DM-TR-10-10, 2010.
- [21] S. Verdú and T. Weissman, “Erasure entropy,” in *IEEE International Symposium on Information Theory (ISIT 2006)*, 2006, pp. 98–102.
- [22] R. G. James, C. J. Ellison, and J. P. Crutchfield, “Anatomy of a bit: Information in a time series observation,” *Chaos*, vol. 21, no. 3, p. 037109, 2011.
- [23] S. A. Abdallah and M. D. Plumbley, “Predictive information and Bayesian surprise in exchangeable random processes,” Queen Mary University of London, Tech. Rep. C4DM-TR10-09, 2010.
- [24] K. Potter, G. A. Wiggins, and M. T. Pearce, “Towards greater objectivity in music theory: Information-dynamic analysis of minimalist music,” *Musicae Scientiae*, vol. 11, no. 2, pp. 295–324, 2007.
- [25] S. A. Abdallah and M. D. Plumbley, “A measure of statistical complexity based on predictive information with application to finite spin systems,” *Physics Letters A*, vol. 376, no. 4, pp. 275 – 281, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0375960111013880>
- [26] A. Robertson, “A Bayesian approach to drum tracking,” in *Proceedings of the 8th Sound and Music Computing Conference (SMC)*, 2011.
- [27] M. Puckette, T. Apel, and D. Zicarelli, “Real-time audio analysis tools for Pd and MSP,” in *Proceedings of the International Computer Music Conference*, 1998, pp. 109–112.
- [28] I. Xenakis, *Formalized music : thought and mathematics in composition*. Stuyvesant, NY: Pendragon Press, 1992.
- [29] M. Boden, *The Creative Mind: Myths and Mechanisms*. Weidenfeld & Nicolson, 1990.
- [30] W. Wundt, *Outlines of Psychology*. Leipzig: Englemann, 1897.