Review of *The Cognition of Basic Musical Structures*[1]

Roger B. Dannenberg

Machine models of music understanding now form an interesting and diverse literature. Problems such as key finding, beat tracking, segmentation into phrases, and harmonic analysis have been tackled using a variety of techniques, ranging from ad-hoc programs to machine learning. Given this background, it is a bold step to embark upon a program to revisit a range of music understanding tasks. David Temperley has done just that. His work is especially interesting because he consistently applies a general approach to a variety of music problems. The approach, essentially, is to search for solutions that optimize a set of preference rules. Temperley shows that this approach is widely applicable to music problems, relatively simple, and generally quite effective. One could argue that a better approach might exist for any single problem, or that Temperley's simple models could be extended to incorporate more musical knowledge. However, it is striking how far Temperley was able to take this approach. By stating assumptions clearly, sticking with very general models, and offering implementations, this work forms an important benchmark against which other work can be compared.

Readers who were frustrated by the lack of formal, specific procedures in Lerdahl and Jackendoff's *Generative Theory of Tonal Music* will be pleased to find algorithms available in source form and objective performance results. (This is not to say that this work is an implementation of GTTM, although it does take inspiration from that work.)

In all, six systems are developed to solve the following tasks: (1) Metrical Structure, the identification of beats at multiple levels, e.g. eighths, quarters, measures, and measure groups, (2) Melodic Phrase Structure, in which notes are grouped into phrases, (3) Contrapuntal Structures, which separate polyphonic structures into voices, (4) Pitch Spelling, a precursor to harmonic analysis and key finding, (5) Harmonic Structure, i.e. harmonic analysis, and (6) Key Structure, the identification of keys and points of key change. These six preference rule systems follow an introductory chapter and constitute Part I of the book.

The preference rule system for Metrical Structure illustrates the approach, which is more-or-less repeated for each other system. Temperley first introduces a set of preference rules. For Metrical Structure, the first rule prefers to align strong beats with note onsets and the second prefers to align strong beats with longer notes. The third favors periodic beats by saying that beats should be evenly spaced. In addition to these preference rules, there are "well-formedness rules" that place constraints on the form of any structure. For Metrical Structures, beats exist at multiple levels, beats at higher levels also exist at lower levels, and each lower level divides the next higher level into two or three divisions. Duple relationships are preferred over triple divisions. Although not implemented in a general way because of interdependencies with the calculation of melodic grouping structures, there is a preference to locate strong beats at the beginning of groups.

Using just these rules, it is possible to search for a metrical structure that best satisfies all preferences. To do this, one must decide how to weigh the various preferences when they are in conflict. Do you hold a very steady tempo to satisfy the even-spacing rule, or do you adjust the tempo freely to align more beats with note onsets to satisfy the first alignment rule? Temperley addresses this issue by using numerical scores that correspond to the degree to which a preference rule is satisfied. Then, an overall score is computed as a weighted sum of these preference rule scores. A nice by-product of this additive approach to scoring is that optimal structures can be found efficiently using dynamic programming. Thus, the best

[1] Published as: Roger B. Dannenberg, ``The Cognition of Basic Musical Structures by David Temperley." (Book Review) *Music Perception* 20(3), (spring 2003), pp. 327-331.

solution can take into account non-local phenomena; the optimal solution for the whole may not be optimal for any local region of a few measures.

While this approach is demonstrated to be very powerful, it should be kept in mind that it does not capture interdependence between preference rules and between structures. For example, melodic phrases and harmony may contribute significantly to the perception of meter and rhythm, but these factors are not considered in the preference rule systems. Within a system, there may be interactions. For example, the steadiness of tempo at the beginning of a musical work might cause listeners to weigh the "steady tempo" preference rule more heavily later in the work. These sorts of interactions are not modeled, partly to keep the systems simple, and partly because the dynamic programming optimization technique cannot handle rule systems where interactions are not additive.

The other preference rule systems are similar in structure, each with a set of rules, a set of scoring functions, and an algorithm that finds the structure with the best score. After presenting each preference rule system, Temperley evaluates each one, most often using excerpts from a music theory workbook by Kostka and Payne. For the metrical structure system, additional tests used unquantized input from expert piano performances (253 measures).

Perhaps the best information in the book is the discussion of each rule system at the end of the chapter. Here, Temperley gives an insightful assessment of the successes and failures discovered in the evaluation, limitations of the system, and possible improvements. Especially interesting are the examples pulled from the tonal music literature, frequently used to raise an issue or explain a point.

Of particular interest is Temperley's model for pitch spelling. Pitch spelling is essentially the determination of whether a pitch is labeled with a sharp (e.g. C#) or flat (Db). This is not a commonly addressed problem, and most researchers would consider this to be a part of or a byproduct of key identification. In contrast, Temperley prefers to treat this problem first, and then use the solution to inform his Harmonic Structure and Key Structure systems.

Temperley introduces a model called "tonal-pitch-class representation," where enharmonics such as (C#, Db) are considered to be in separate pitch classes. As a result, the familiar circle of fifths becomes a line of fifths, which descends by fifths from left to right, resulting in sharp spellings on the left and flats on the right. One significant advantage of the line of fifths is that distances and average positions are well defined here, which is not the case when the line is mapped onto a circle. Whether or not the line of fifths has a psychological reality and how that reality is manifest is an interesting question.

Part II of the book explores the imperfections of the preference rule concepts introduced in Part I. Chapter 8, "Revision, Ambiguity, and Expectation" considers music listening as a *diachronic* ("across time") process. Dynamic programming algorithms have the property that they build solutions incrementally. Since all of Temperley's algorithms consider music events in temporal order, they are able to present a "best so far" interpretation of the data. This forms the basis for considering various listening phenomena. No comprehensive model of these phenomena is offered. Instead, Temperley reviews the literature, discusses the issues, and merely suggests how his preference rule systems might be viewed as consistent with listening phenomena.

Chapter 9 considers "Meter, Harmony, and Tonality in Rock." The preference rules introduced in Part I were developed for common-practice music and are not directly applicable to rock music. Temperley argues that rock contains significant differences that can be accommodated by a preference rule system, and he offers an interesting discussion. In particular, rock contains syncopations where important notes tend to anticipate strong beats rather than align with them. One solution offered is to consider the "surface" rhythms to be a transformation of a "deep" structure where the melody and rhythm are more closely aligned. Temperley offers other interesting insights into rock harmonic structures, the use of modes, and the use of pentatonic scales. Unfortunately, these ideas are not backed up by specific

implementations, so this is an area where much more work is needed before drawing too many conclusions.

Similarly, Chapter 10, "Meter and Grouping in African Music" explores how preference rule systems might apply to African music, using ethnomusicological studies as a source of information about African musical practice. Temperley is careful to note that "African Music" encompasses many styles and cultures. He is concerned with "general features of African rhythm," and uses studies selectively. Overall, this chapter argues that there are strong similarities between Western and African music and that preference rule systems can model the music cognition of both.

Chapter 11, "Style, Composition, and Performance," looks at the generation of music and argues that similar processes of cognition are involved in listening, composition, and performance. After all, music generation must work within the cognitive processing capabilities and expectations of listeners if music producers expect to communicate with listeners via a shared musical and cultural framework. In this chapter, the idea of style is also explored. Temperley suggests that, after analysis with a preference system, the overall score suggests how well a piece fits the style implied by the preference rules. Here, it is also suggested that we might learn rules by discovering rules and "parameter settings" that do a better job of producing high-scoring analyses. One paradox is that music seems to always violate its own rules with intention. If we were to compose music with very high analysis scores, the music would be completely boring, and even outside of "the style." Temperley suggests that composers use less-analyzable passages to create tension, which is indicated by lower analysis scores, although not necessarily to the point of ambiguity. It seems to me that if "good" music generates lower scores, then something is either wrong with the model, or the model is not really measuring style but some aspect of style, such as an abstract "channel" which is "modulated" to communicate musical meaning. Certainly, much more study is needed to resolve these issues. The chapter concludes with interesting observations and examples of how performers can use expressive timing and articulation to make subtle changes in analysis scores, thereby making analysis more or less ambiguous for musical effect.

Chapter 12 is titled "Functions of the Infrastructures." By "infrastructures," Temperley means the basic conventions of common practice music as expressed in the preference rule systems. This chapter explores the question of why such an infrastructure exists. The infrastructure itself is not particularly meaningful or enjoyable by itself, but it does give rise to a rich set of expressive possibilities in music. Temperley reviews a number of concepts in music perception that fall outside those covered by his preference rules. In other words, this chapter considers structure in music, such as motivic structure, that rides above the infrastructure addressed in Part I.

An interesting section of this chapter considers the effect of making various minor changes to the opening measures of the minuet from Beethoven, Sonata Op. 10, No. 3. Ten "recompositions" are offered an evaluated in terms of harmonic, rhythmic, and phrase structure. This section illustrates connections between compositional decisions and the preference rule systems, but of course, the preference rules only go so far in explaining Beethoven's choices.

Overall, *The Cognition of Basic Musical Structures*, is brilliantly conceived and executed. It advocates a clear theory, exploring many facets, but stops short of making extravagant claims or unwarranted generalizations. One could write pages on the limitations and problems of the preference rule model and implementation details, but Temperley has already done that with great insight and honesty. Perhaps the biggest question is whether this is really a book on cognition, or is this "computational music theory?" Temperley is careful to separate the general framework of preference rules from the specific implementation in terms of data structures, optimization, and so forth. Certainly, the implementations tell us little about how, specifically, the mind might actually process musical information. However, Temperley argues that it is plausible that preference rule systems are somehow incorporated in our music processing. At the outset, Temperley suggests that preference rule systems offer "promising hypotheses about the cognition of basic musical structures." In my opinion, the word "cognition" in the title is

justified and the content is bound to stimulate a great deal of thought and additional research on music perception and cognition.