Available online at www.sciencedirect.com

ScienceDirect

Artificial Intelligence 171 (2007) 1118–1121

Artificial Intelligence

www.elsevier.com/locate/artint

Book review

David Temperley, Music and Probability, MIT Press, 2007.

Roger B. Dannenberg

School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 15213, USA

Available online 10 October 2007

Connections between music and probability have been around for a long time. Music theorists and composers have used probabilistic systems for music generation since the early days of computers (not to mention Mozart's Musikalishes Würfelspiel, a system for composing by rolling dice). Markov models have been used by many to generate music after training transition and output probabilities on various collections of machine-readable music. More recently, we have seen extensive use of probabilistic methods in music analysis and information retrieval. David Temperley's *Music and Probability* offers an introduction to probabilistic analysis of music, with an emphasis on Bayesian methods.

Some of the motivation for this book comes from Temperley's earlier work, culminating in his 2004 book, *The Cognition of Basic Musical Structures* [2]. In this previous work, algorithms are developed to solve various music analysis tasks such as identifying keys and chords, finding beats in music performances, and grouping notes into phrases. This was accomplished using preference rules to describe the desired outcomes and optimization algorithms to find solutions that best satisfy the preferences. Two examples of preference rules are that beat locations should coincide with note onsets, and beats should be about evenly spaced.

One of the problems with the preference rules approach is that it relies on a number of weights and hand-tuned parameters. So even though the results have been quite good, the use of many arbitrary constants is troubling. This sets the stage for probabilistic methods, and Bayesian methods in particular. In Bayesian terms, the goal of these music analysis tasks is to determine an underlying *structure* of the music, such as key, meter, or harmony, given the observable *surface* of the music, which is for the most part a list of notes, their pitches, onset times, and durations.

Taking the standard Bayesian approach, Temperley presents generative models of music that can be used to estimate the probability of observing the surface given a structure, or $P(\text{surface} \mid \text{structure})$. These generative models tend to be simple to allow the estimation of parameters, and not rich enough to really generate any convincing music. For example, if the structure is a key (such as C major), the model is just the probability of observing different pitches given the key (in C major, an F is fairly likely, but F# is unlikely).

Continuing with the Bayesian approach, Bayes' rule tells us that $P(\text{structure} \mid \text{surface}) \propto P(\text{surface} \mid \text{structure})$ $P(\text{structure})$. This tells how to estimate the structure (e.g. determine the key) from the actual music (the surface), which is the essence of music analysis. In most cases, the solution is to simply estimate $P(\text{surface} \mid \text{structure})$ for all possible structures and choose the structure with the highest probability (or maximum likelihood).

Temperley does an excellent job of introducing the basic concepts of probability and Bayesian methods. For experts, this will seem elementary, but the writing is direct and clear. I would certainly recommend the book to students interested in music analysis by computer or to anyone coming into the field. For AI and machine learning researchers with some knowledge and interest in music, this book takes a very elegant approach to music analysis and offers interesting insight into the nature of both music perception and composition from a probabilistic perspective. The book generally assumes a basic knowledge of music theory: you will need to be familiar with keys, chords, scales, and a bit of rhythmic notation. Some of the discussion sections refer to more advanced music theory. So, for example, those

E-mail address: rbd@cs.cmu.edu.

doi:10.1016/j.artint.2007.10.002

familiar with common devices for modulation will better appreciate the discussion of pivot chords, but it should be possible to follow almost everything without expert knowledge of music theory.

The book is commendable in that the examples address realistic music analysis problems. For example, detecting beats and meter is a long-standing and difficult problem. Key finding was a Music Information Retrieval Evaluation eXchange (MIREX) evaluation task in 2005. (http://www.music-ir.org/mirex2005) (MIREX is an opportunity for researchers working on audio and symbolic music analysis to compare different algorithms and systems on databases that are shared within the music information retrieval community.) Temperley's examples are tested using standard databases such as the Essen Folk Song Collection, a corpus of over 6000 European folk songs, and expertly labeled examples from a harmony textbook by Kostka and Payne. (1995. *Tonal Harmony*, New York: McGraw-Hill). Although many of the examples in the book have appeared as conference publications, this book provides much more introduction, interesting discussion, and speculation than the corresponding papers.

Music and Probability introduces some interesting new ideas for probabilistic reasoning, thinking about communication, and music theory. Two ideas on probabilistic reasoning struck me as especially interesting. The first is a way to think about tonal music generation. The concept of a pitch class profile or histogram of pitch frequencies has been around for a long time. In essence, if you play in C major, white keys are more common than black keys. This concept can be refined, quantified, and used to generate music in a given key or to estimate the key, given a list of note pitches. Another concept used mostly in algorithmic composition is the distribution of pitch intervals. It is known that small pitch intervals are more common than large ones, and composers have often used "$1/f$ noise," Brownian motion, and related concepts to create pleasing pitch sequences. Temperley adds a new distribution called the *range profile*. The idea is that melodies typically span an octave or so, and tend to favor notes toward the middle of that pitch range. Different melodies have different ranges, perhaps because different voices have different ranges, so the range profile is conditioned by a central pitch, which has its own probability distribution.

Temperley shows how all of these distributions and profiles can be integrated in a probabilistic fashion. Although he uses it mainly for key finding, this model could be interesting for algorithmic composition, for creating and exploring new musical styles, and for studies of music style and expectation. In fact, a chapter is devoted to music expectation and error detection. How is it that we often know someone plays a wrong note even if we have no prior knowledge of the music? Temperley suggests that we have general models of music that tell us some notes are likely but others are so unlikely that they must be mistakes.

The second idea on probabilistic reasoning that I found especially interesting is Temperley's algorithm for rhythmic parsing, which is based on dynamic programming. Dynamic programming has been used often for music similarity (e.g. comparing two melodies) and music score following (tracking a live performance by matching it to a sequence of notes in a score). One of the problems of dynamic programming is that it is hard to incorporate duration constraints. Often, we would like to say "these sequences should match while maintaining a fairly steady tempo." A non-musical example is "I want to match a sequence of landmark observations to my map, assuming that my robot moves at a fairly constant velocity." Hidden Markov models are often suggested for this sort of task, but HMMs have a limited ability to incorporate models of tempo, velocity, or acceleration because of the Markov assumption and because states are discrete. A solution offered by Lorin Grubb [1] is to represent position by a continuous probability density function, which allows one to model velocity or tempo as a continuous or real number. In Temperley's solution, time is quantized into "pips" of 50 ms, which is fine enough for at least a rough estimate of tempo, and the dynamic programming solution takes into account the probability distribution of beat-to-beat tempo change. This allows the solution to prefer a steady tempo. I think this could have applications beyond the rhythmic parsing model where it is used in the book. A similar model was used in Temperley's work with preference rules, but the representation is subtle and its implications are easily overlooked.

While the first part of the book is fairly scientific and experimental—models, hypotheses, experiments, results, conclusions—the later chapters are more speculative. In some cases, this is disappointing. I was hoping for some objective experiments to resolve some of the very interesting questions. For example, Section 9.3, "Testing Schenkerian Theory," describes in basic terms the influential theory of Heinrich Schenker. Temperley suggests that probabilistic models of this theory could be used to test it. If the theory is valid, it should make predictions about musical surface, and probabilistic methods such as cross-entropy should enable us to measure whether these predictions are meaningful. However, a computational or probabilistic model of Schenkerian theory does not exist, so the section is mostly a description of the difficulties of this problem.

The following chapter, "Communicative Pressure" offers some interesting insights into the nature of music perception and music composition. Music researchers usually think of music as something that is given in a fixed form such as the corpus of Bach chorales or Mozart piano sonatas. The challenge is to discover the structure, style, or "language" of these collections. But one could also consider that the composers of these works intended to communicate something. Composers must work within a framework where there is some "structure" to communicate via a "surface" and where listeners must infer the structure from the surface. Therefore, the whole nature of music can be reconstructed not as a given language within which composers and listeners work, but as a *manufactured* language that evolves in ways that are constrained by the need for listeners to infer structure from surface. There are some interesting examples dealing with rhythm: in jazz, rhythms tend to be complex, with many notes occurring off the beat, but the tempo tends to be very steady. Presumably, the steady tempo enables jazz listeners to parse the complex rhythms. In contrast, romantic period music tends to be played with wide, expressive tempo variations, but the music tends to be structured without much syncopation or other rhythmic complexity, again enabling listeners to parse the music. In this case, the complexity comes in the form of tempo variations. One can also take into account performance practice. What does the performer have to do to communicate structure? Or alternatively, what (new) performance structure can the performer add to the work without destroying the listener's ability to perceive both the intentions of the composer and the expression of the performer? The implied circularity of all this is enough to make your head swim, but I suspect these ideas will become increasingly important in music theory and perception research.

In the realm of music theory, Temperley's models immediately offer new tools. Chapter 7 includes an analysis of the opening of a Chopin Mazurka in terms of probability scores for each measure. The log probability plot dips and then rises (see Fig. 1). Could this be an expression of dramatic tension and release communicated through the probability of the music surface? In another example, Temperley shows that the dominant to tonic cadence in major keys is less tonally ambiguous if the dominant seventh chord is used, but in minor keys, the dominant without the seventh is unambiguous. Therefore, you might expect composers to use the dominant seventh chord more in major keys than minor keys, assuming the composers' goal is to make these cadences strong, unambiguous musical statements. Statistics on Haydn, Mozart, and Beethoven piano sonatas show that this is indeed the case. The tendency to use the dominant seventh is stronger in major keys as predicted. Another example sheds some light on why the subdominant (IV) chord often precedes the dominant (V) before resolving to the tonic (I) chord. Tonic ambiguity seems to play a role and offers an explanation to a question that has long challenged theorists.

Music and Probability offers a great introduction to an important area of music research. The book combines clear and detailed explanations of basic concepts with some interesting speculation about applications and implications of the probabilistic approach. As music theory, this work suggests new, quantitative approaches to thinking about music analysis. Even for non-musician researchers who might encounter probabilistic methods in other domains, the suggested links between music perception, music composition, and probability are fascinating to consider.

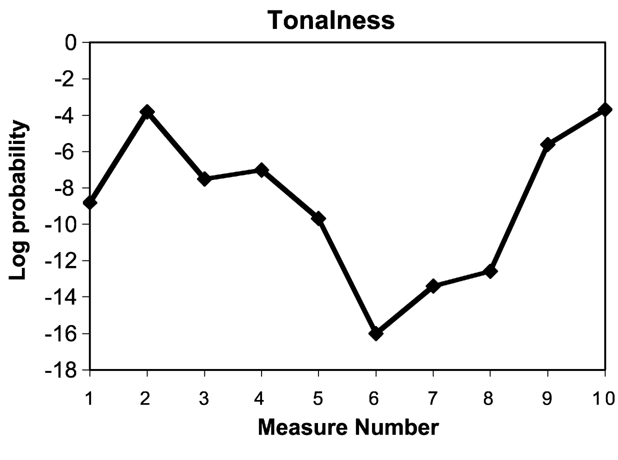

Fig. 1. The probability of pitch sets in the first 10 measures of Chopin, Mazurka Op. 6, No. 1, showing a sharp decrease in measures 5–8. The probability can be considered a measure of how much each measure is "tonal" or consistent with the local key (F# minor). (Adapted from Temperley, Fig. 7.7, p. 113.)

References

[1] L.V. Grubb, A probabilistic method of tracking a vocalist, Ph.D. Thesis, Report CMU-CS-98-166, Carnegie Mellon University Computer Science Department, Pittsburgh, 1998.
[2] D. Temperley, The Cognition of Basic Musical Structures, MIT Press, Cambridge, MA, 2004.