A Bayesian Approach to Key-Finding

David Temperley

Department of Music Theory
Eastman School of Music
26 Gibbs St.
Rochester, NY

dtemp@theory.esm.rochester.edu

Abstract. The key-profile model (originally proposed by Krumhansl and Schmuckler, and modified by Temperley) has proven to be a highly successful approach to key-finding. It appears that the key-profile model can be reinterpreted, with a few small modifications, as a Bayesian probabilistic model. This move sheds interesting light on a number of issues, including the psychological motivation for the key-profile model, other aspects of musical cognition such as metrical analysis, and issues such as ambiguity and expectation.

1. Introduction

How do listeners determine the key of a piece? This question has been the subject of considerable attention in recent years. A number of computational models have been proposed, reflecting a variety of approaches ([11], [1], [2], [9], [16]). One important proposal has been the key-profile model of Krumhansl and Schmuckler [8]. In this model, the distribution of pitch-classes in a piece is compared with an ideal distribution or "key-profile" for each key; the key whose profile best matches the pitch-class distribution of the piece is the preferred key. Elsewhere ([13], [14]) I have proposed a modified version of the original Krumhansl-Schmuckler (hereafter K-S) model which yields significantly improved performance; both the K-S model and my modified version will be discussed further below.

The purpose of the current paper is to examine the connection between the key-profile model of key-finding and the Bayesian method of cognitive modeling. Examination of these two approaches shows that they have a great deal in common; indeed, it could practically be argued that the key-profile model (particularly my own version of it) simply *is* a Bayesian probabilistic model. If this is true, then it is well worth noting for several reasons. First, the Bayesian perspective provides a more compelling psychological motivation for the key-profile model than anything that has been provided before. Secondly, the connection explored here may point to a broader connection between Bayesian modeling and a general approach in music cognition, the preference rule approach, which has been applied to a variety of problems in musical analysis, such as grouping, metrical analysis, harmonic analysis, and stream segregation. The Bayesian perspective may also be relevant to other aspects of

C. Anagnostopoulou et al. (Eds.): ICMAI 2002, LNAI 2445, pp. 195-206, 2002.

musical cognition, such as ambiguity and expectation. We will return to these ideas at the end of the paper.

I will begin by presenting the key-profile model, followed by a brief introduction to Bayesian modeling. I will then consider how the key-profile model might be reinterpreted within the Bayesian framework.

2. The Key-Profile Model of Key-Finding

The original key-profile model is based on a set of twelve-valued vectors, called key-profiles, representing the stability or compatibility of each pitch-class relative to each key. Key-profiles for C major and C minor are shown in Figure 1; profiles for other keys can be generated by simply shifting the values over by the appropriate number of steps. (For example, while in C major the value for C is 5.0 and the value for C# is 2.0, in C# major the value for C# would be 5.0 and the value for D would be 2.0.) These are the profiles used in my modified version of the key-profile model; they differ slightly from those used in Krumhansl and Schmuckler's original version, which were based on experimental data. (See [14] for an explanation of why these modified values were proposed.) It can be seen that the key-profiles reflect basic theoretical principles of tonal music. In the major profile, the values for the major scale are higher than those of chromatic pitches; within the major scale, values for pitches of the tonic triad are higher than for other diatonic pitches, and the value for the tonic is highest of all. The same is true of the minor profile (assuming the harmonic minor scale).

Given these profiles, the model judges the key of a piece by generating an "input vector" for the piece; this is, again, a twelve-valued vector, showing the total duration of each pitch-class in the piece. The correlation value is then calculated between each key-profile vector and the input vector. In essence, this involves taking the product of each key-profile value with the corresponding input-vector value, and summing these products.[1] The key yielding the maximum correlation value is the preferred key. Informally speaking, if the peaks in the input vector correspond with the peaks in the key-profile vector, the score for that key will be large. Figure 2a shows a simple example; for this passage, the model chooses C major as the correct key, just as it should.

My own tests of the Krumhansl-Schmuckler model revealed several problems with it, which were addressed in a modified version of the model (see [14] for details). One problem with the original model was that repeated notes appeared to carry too much weight. In a case such as Figure 2b, the repetitions of the E give a strong advantage to E major and E minor, even though C major is clearly the correct judgment. (A similar problem occurs when a pitch-class is duplicated in different octaves.) It seems that,

[1] This is a simplified version of the correlation formula, but appears to give essentially the same result. For discussion, see [14], pp. 173-6.

for a small segment of music at least, what matters most is the pitch-classes that are present, rather than how much each one is present. This problem was addressed in the following way. The model assumes some kind of segmentation of the piece which has to be provided in the input. (It works best to use segments of one to two seconds in length. In the tests reported below, I used metrical units—measures, half-measures, etc.—always choosing the smallest level of metrical unit that was longer than 1 second.) In constructing the input vector for a segment, the model simply gives each pitch-class a value of 1 if it is present in the segment and 0 if it is not. It then uses this "flat" input vector to calculate the correlations with the key-profiles. Since the input vector values are all 1 or 0, this simply amounts to adding the key-profile values for the pitch-classes that score 1 in the input vector. Consider Figure 2a; in this case, the score for C major is produced by adding the values in the C major profile for C, D, E, and F, yielding a score of 5.0 + 3.5 + 4.5 + 4.0 = 17.0. This "flat-input" approach proved to achieve substantially better results than the "weighted-input" approach of the original K-S model.

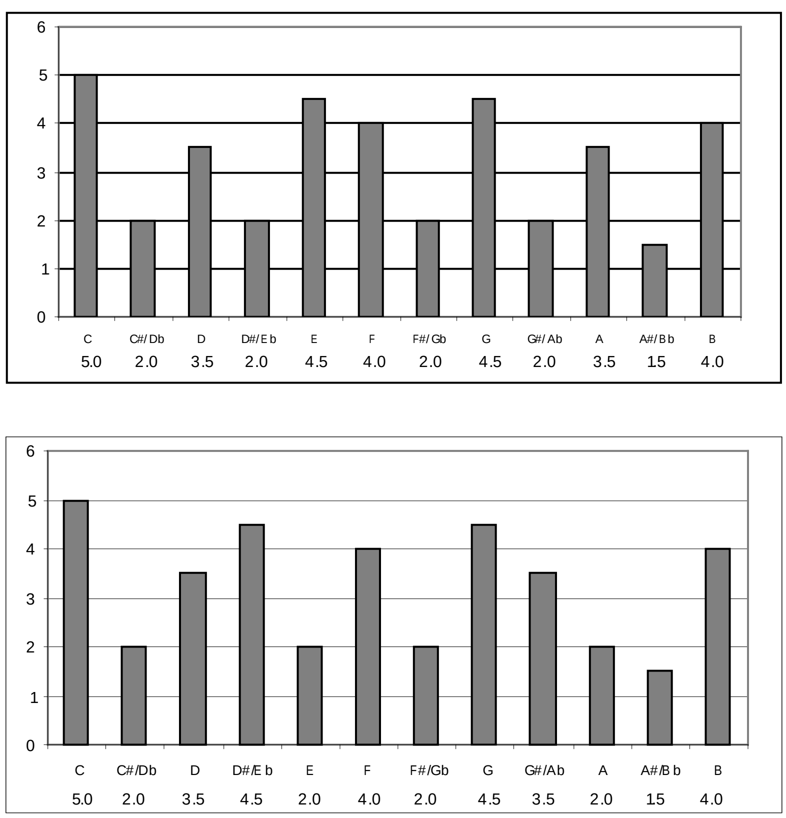

Fig. 1. Key-Profiles for C major (above) and C minor (below).

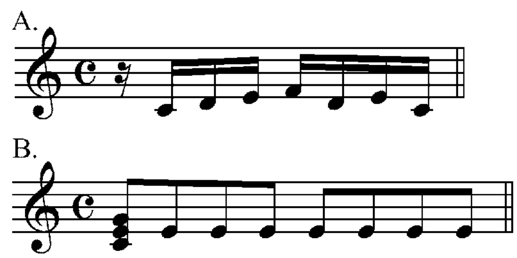

Fig. 2. (A) Bach, Invention No. 1, m. 1. (B) A problematic passage for the K-S model.

A further modification of the original key-profile model concerned modulations (changes in key). The original key-profile model is incapable of handling modulation; it simply produces a single key judgment for the entire input. Since the modified version of the model requires a segmentation of the piece in the input, one solution would be to judge the key of each segment independently; then modulations would simply emerge at points where the key of one segment differed from that of the previous one. However, this is not very satisfactory. Intuitively, key has a kind of inertia; once we are in a key, we prefer to remain in that key unless there is strong evidence to the contrary. To handle this, the model evaluates each segment independently, but imposes a "change penalty" if the key for one segment differs from the key for the previous segment. These penalties are then combined additively with the key-profile scores to choose the best key for each segment.

Once the change penalty is introduced, this means that the analysis for a particular segment is sensitive to context: the optimal analysis for one segment might depend on what precedes it. Moreover, if the model's goal is to find the highest-scoring analysis overall, the analysis of a particular segment might depend on what follows, as well. If the model is to be assured of finding the optimal analysis then, it is necessary for it to consider all possible analyses of the entire piece, where an analysis is simply a labeling of every segment with a key, and choosing the highest-scoring one overall. Since the number of possible analyses increases exponentially with the number of segments, this approach is not feasible in practice, either for a computer or for a human mind. In the computational implementation I devised for the model, dynamic programming is used to find the highest-scoring analysis without actually generating them all.

The model was tested using the workbook and instructors' manual accompanying the textbook *Tonal Harmony* by Stefan Kostka and Dorothy Payne ([7], [6]). The workbook contains a number of excerpts from pieces in the tonal repertory; the instructors' manual provides harmonic analyses done by the authors, showing keys and modulations. The model was run on 46 excerpts from the workbook, and its output was compared with the analyses in the instructors' manual. (The key-profiles were set prior to this test. However, the change penalty was adjusted to achieve optimal performance on the test.) Out of 896 segments, the program labeled 751 correctly, a rate of 83.8% correct. Inspection of the results showed that the program's errors were mainly due to three things. First, the program's rate of modulation was sometimes wrong: it sometimes modulated where the correct analysis did not (perhaps treating something only as a secondary harmony or "tonicization" instead), or vice versa. Second, the program frequently had trouble with chromatic harmonies such as augmented sixth chords. It might be desirable to build special rules into the program for handling such cases, though this has not so far been attempted. A third source of error in the model concerned pitch spelling. The original key-profile model did not distinguish between different spellings of the same pitch: for example, A*b* and G#. (I have called these categories "tonal pitch-classes" as opposed to the "neutral pitch-classes" of conventional theory.) However, it seemed likely that allowing the model to make such distinctions—so that, for example, E is more compatible with C major than Fb is—would improve the model's performance. A version of the model was developed which recognized such distinctions, and its level of performance was indeed slightly higher (87.4% on the Kostka-Payne corpus). However, if our aim is to model perception, giving the program such information could be considered cheating, since the spelling of a pitch might in some cases only be inferable by using knowledge of the key. In the tests that follow, the "neutral-pitch-class" version of the key-profiles will be used (exactly as shown in Figure 1), thus avoiding this problematic issue.

We now turn to a brief introduction to the concepts of Bayesian modeling. We will then consider the connection between Bayesian models and the key-profile model.

3. Bayesian Modeling

Communication generally involves the transmission of a message of some kind from a producer to a perceiver. As perceivers, we are often given some kind of surface representation of a message (what I will simply call a *surface*); our task is to recover the underlying content that gave rise to it—the information that the sender was trying to convey—which I will simply call a *structure*. The problem is probabilistic in the sense that a single surface might arise from many different structures. We wish to know the structure that is most probable, given a particular surface—in the conventional notation of probability, we need to determine

$$\text{argmax}_{\text{structure}}\ p(\text{structure} \mid \text{surface}) \qquad \textbf{(1)}$$

A solution to this problem lies in Bayes' rule. Bayes' rule states that, for any two events A and B, the probability of A given B can be computed from the probability of B given A, as well as the overall probabilities (known as the "prior probabilities") of A and B:

$$p(A \mid B) = \frac{p(B \mid A)\, p(A)}{p(B)} \qquad \textbf{(2)}$$

In our terms, for a given surface and a given structure:

$$p(\text{ structure} \mid \text{surface }) = \frac{p(\text{surface} \mid \text{structure})\, p(\text{structure})}{p(\text{surface})} \qquad \textbf{(3)}$$

To find the structure that maximizes the left side of the equation, we need only find the structure that maximizes the right side—and this turns out to be easier. Note, first of all, that "p(surface)"—the overall probability of a given surface—will be the same for all values of "structure". This means that it can simply be disregarded. Thus

$$\begin{aligned} &\text{argmax}_{\text{structure}}\ \text{p(structure} \mid \text{surface)} \\ &= \text{argmax}_{\text{structure}}\ \text{p(surface} \mid \text{structure) p(structure)} \end{aligned} \quad \textbf{(4)}$$

Thus, to find the most probable structure given a particular surface, we need to know—for every possible structure—the probability of the surface given the structure, and the prior probability of the structure.

The Bayesian approach has proven useful in modeling a number of perceptual and cognitive processes. One example is speech recognition. In listening to speech, we are given a sequence of phonetic units—phones—and we need to determine the sequence of words that the speaker intended. In this case, then, the sequence of phones is the surface and the sequence of words is the structure. The problem is that a single sequence of phones could result from many different words. Consider the phone sequence [ni] (this example is taken from [5]). Various words can be pronounced [ni], under certain circumstances: "new", "neat", "need", "knee", and even "the". However, not all of these words are equally likely to be pronounced [ni]; p(surface | structure) may be higher for some words than others. The words also differ in their prior probabilities—that is to say, p(structure) is higher for some words than others. Once we know p(surface | structure) and p(structure) for each word (relative to the surface [ni]), we can take the product of these values; the structure maximizing this product is the most likely structure given the surface.

4. The Key-Profile Model as a Bayesian Model

Like speech recognition or other perceptual processes, key-finding requires inferring a structure from a surface. In this case, the structure is a sequence of keys; the surface is a pattern of notes. The problem is to infer the most likely structure, given a particular surface. According to Bayes' rule, we can do this if we know—for all possible structures—the probability of the structure itself and the probability of the surface given the structure.

First consider the probability of a structure itself: a labeling of each segment with a key. (We will continue to assume a segmentation of the piece in the input.) Assume that for the initial segment of a piece, all 24 keys are equally probable. For subsequent segments, there is a high probability of remaining in the same key as the previous segment; switching to another key carries a lower probability. (We consider all key changes to be equally likely, though this may be an oversimplification.) The probability of a given key structure can then be calculated as the product of these "modulation scores" (S_m) for all segments. Let us assume, for any segment except the first, a probability of .8 of remaining in the same key as the previous segment, and a probability of .2/23 of changing to any other key. For a structure of four segments, C major - C major - C major - G major, the score will be

$$1/24 \times .8 \times .8 \times .2/23 = .000232 \quad \textbf{(5)}$$

Now, how do we calculate the probability of a surface given a structure? This problem could be solved in several different ways. I will propose one solution here, and then consider a second possibility later on. Let us suppose that, in each segment, the composer makes twelve independent decisions as to whether or not to use each pitch class. These probabilities can be expressed in a key-profile. We could base these key-profiles on actual data as to how often each pitch-class is used in segments of a particular key. Such data is shown in Table 1 for the Kostka-Payne corpus. As with the original key-profile model, the data is collapsed over all major keys and all minor keys, so that the profiles represent pitch-classes relative to keys—scale degrees, essentially. As an example, scale degree 1 (the tonic) occurs in .748 (74.8%) of segments in major keys; scale degree #4, by contrast, occurs in only .096 (9.6%) of segments. (It can be seen that the basic hierarchy of scale degrees in the profiles of Figure 1—with the tonic at the highest level, then other degrees of the tonic chord, then other diatonic degrees, then chromatic degrees—is reflected in these key-profiles as well; one odd exception is that 5 scores higher than 1 in minor.)

Table 1. The frequency of occurrence of each scale degree (relative to the current key) in the Kostka-Payne corpus, for major and minor keys. Numbers represent the proportion of segments in which the scale-degree occurs.

Scale degree	Major keys	Minor keys
1	.748	.712
#1/b2	.060	.084
2	.488	.474
#2/b3	.082	.618
3	.670	.049
4	.460	.460
#4/b5	.096	.105
5	.715	.747
#5/b6	.104	.404
6	.366	.067
#6/b7	.057	.133
7	.400	.330

The probability of a scale degree *not* occurring in a segment is, of course, 1 minus the score in the profile: for scale degree 1 in major keys, 1–.748 = .252. For a given key, the probability of a certain pitch-class set being used is then given by the product of the key-profile values—we could call these "pc scores" (S_{pc})—for all pitch-classes present in the segment (p), multiplied by the product of "absent-pc" scores ($S_{\sim pc}$) for all pitch-classes not present (~p).

$$\text{key-profile score} = \prod_{p} S_{pc} \prod_{\sim p} S_{\sim pc} \qquad \textbf{(6)}$$

To find the most probable structure given a surface, we need to calculate p(structure) p(surface | structure). This can be calculated, for an entire piece, as the product of the modulation scores (S_m) and the key-profile scores for all segments (s):

$$\text{p(structure) p(surface | structure)} = \prod_s \left(S_m \left(\prod_p S_{pc}\right)\left(\prod_{\sim p} S_{\sim pc}\right)\right) \quad \textbf{(7)}$$

A standard move in Bayesian modeling is to express such a formula in terms of logarithms. Since the function *ln x* is monotonic, two values of *ln x* will always have the same ranking of magnitude as the corresponding values of x; if our only aim is to find the maximum value of *x*, then using *ln x* instead works just as well. The logarithm for the formula above can be expressed as

$$\sum_s \left(\ln S_m + \sum_p \ln S_{pc} + \sum_{\sim p} \ln S_{\sim pc}\right) \quad \textbf{(8)}$$

Now the score is a sum of segment scores; each segment score is itself the sum of a modulation score, pc scores for present pc's, and absent-pc scores for absent pc's.

It can be seen that this is very similar to the key-profile model proposed earlier. If we pretend that the key-profile values and modulation penalties from the earlier model are really logarithms of other numbers, then the two models are virtually identical. There are some superficial differences. Since the scores in the Bayesian model are all logarithms of probabilities (numbers between 0 and 1), they will all be negative numbers. Also, the Bayesian model adds modulation scores for all segments, not just modulating segments. These are simply cosmetic differences which could be removed by scaling the values differently in the original model, without changing the results. There is one significant difference, however. In the earlier model, I simply summed the key-profile scores for the pc's that were present. In the Bayesian model, we also add "absent-pc" scores for pc's that are absent. It appears that this may be a significant difference between the two models.

A Bayesian version of the key-profile model was developed, exactly as just proposed. I used the key-profiles generated empirically from the Kostka-Payne corpus, as shown in Table 1. (Normally the corpus used for estimating the parameters should not be used for testing, but this was deemed necessary given the small amount of data available.) The only parameter to be set was the change penalty. Different values of the change penalty were tried, and the one that yielded optimal results was chosen. On the Kostka-Payne corpus, the Bayesian key-finding model achieved a correct rate of 77.1%, somewhat lower than correct rate of the earlier version (83.8%). Again, the main difference between the two models appears to be that in the second model, "absent-pc" scores are added to the key-profile scores. (The key-profiles are also different, but the key-profiles in the Bayesian model were generated directly from the test data; thus one would expect them to be optimal.) It is unclear why this would result in lesser performance for the Bayesian model. I intend to investigate this further.

One issue to consider, before continuing, is the probabilistic interpretation of key-profiles. In the model above, a key-profile is treated, essentially, as 12 independent probability functions indicating the probability of each scale degree occurring in a segment (and, thus, the probability of each pitch-class relative to each key). This approach is not ideal, since it requires a prior segmentation of the piece; there is little reason to think that such a segmentation is involved in human key-finding. An alternative approach—simpler, in some ways—would be to treat each key-profile as a single probability function (so that the 12 values of the profile would sum to 1). This function could then be used to estimate the scale-degree probabilities of an event given a certain key. Events could be treated as independent; the probability of a note sequence given a key would then be given by the product of the key-profile scores for all events—or, in logarithmic terms, the sum of scores for all events. This method resembles the "weighted-input" approach of Krumhansl and Schmuckler's original model, discussed earlier, in which the input vector reflects the number and duration of events of each pitch-class. The problem with this approach has already been noted: it tends to give excessive weight to repeated events. Initial tests of the key-profile model showed significantly better performance for the flat-input model than the weighted-input model. Thus it appears that treating the key-profiles as probability functions for independent events is unlikely to work very well. (Intuitively, in Figure 2b, the weighted-input approach assumes a generative model in which the composer decides to use C and G once, and then makes eight independent decisions to use E. But a more plausible model is that the composer decides to use certain pitch-classes, and then decides to repeat one of them.)

It is possible, however, that a more successful model could be developed based on the "weighted-input vector" idea. One way would be to assume that a musical surface is generated from a sparser, "reduced" representation of pitches, something like a "middleground" representation in a Schenkerian analysis. In such a representation, immediate repetitions of pitches such as those in Figure 2b (and also perhaps octave doublings and the like) would be removed. Possibly, a "weighted-input" model applied to such a reduction would produce better results; and it would also avoid the arbitrary segmentation required by the "flat-input" model. However, such an approach would present serious methodological problems, since it would require the middleground representation to be derived before key-finding could take place.[2]

5. Further Implications

My main aim in this paper has been simply to show that the key-profile model as proposed in [14] can be construed, with some small modifications, as a Bayesian probabilistic model. This connection is of interest for several reasons. First, the Bayesian approach provides a new perspective on the key-finding process; in a sense, it explains why the key-profile method is a sensible and effective way of determining key. The Bayesian viewpoint suggests that we can think of composition as a stochastic process in which a sequence of keys are generated and musical surfaces are then generated from these keys. (It is not *really* a stochastic process, but it can be effectively modeled in this way). From the perceiver's point of view, given knowledge of scale-degree distributions and the likelihood of key changes, the intended sequence of keys can then be recovered. In this way, the key-profile model emerges as a very natural and logical way of recovering key structure.

---

[2] It is also instructive to compare the Bayesian formula with the correlation formula used in Krumhansl and Schmuckler's original model (which, as already mentioned, differs slightly from the formula used in my model). However, space limitations prevent such a comparison here.

The relevance of Bayesian modeling to music cognition may go well beyond key-finding. The key-profile model as proposed in [14] can be viewed as a preference rule system—a model involving several rules which are used to evaluate many possible representations and select the preferred one. Preference rule models have been proposed for a variety of aspects of musical perception, including metrical analysis, grouping analysis, pitch reduction, harmonic analysis, and stream segregation ([10], [14]). In the case of the key-profile model, just two preference rules are involved:

> Key-Profile Rule: Prefer to choose a key for each segment which is compatible with the pitches of the segment (according to the key-profiles);
>
> Modulation Rule: Prefer to minimize the number of key changes.

It appears, in fact, that preference rule models generally can be construed as Bayesian models. One insight offered by the Bayesian view is that it suggests a distinction between two categories of preference rules. Some rules relate to the probability of a certain structure; we could call these “structure rules”. Others relate to the probability of a surface given a structure; we could call these “surface-to-structure rules”. In the case of the key-profile model, the Modulation Rule is a structure rule; the Key-Profile Rule is a structure-to-surface rule. Consider another example: metrical analysis. In this case, the structure is a row of beats (or a framework of levels of beats, but we will consider just a single level of beats for now), and the surface is once again a pattern of notes. The metrical model proposed in [14] (see also [15]) involves three main rules:

> Event Rule: Prefer for beats to coincide with event-onsets;
>
> Length Rule: Prefer for beats to coincide with longer events;
>
> Regularity Rule: Prefer for beats to be roughly evenly spaced.

The process of deriving a row of beats involves optimizing over these three rules: choosing the metrical level which aligns beats with as many events as possible, especially long events, while maximizing the regularity of beats. In this case then, the Regularity Rule is a structure rule, indicating the probability of structures (more regular structures are more probable); the Event Rule and Length Rule are structure-to-surface rules, indicating the probability of surface patterns given a certain structure (patterns are more probable which align events with beats, especially longer events). The model in [14] evaluates a possible analysis by assigning to it scores from each of these three rules and summing these scores. It can be seen how a model of this kind could be reconstrued as a Bayesian model, much as we have reinterpreted the key-profile model in Bayesian terms. (Cemgil et al. ([3], [4]) propose a Bayesian model of metrical analysis, somewhat along these lines.)

Aside from problems of musical information extraction, I have discussed several other applications of preference rule models [14]. One is in the representation of ambiguity. With regard to key-finding, ambiguity is a phenomenon of recognized importance. Some pitch-sets are relatively clear in their tonal implications, such as major or minor scales or triads; others are ambiguous, in that they are compatible with several different scales (such as a diminished seventh chord, C-Eb-F#-A). In the model proposed in [14], the ambiguity of a passage is captured in the numerical scores output by the model for different analyses: an ambiguous passage is one in which two or more different analyses are more or less “tied for first place”. Once again, this aspect of the key-profile model transfers straightforwardly to the Bayesian version; an ambiguous passage is one in which several analyses are more or less equally probable.

Another useful aspect of the Bayesian approach concerns the estimation of probabilities of musical surfaces. Bayesian theory tells us that the probability of a surface and a structure occurring in combination equals p(surface | structure) p (structure). It further tells us that the overall probability of a surface is equal to its probability in combination with a structure, summed over all possible structures:

$$p(\text{surface}) = \sum_{\text{structure}} p(\text{surface} \mid \text{structure})\, p(\text{structure}) \qquad \textbf{(9)}$$

Recall that the quantity p(surface | structure) p(structure) is exactly what must be generated, for all possible structures, in order to calculate the most probable structure for the surface. Thus it is not implausible to suppose that perceivers generate a measure of the probability of the surface itself, by summing these scores. In terms of key structure, a highly probable surface is one for which there is one (or perhaps more than one) highly probable key structure—one with relatively few modulations—which might, with high probability, give rise to it. Other surfaces are less probable, because no such structure exists. (Imagine a passage with a great deal of chromaticism, or rapid modulations, or both.) One application of this idea concerns expectation. It is well known that, in hearing a piece of music, listeners generally form expectations as to what is coming next; and some theorists have argued that the way music fulfills and denies expectations is an important part of its meaning and effect. It is natural to suppose that listeners’ expectations are governed by the same probabilistic models which (by hypothesis) govern their processing of musical input. With regard to key structure, we expect a continuation which is relatively probable given the probable structure—that is, one which adheres to the scale of the currently established key (though undoubtedly other constraints are involved in expectation as well). (Indeed, something along these lines has already been experimentally demonstrated; Schmuckler [12] has shown that listeners’ expectations for melodic continuations correspond closely with the key-profiles of the original K-S model.) The Bayesian approach would seem to offer a very natural method for modeling this process.[3]

Thus the Bayesian approach appears to offer a number of benefits. It provides a probabilistic basis for preference rule models generally, not merely for key-finding but for metrical analysis and other aspects of structure as well. It also suggests a way of modeling ambiguity and expectation. While many of the ideas in the paper are conjectural, the Bayesian framework seems to offer a promising avenue for modeling music cognition which deserves further exploration.

---

[3] It might also be interesting to consider the probabilities of actual pieces or passages according to the key-profile model. This would indicate the probability of a passage actually being generated by the key-profile model—reflecting, perhaps, the “tonalness” of the passage. (A similar, but less satisfactory, method of measuring this was proposed in [14].)

References

1. Bharucha, J. J.: Music Cognition and Perceptual Facilitation: A Connectionist Framework. Music Perception 5 (1987) 1-30
2. Butler, D.: Describing the Perception of Tonality in Music: A Critique of the Tonal Hierarchy Theory and a Proposal for a Theory of Intervallic Rivalry. Music Perception 6 (1989) 219-42
3. Cemgil, A. T., Desain, P., Kappen, B.: Rhythm Quantization for Transcription. Computer Music Journal 24(2) (2000) 60-76
4. Cemgil, A. T., Desain, P., Kappen, B, Honing, H.: On Tempo Tracking: Tempogram Representation and Kalman Filtering. Journal of New Music Research 29 (2000) 259-73
5. Jurafsky, D., Martin, J. H.: Speech and Language Processing. Prentice-Hall, Upper Saddle River, NJ (2000)
6. Kostka, S.: Instructor's Manual to Accompany Tonal Harmony. McGraw-Hill, New York (1995)
7. Kostka, S., Payne, D.: Workbook for Tonal Harmony. McGraw-Hill, New York (1995)
8. Krumhansl, C.: Cognitive Foundations of Musical Pitch. Oxford University Press, Oxford (1990)
9. Leman, M.: Music and Schema Theory. Springer-Verlag, Berlin (1995)
10. Lerdahl, F., Jackendoff, R.: A Generative Theory of Tonal Music. MIT Press, Cambridge MA (1983)
11. Longuet-Higgins, C., Steedman, M.: On Interpreting Bach. Machine Intelligence 6 (1971) 221-41
12. Schmuckler, M.: Expectation in Music: Investigation of Melodic and Harmonic Processes. Music Perception 7 (1989) 109-50
13. Temperley, D.: What's Key for Key? The Krumhansl-Schmuckler Key-Finding Algorithm Reconsidered. Music Perception 17 (1999) 65-100
14. Temperley, D.: The Cognition of Basic Musical Structures. MIT Press, Cambridge MA (2001)
15. Temperley, D., Sleator, D.: Modeling Meter and Harmony: A Preference-Rule Approach. Computer Music Journal 23(1) (1999) 10-27
16. Vos, P., Van Geenen, E. W.: A Parallel-Processing Key-Finding Model. Music Perception 14 (1996) 185-224