

# **A Computational Model of Meter Cognition During the Audition of Functional Tonal Music: Modeling A-Priori Bias in Meter Cognition**

Jonathan Berger  
CCRMA, Stanford University  
Stanford CA 94305, USA  
brg@ccrma.stanford.edu

Dan Gang  
Institute of Computer Science, Artificial Intelligence Laboratory  
Hebrew University  
Jerusalem, Israel  
dang@cs.huji.ac.il

## **Abstract**

We describe a series of experiments using sequential neural networks to model the effect of contextual bias in music cognition. The model quantifies the strength and specificity of a virtual listener's expectations while listening to functional tonal harmonic chord sequences. The network integrates pools of duple and triple metric units with pitch class representations of chords. The 'listener' is then exposed to new chord sequences. We interpret the output of each sequential vector as the expectation for the next event. By representing segregated duple and triple metric beat units we visualize the process of metric cognition, and the mutual reliance of metric and functional harmonic expectations in establishing a percept of meter and a context for expecting consequential harmonic activity.

## **1. Background**

Recent studies in cognition address the interdependence and mutual influence of multiple schemas in creating contexts. Schema based studies and models of music cognition [Leman, 1998] have applied Gestalt approaches to many aspects of the musical experience. However, few of these studies consider premonitory conditioning in schema selection. In this paper we show that a-priori contextual bias can strongly influence a sequential neural network model of music cognition and propose that this influence carries strong implications in understanding how a listener arrives at metric awareness (or, at the very least, suggests that this factor must be taken into consideration in modeling).

The interplay between conditioning and habituation in attending to, suppressing, repressing or rejecting contexts has been studied both from biological [Ricker et al, 1993] and ecological approaches, also see [Wagner, 1989] for a neural network model of conditioning. Bias in auditory priming has been studied in speech perception [Ratcliff, et al 1996].

Dessain and Honing's study of beat induction marvels at the human ability to induce a strong sense of beat "Only after a few (5-10) notes" in a bottom-up process. They go on to note the simultaneous presence of a 'top-down' process that creates a metrical framework to build expectations. "When in a change of meter the evidence for the old percept becomes to meager, a new beat interpretation is induced." [Dessain and Honing 1994]. A-priori expectations can be among the factors in the speed and efficiency of beat induction (is a few 5 or 10?). The time that it takes a listener to be cognizant of the meter is one possible measure of metric

clarity.

Although there have been studies regarding the influence of musical contexts on meter processing [Keller, 96] no studies that we are aware of have considered the role of a-priori expectations for meter on metric cognition. We call this phenomena contextual bias. Bias plays an important role in establishing the degree of realized expectation (DRE). Examples of a-priori contextual bias include:

1. Biased expectation for duple meter suggested by [Keller, 96].

In the absence of contextual hints there is empirical evidence of an initial assumption of duple meter.

2. The expectation from prior experience that a work or movement will be in a given meter.

An experienced listener, for example, would expect the second movement of a Haydn string quartet to be a minuet. Although each of these particular minuets contain peculiarities that play on expectations, one (op 77 no. 4) begins in duple meter, creating an immediate conflict with premonitory expectations. 3. A metric change or diversion in the course of a piece.

Once a listener 'selects' a metric framework, any subsequent change in meter will require a certain amount of reconditioning (based upon how abrupt the change is). Such a change can include a prolonged metric shift (example x) or a temporary metric diversion (for example, a hemiola) .

The range of the above examples suggests that metric bias deserves careful attention and must be integrated into any study or model of beat induction.

## **2. Design Issues, Architecture and Representation**

Our approach to study the complex issue of premonitory conditioning was to model it with simple means. In previous publications [Gang and Berger, 97], [Gang and Berger, 96], [Berger and Gang, 96] we described a neural network model of the interaction of duple and triple metric schemas with isochronous harmonic progressions. In this model we trained a sequential neural network with a repertoire of metered tonal progressions in duple and triple meter. We then introduced unambiguous, ambiguous and anomalous progressions to our virtual listener and studied the interaction and mutual influence of metric and harmonic expectations. In this paper we describe and compare two sets of experiments that model a-priori metric bias of our virtual listener achieved by varying the training strategy we were able to control bias.

Our model uses a sequential neural network with two pools of metric units (3 units for triple and 4 units for quadruple meter) and a pool of 12 units representing normalized pitch class representations of chord tones. The state layer is composed of the two pools of metric units and the pool of PCs. The state units are used to establish a context that influences the prediction of the next element of the sequential information. The output layer contains the same pools of units as the state layer. The metric units represent the predicted interpretation of the net for the current metric position. The 12 PC units in the output layer, represent the prediction for the subsequent chord. In the case of the metric units, the output is fed back into the corresponding pools in the state and added to the context. In the case of the PC units, the context is updated with the target, instead of the actual output. The metric pools in the net's state units are fully connected to the hidden layer together with the pool of pitch classes, actually implementing the integration of the mutual influences of meter and harmony. The hidden units are fully connected to the output layer.

In the case of the metric pool the update rule of the state is:

$\text{State\_meter}(t) = \text{State\_meter}(t-1) * \text{Decay\_meter} + \text{Output}(t-1).$

Where the Decay\_meter is between 0 to 1.

This rule simulates the fact that the listener is unassisted in its metric interpretation. In the learning phase we feed back the actual output but use the target meter to train the net. In the generalization phase the meter is unknown, hence there is no target.

The state update rule for the pitch class pool is:

$\text{State\_harmony}(t) = \text{State\_meter}(t-1) * \text{Decay\_harmony} + \text{Target}(t-1).$

This rule simulates the fact that the listener is concurrently processing the present chord and expecting the chord to follow. Thus we feed the actual "heard" chord (target at t-1) and not the expectations of the chord at time t-1, into the state.

Our model quantifies the strength and specificity of a virtual listener's expectations while listening to functional tonal harmonic chord sequences. The network integrates pools of duple and triple metric units with pitch class representations of chords. The 'listener' is then exposed to new chord sequences. We interpret the output of each sequential vector as the expectation for the next event. By representing segregated duple and triple metric beat units we visualize the process of metric cognition, and the mutual reliance of metric and functional harmonic expectations in establishing a percept of meter and a context for expecting consequential harmonic activity.

We classify expectations according to predictive qualifiers.

The degree of realized expectation (DRE) is a measure of the correspondence between a prediction and the associated event. DRE is a composite of two indices, one of surprise (DSp) and the other of ambiguity (DA).

A high DRE results from frequency of occurrence (normative) and from normative placement. The strength and specificity of the expectations must correspond with that of the expected consequent. Low DRE's represent surprises and/or ambiguities. The ability of neural networks to generalize time ordered sequential data and deduce patterns of varying degrees of abstraction make this approach a suitable one for modeling these predictive states.

### **3. On Meter**

Meter is fundamental to formulating musical expectations. In [Berger and Gang, 1995] we describe a model of music cognition in which a metric counter was critical in building meaningful expectations. In a subsequent paper we described a model of the mutual influence of meter and harmonic rhythm [Gang and Berger, 1996]. Our research proceeded to examine the mechanism of metric perception [Berger and Gang, 1997]. We proceed here to examine the role of context on formulating metric expectations.

We note the following attributes of meter:

1. Meter is repetitive - its periodicity allows for organization.
2. Meter is hierarchical.
3. In tonal music functional harmonic relationships generally adhere to placement constraints according to this hierarchy.

4. Meter is data reductive - it rules out certain interpretations thus greatly reducing ambiguities. This feature is an outgrowth of items 2 and 3.
5. Metric determination comprises creating (sometimes complex) interactions of differing temporal structures. The sense of metric conflict resulting from (for example) a misalignment of harmonic rhythm and the metric framework is both a cause and a resultant of this property.
6. Conflicts with and variations within a metric framework reduces habituation.
7. Interactions between metric organization and other musical parameters (in the case of our research, harmonic rhythm) can pose contradictions to meter thus promoting the formulation of hypotheses which constitute expectations.

### **3. On Measuring DS, DA and DRE:**

We classify output activations in terms of their strength and specificity. The strength is graphically represented by the size of activation in the output. The specificity describes the distribution of output activations. Output activations can be:

strong and specific  
 strong and unspecific  
 weak and specific/unspecific

A strong and specific output that correctly predicts the input constitutes a high DRE.

Musical surprises and/or ambiguities are represented by conditions in which:

1. strong and specific activations do not match the corresponding target vector. An example of this situation occurs in beat 5 of example 1.1 in which the output is strong and specific for [0,5,9] (IV) but not matching the target [0,4,9] (vi). This constitutes a surprise.
2. activations are unspecific. An example of unspecific activations occurs in beat 6 of example 1.1 in which activations for [0,4,5,7,9] (with a weak activation for 2) fail to represent a singular expectation (possible inferences include I,IV or vi). Since specificity quantifies the distribution of activations, a qualitative description of a lack of specificity would be ambiguity.

Metric expectations are classified as follows:

A normative metric event is a strong and unique activation (only one metric pulse activated at one moment) that bears sequential consequence (a beat is followed by the next beat) and is a member of at least one period of a metric group.

Two situations of metric ambiguity exist. These are:

1. event could be the same beat position in more than one metric pools, beat could
2. event could be different beat positions in more than one metric pools

There are 2 types of metric specificity -

two or more activations in a single event within a single meter pool  
 one or more activations simultaneously in multiple meter pools

metric shift is not simultaneous - that is it occurs between pools and in consequential beat units.

## **4. Description of the Experiments**

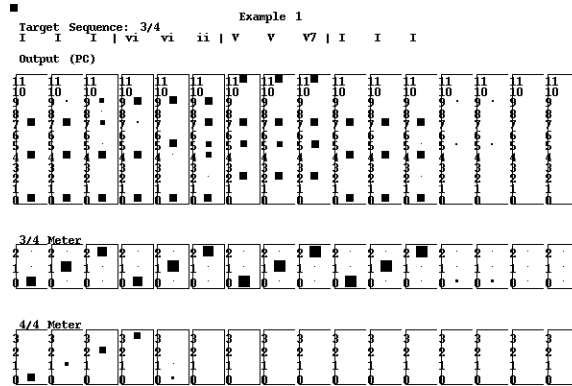
### **4.1 Experiment 1: Unambiguous tonal progressions:**

We first describe cognition of a tonal sequence that is unambiguous both in its harmonic progression as well as in the metric placement of harmonic events. (That is, both what occurs and where it occurs in time conforms to experientially derived expectations).

We proceed to describe the output of our biasing experiments in terms of DS and DA.

In experiment 1 we analyzed the output of the network on a a four measure progression in triple meter [I I I | vi vi ii | V V V7 | I I I] after training the network with a learning set whose order was first randomized, next biased to triple meter and finally biased towards duple meter.

#### 4.1.a. example 1.1 Triple meter, high DRE, minimal bias

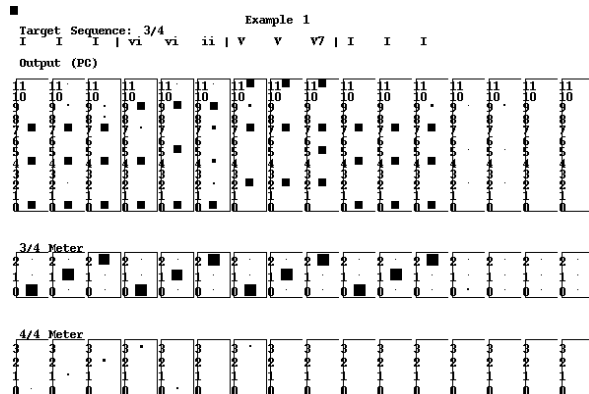


In this example bias has been minimized by randomizing the training set.

Note that the output of both meter pools are activated at the start, representing openness for either duple or triple meter. The recurrence of vi on beat 5 squelches the continuation of expectation of duple meter after the weak activation of a downbeat. This plausible listening strategy is entirely consistent with the harmonic rhythm since the network is not trained with harmonic rhythms that cross measure boundaries.

Of particular interest is the distribution of activations in beat 3 in which the lack of a-priori metric preference contributes to expectations for change. The change to a submediant in beat 4 weakens the plausibility of quadruple meter. The repetition of this harmony in beat 5 completely obliterates activations in the quadruple pool.

#### 4.1.b. example 1.2 Triple meter, high DRE, triple bias

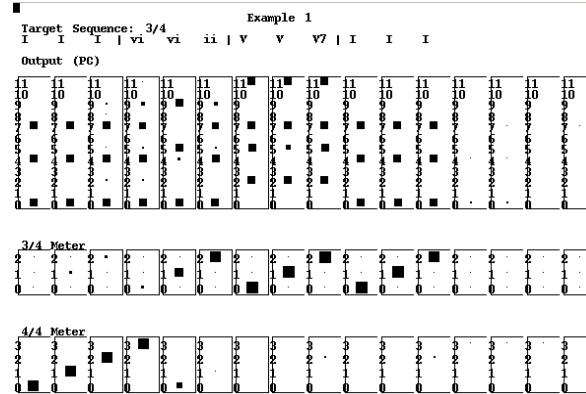


In this example the same progression is introduced to the network, which, except for the fact that the training method creates a biased preference for triple meter, is identical to that described above.

In beat 3 of example 1.2 the strong preference for triple meter reduces the ambiguity noted in example 1.1.

Although example 1.1 doesn't reflect a strong musical surprise, the openness of the virtual listener to either meter creates a subtle but important conflict that is sharply reduced in the model of a listener who, in essence, can correctly tap her foot before the music event starts.

#### 4.1.c. example 1.3 triple meter, high DRE, quadruple bias

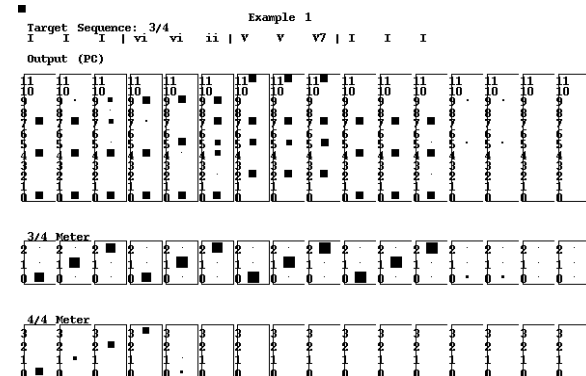


In this example the same triple metered harmonic progression is presented to the network which is biased towards quadruple meter. The incorrect prediction that a tonic harmony will continue in beat four is a direct result of this bias. This situation represents a listener who has a strong premonitory belief that the piece will adhere to a quadruple meter. The ability of the 'listener' in beat 5 to readily adapt by switching activations across meter pools from an incorrect prediction of beat 4 in 4/4 to a correct prediction of beat 2 in 3/4 visualizes the mutually influential organizational power of harmonic rhythm and meter.

## 4.2 Experiment 2: Ambiguous harmonic rhythmic progression

In experiment two we analyzed the output of the network when input with a far more problematic progression. The progression [I IV V I vi V7 I IV ii V V I I I I I] is ambiguous from beat one through beat seven, being plausibly parsable in both triple and quadruple meter. The dominant-tonic pair in beats 4-5 hints at triple meter. This is substantiated in the V-I progression in beats 6-7. However the arrival at the tonic on beat 12 would violate expectations in that it falls on the third beat of a triple measure.

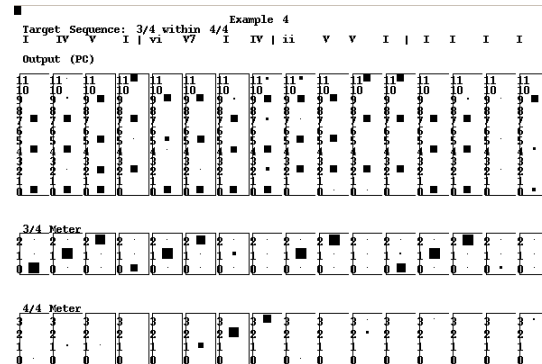
#### 4.2.a. example 2.1 high metric DA, minimal bias



The unbiased 'listener', influenced by the first dominant-tonic pair, rejects the duple interpretation in the

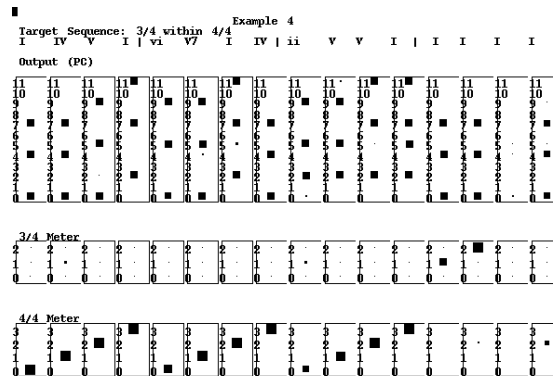
first measure (note the strong and specific expectation for beat 2 in triple meter at beat 5). However the tonic arrival on beat 12 creates a situation in which metric expectations are weak and unspecific.

#### 4.2.b. example 2.2 high metric DA, triple bias



The triple-biased 'listener' seems (oddly at first glance) to be more resistant to a triple meter interpretation than the unbiased listener was. The reason for this lies in the influential interaction of harmonic prediction and metric interpretation. Of particular interest is the shift in meter resulting from the dominant-tonic resolution in beat 12. In this case, there is a strong and specific triple interpretation albeit shifted to a new accentual position.

#### 4.2.c. example 2.3 high metric DA, quadruple bias



Until the surprise sounding of the tonic on beat 12 (the fourth beat in a 4/4 parsing), the quadruple-biased listener manages to maintain a persistent duple interpretation. Of interest in this example is the weakened activation in the downbeat of measure 3 (beat 9) resulting from the disparity between predicted harmonic function and the target chords from beat 6.

## 5. Conclusion and Summary

Significant and highly interpretable differences in output is evident when simple generalization strategies are varied. Research on algorithms to minimize errors during the generalization phase [Gish, 1992] focus

primarily on size of training sets needed during the learning phase. Although we are not concerned directly with efficiency of learning a useful measure taken from these studies is the Maximum likelihood (ML) criterion [cite] which is a distance measure from data points to decision boundary. [Rumelhart and McClelland, 1986] describe a competitive learning algorithm (discussed in musical contexts in [Bharucha and Todd, 1989]).

The cognitive and music analytic implications of these differences suggest that meter acquisition specifically, and, perhaps, any schema identification process is highly influenced by a-priori expectations. This observation challenges many assumptions made in cognitive modeling studies.

We propose that a listener incorporates extra musical expectations in presuming the initial metric organization of the music being attended to. Contributors to this bias could include biological or experiential preference for a given metric schema, the influence of metric cues heard prior to audition of the work, or veridical expectations based upon style or genre. We feel that the establishment of metric awareness must adapt to or overcome these biases while simultaneously searching for pattern based periodicities within the music in order to recognize metric organization. By controlling the order of training we model the premonitory contextual bias and observe how the musical patterns either reinforce or subvert these a-priori preferences.

## **6. References**

[Berger and Gang, 96] Modeling the Degree of Realized Expectation in Music: A Study of Perceptual and Cognitive Modeling Using Neural Networks, Proceedings of the International Computer Music Conference, Hong Kong, August 1996.

[Bharucha and Todd, 1989] Bharucha, Jamshed and P. Todd. Modeling the Perception of Tonal Structure with Neural Nets. Computer Music Journal, Vo. 13, No. 4, Winter 1989. Reprinted in Todd and Loy, Music and Connectionism, page 128.

[Dessain and Honing, 1994] Desain, Peter., and H. Honing. Foot-Tapping: a brief introduction to beat induction. In Proceedings of the 1994 International Computer Music Conference. Page 78..

[Gang and Berger, 97] A Neural Network Model of Metric Perception and Cognition in the Audition of Functional Tonal Music. Proceedings of the International Computer Music Conference, Thessaloniki, September, 1997.

[Gish, 1992] Gish, Herbert. A Minimum Classification Error, Maximum Likelihood, Neural Network. Proceedings of the 1992 International Conference on Acoustics, Speech and Signal Processing. IEEE 92CH3103-9, Page 289.

[Leman, 1998] Leman, Marc, editor. Music Gestalt and Computing: Systems in Cognitive and Systematic Musicology, Springer, 1998.

[Rumelhart and McClelland, 1986] Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1. Cambridge, MIT Press, 1986.



