ECRONICON OPEN ACCESS

EC NEUROLOGY
Review Article

A Dynamic Model Based on Bayesian Inference for Continuous Musical Expectation

Keyvan Yahya*

Chemnitz University of Technology, Chemnitz, Germany

***Corresponding Author:** Keyvan Yahya, Chemnitz University of Technology, Chemnitz, Germany.

Received: July 30, 2020; **Published:** November 28, 2020

Abstract

This paper aims to give an insight about continuous expectation in music based on a relatively broad range of the relevant literature. After providing a preliminary framework that holds on to the established liaison of music expectation and active inference, we put forth a set of hypothesis that altogether suggest the possibility of a biologically plausible dynamic model that could fully account for the inference of expectations in monophonic music. We also put a special emphasis on the Bayesian theory of music perception that underpins these types of models that could infer what happens next during the successive time windows of a given melody on a basis of prior context or technically speaking, generative models denoted by $\sum P(x \mid n)$ containing a normal distribution and will be updated over time. Finally, in regard with the active actions one must take to infer the hidden parameters of a melody to perform the melodic expectation tasks, we point out to the significance of the concept of m-grams in a more precise and wider building and improvement of the hidden Markov chains representing dynamic expectations. That aside, we will propose that the predictive coding models of monophonic melodies could reveal some certain patterns that would be listened on a regular basis and construct the distinctive quintessence of various local music.

Keywords: *Dynamic Model; Bayesian Inference; Continuous Musical Expectation*

Introduction

Expectation and prediction of events plays a key role in music perception and cognition [1] and more generally is regarded as a fundamental property of the human brain [2]. Music expectation has been extensively studied from different viewpoints including cognitive neuroscience [3], information theory [4-7], or neuropsychology of music [8-10]. However, there are only few works concerned with dynamic and continuous ratings of expectation in monophonic melodies [10]. In particular, it was shown that probability-based models, despite some limitations, are more sensitive to capture the variability in continuous ratings of humans for monophonic melodies [6].

There exist strong behavioral and physiological evidence that the nervous system maintains probabilistic distributions that are updated by sensory information using probabilistic, in particular Bayesian, inference [11,12]. Bayesian inference has been either implicitly or explicitly applied to model several aspects of musical processes such as key-finding [13], melodic and rhythmic expectation [14], harmonic analysis (Raphael and Stoddard, 2004), improvisation (Mavromatis 2005), and learning of melodic patterns [15]. In addition, neural theories derived from Bayesian updating schemes, such as predictive coding (Rao and Ballard 1999) have been used to providean account for musical functions such as pleasure (Gebauer., *et al*. 2012) and rhythm perception [16].

In this work, we propose a biologically inspired Bayesian model to account for melodic expectation in monophonic music. More precisely, this model is based on encoding the probabilities of perceived musical notes as internal belief with a number of generative models and incorporates Bayesian inference to capture the continuous variations in dynamic musical expectation. Ideally, we intend to show that given optimal parameters, this model can predict the responses of the subjects for the dynamic ratings of the melodies in our experimental data.

A Bayesian model for musical expectation

In order to provide an account for dynamic melodic expectation, we propose a Bayesian model, which is consisted of three major parts: input process, generative models, Bayesian inference. A generative model which is used to encode the internal representation for sensory observations of the input process based on some hidden parameters. In other words, the Bayesian model compares the predictions of this generative model to the observed sensory input in order to infer the properties of the presented stimulus. A set of belief values are resulted from this comparison and Bayesian inference. The internal belief values represent the probability of the stimulus that is caused by sensory input (Bitzer., *et al*. 2014).

In musical terms, there are a number of generative models, one for each musical pitch, that encode the internal belief about that specific class of musical pitch. These generative models indicate how probable is the observed stimulus, given each musical note. Through Bayesian inference, the model will derive the probability of the internal belief (perceived pitch) encoded by the nervous system given the musical note as input. For the sake of simplicity, we consider the tonality of input melody is known, or found using a key-finding algorithm. The components of the Bayesian are described more formally below.

Input process

The input process models sensory information, or feature values x_t, which the brain translates from auditory stimulus. The auditory stimulus are derived from various sequences of notes encoded as MIDI files. The auditory stimuli are two groups of melodies used for two behavioral experiments: first group consists of 40 real melodies with two sources of unpredictability, whereas the second group consists of 27 artificial synchronous sequence generated with variable degrees of predictability [6].

Generative models

Perception in the Bayesian terminology is the inversion of generative model used to infer causes from observations. For internal representation of each note, we attribute a generative model that is the probability of observing input when a certain cause (input property) is in effect. These generative models are adapted to infer the properties of input. A generative model of an abstracted auditory input, for each perceived pitch, is defined as a Gaussian density:

= ($\Delta t \sigma^{*2}$)

Where μ^* stands for the mean and is the internal uncertainty for the sensory input. These parameters are

calculated based on a big data set of folk melodies [6].

Bayesian inference

Bayesian inference is a method of performing computations that lead to posterior belief over the internal representation of a musical pitch given the sensory observations. Since the auditory stimuli are melodies presented as a sequence of musical notes, the posterior belief is demonstrated as where = $\{x_1, x_2, \ldots, x_t\}$ denotes the sequence of observed musical notes. The posterior belief is then computed through applying Bayesian inference in a recursive manner:

$$p(N_i|x_1) = \frac{p(x_t|N_i)p(N_i)}{\sum_{j=1}^{M} p(x_1|N_j)p(N_j)}$$

$$p(N_i|X_{1:t}) = \frac{p(x_t|N_i)p(N_i|X_{1:t-1})}{\sum_{j=1}^{M} p(x_t|N_j)p(N_j|X_{1:t-1})}$$

Where M is the number of possible musical notes (including the rest note). In other words, the equations above indicate how the belief is integrated through time using the posterior belief on the last observed sequence (excluding the last observed note).

Music expectation

Finally, musical expectations in the Bayesian model are based on the posterior belief over the sequence of observed musical notes $p(N_i, X_{1:t})$

$$\max_i p(N_i|X_{1:t})$$

We are also willing to remark that building an active inference model of local monophonic music has an extra advantage that is, revealing an interesting difference between the two structural approaches that must be properly taken into account. In the western music, the epistemic offering of music fits in well with decreasing in top-down signals and increasing in bottom up signals from higher down to lower levels. Conversely, in monophonic music containing irregularities that are more likely to be detected, we will be dealing with increasing top down signals versus decreasing bottom up signals (See figure 1). Interestingly, a set of beliefs in a model of learning monophonic music will be formed and fostered as a knowledge network in the same term detailed in active inference models of selective visual attention [17,18] containing the diagnostic features that will facilitate the inference process.

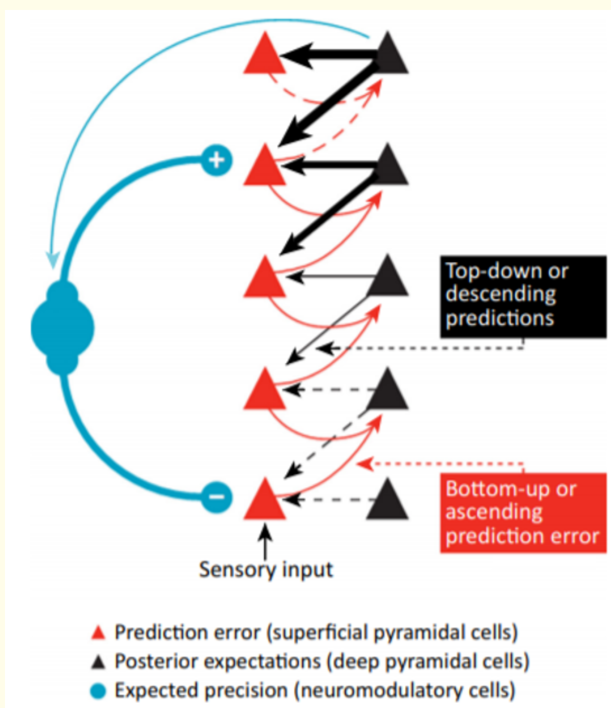

Figure 1: *Adapted from (Kolesh et al. 2019) shows the hierarchy of active inference system with increasing top down signals at higher levels and increasing bottom up signals at lower levels. In folk monophonic music, because of the existence of regular and frequent similar patterns, the tendency to resolve the uncertainty during listening to a melody is much faster than the western music.*

Simulation and experiment

The data are collected in a previous study [6]. These data sets are being used to study the dynamic ratings of melodic predictability in humans. In order to perform the simulation and generate the responses using the Bayesian model, we need to use discrete time-steps with the length of 100 ms. The prior probabilities are initiated using the calculations on ground-truth melody data-sets and also principles of music theory.

Discussion and Conclusion

The proposed Bayesian model provides a biologically inspired account for musical expectation, which is based on generative models, internal representation of the sensory input, and Bayesian inference. Essentially, the model, with efficient parameter setting, should be able to capture the dynamics of ratings of melodic expectation in humans.

Moreover, our results may contribute to models describing neural theories underlying music expectation. It has been widely known that humans perform Bayesian inference in a near-optimal fashion and this is performed through the representation of probability distributions and combining these distributions using Bayes's rule [12]. Furthermore, evidence about tonotopic organization of pitch in human auditory cortex might suggest that there are separate populations of neurons in the brain encoding the features of auditory stimuli for different pitches [19]. On the other hand, a recent study [20] shows that pitch is represented in the brain through the interaction of different cortical levels such as primary auditory cortex and the area Heschl's gyrus which is consistent with the mechanisms of predictive coding and Bayesian theories of the nervous system. Our model can shed some light on how neural mechanisms underlying music expectation can be explained by Bayesian inference and can be embedded in the predictive coding framework.

We have so far indicated the possibility of building up such predictive model that would explain how the brain comes to tackle and improve our expectations during listening to a monophonic melody. We used a set of beliefs elicited over the last observed sequence. Furthermore, we also argued that posterior perception of a monophonic melody would take place along with finding some regular patterns throughout the melody. These regular patterns include a wide range of melody progression. For example, they could appear as a series of notes, which a melody frequently revolves around or moves back to it. In some local music structures like Persian music, they have long been named and theoretically identified as Shahed (observer) or Ist (full stop) notes. This duality could also identify and recognize most of the non-western music frameworks. Therefore, using predictive coding to recognize these regular patterns will highly likely to lead us to faster improving the performance of expectations task during our perception of a melody. It also implies the possibility of the representation of a probabilistic model of a melody based on these regular patterns, which could also engage in updating generative models. Furthermore, using these regular patterns could quickly cast out irrelevant prior knowledge. Studying these assumptions could be very fruitful especially when we are dealing with massive and error-prone multi-layer Markov chains. The same inhibitory-excitatory relationship could be found in other Bayesian-based models such as the predictive coding model of selective visual attention [17]. On the other hand, we have not yet fully attribute the whole process of active inference with fitting neural correlates. Hence, finding the changes in brain using imaging techniques might unfold the neurochemical processes in the brain. We guess that in addition to dopamine thought to depict prediction error, the complete neural correlates of melody expectations might involve the thalamocortical recurrent loops beginning from cortex passing across the basal ganglia and finally ending to the same cortical area. Unlike visual template learning, audio template learning has yet to be fully accounted. It might be interesting to know whether music perception could also rest on a few templates and therefore fall upon the information integration type of template learning systems. We could also set into question whether similar to visual perception the brain uses different template learning systems in perception of different elements of music such as rhythm, melody, chords and so on. These issues as such could point out to the direction of future studies [21].

Bibliography

1. Rohrmeier MA and Koelsch S. "Predictive information processing in music cognition. A critical review". *International Journal of Psychophysiology* 83 (2012): 164-175.

2. Dennett D. "Consciousness Explained". Boston, MA: Little, Brown (1991).

3. Pearce MT Wiggins GA. "The role of expectation and probabilistic learning in auditory boundary perception: A model comparison". *Perception* 39 (2010): 1367-1391.

4. Meyer LB. "Meaning in music and information theory". *Journal of Aesthetics and Art Criticism* 15.4 (1957): 412-424.

5. Pearce MT and Wiggins GA. "Expectation in melody: The influence of context and learning". *Music Perception* 23 (2006): 377-405.

6. Eerola T., *et al*. "Real-time prediction of melodies: Continuous predictability judgments and dynamic models". In Proceedings of the 7th International Conference on Music Perception and cognition Sydney". Australia: Causal Productions (2002): 473-476.

7. Abdallah SA and Plumbley MD. "Polyphonic transcription by non-negative spars coding of power spectra". In Proceedings of the 5th International Conference on Music Information Retrieval (2004).

8. Krumhansl CL., *et al*. "Expectation in Finnish spiritual folk hymns: Convergence of statistical, behavioral, and computational approaches". *Music Perception* 17 (1999):151-195.

9. Narmour E. "The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model". Chicago: University of Chicago Press (1990).

10. Eerola T., *et al*. "An investigation of pre-schoolers' corporeal synchronization with music". In Baroni, M., Addessi, A. R., Caterina, R. and Costa, M. ICMPC and ESCOM (2006): 472-476.

11. Kenji Doya., et al. "Bayesian Brain. Probabilistic Approaches to Neural Coding". The MIT Press; 1 edition (2007).

12. Pouget A., *et al*. "Probabilistic brains: knowns and unknowns". *Nature Neuroscience* 16 (2013): 1170-1178.

13. Temperly D. "A Bayesian Approach to Key-Finding". In: Proceeding ICMAI '02 Proceedings of the Second International Conference on Music and Artificial Intelligence (2002):195-206.

14. Temperly D. "Music and Probability". Cambridge, Massachusetts: MIT Press (2007).

15. Saffran JR., *et al*. "Statistical learning by 8-month-old infants". *Science* 274 (1996): 1926-1928.

16. Vuust P and Witek MAG. "Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music". *Frontiers in Psychology* 5 (2014): 1111.

17. Abadi AK., *et al*. "Excitatory versus inhibitory feedback in Bayesian formulations of scene construction". *Journal of the Royal Society Interface* 16.154 (2019): 20180344.

18. Yahya K., *et al*. "12th Biannual Conference of the German Cognitive Science Society". *Cognitive Processing* 15.1 (2014): 107-107.

19. Saenz M and Langers DRM. "Tonotopic mapping in human auditory cortex". *Hearing Research* (2013).

20. Kumar S., *et al*. "Predictive Coding and Pitch Processing in the Auditory Cortex". *Journal of Cognitive Neuroscience* 23 (2011): 3084-3094.

21. Wei JM., *et al*. "Bayesian inference with probabilistic population codes". *Nature Neuroscience* 9.11 (2006): 1432-1438.

Volume 12 Issue 12 December 2020
©All rights reserved by Keyvan Yahya.