GLOBAL PROPERTIES OF EXPERT AND ALGORITHMIC HIERARCHICAL MUSIC ANALYSES

Phillip B. Kirlin
Department of Mathematics and Computer Science, Rhodes College
kirlinp@rhodes.edu

ABSTRACT

In recent years, advances in machine learning and increases in data set sizes have produced a number of viable algorithms for analyzing music in a hierarchical fashion according to the guidelines of music theory. Many of these algorithms, however, are based on techniques that rely on a series of local decisions to construct a complete music analysis, resulting in analyses that are not guaranteed to resemble ground-truth analyses in their large-scale organizational shapes or structures. In this paper, we examine a number of hierarchical music analysis data sets — drawing from Schenkerian analysis and other analytical systems based on *A Generative Theory of Tonal Music* — to study three global properties calculated from the shapes of the analyses. The major finding presented in this work is that it is possible for an algorithm that only makes local decisions to produce analyses that resemble expert analyses with regards to the three global properties in question. We also illustrate specific similarities and differences in these properties across both ground-truth and algorithmically-produced analyses.

1. INTRODUCTION

Music analysis refers to a set of techniques that can illustrate the ways in which a piece of music is constructed, composed, or organized. Many of these procedures focus on illustrating relationships between certain types of musical objects, such as *harmonic analysis*, which can show how chords and harmonies in a composition function in relation to each other, or *voice-leading analysis*, whose purpose is to illustrate the flow of a melodic line through the music. Some types of analysis are explicitly *hierarchical*, in that their purpose is to construct a hierarchy of musical objects illustrating that some objects occupy places of higher prominence in the music than others. Different kinds of hierarchical analysis have different methods for determining the relative importance of objects in the hierarchy. The most well-known of these hierarchical procedures is *Schenkerian analysis*, which organizes the notes of a composition in a hierarchy according to how much each note contributes to the overall musical structure. The procedure also illustrates how notes at one level of the hierarchy function in relation to surrounding notes at higher and lower levels. While Schenkerian analysis is the most common type of hierarchical analysis in the music theory community, there are a number of similar procedures that were developed in the music and linguistics community, specifically the procedures put forth in the book *A Generative Theory of Tonal Music* by Lerdahl and Jackendoff [10], abbreviated here as *GTTM*. GTTM applies an explicitly hierarchical view to multiple aspects of a musical composition, leading to two types of analysis known as *time-span reductions* and *prolongational reductions*, both of which are similar to Schenkerian analysis in that all three analytical methods organize the pitches of the music in hierarchies of relative importance, allowing a person to view a musical composition at multiple levels of abstraction.

None of these types of musical analysis were originally developed as computational algorithms, and so all contain certain ambiguities in their definitions. Schenkerian analysis, in particular, is known for originally being defined primarily through examples and not via a step-by-step procedure. Similarly, the two GTTM reductional analysis systems contain preference rules that can conflict with each other; the authors explicitly state that there is not enough information in GTTM to provide a "foolproof algorithm" for analyzing a composition. Nevertheless, there are now a number of automated computational systems that can construct analyses in a Schenkerian fashion [7, 11] or by following the rules of time-span or prolongational reductions in the GTTM formalism [4]. The most common computational technique underlying these systems and others like them is the *context-free grammar*: such a formalism is widely-adopted because such grammars are easily applied to musical objects, are inherently rule-based, can be adapted to work with probabilities, and admit a computationally-feasibly $O(n^3)$ parsing algorithm that can be used to find the best analysis for a piece of music.

The largest downside to context-free grammars is precisely that they are *context-free*: there are restrictions on how much musical context can be used when applying the rules of a grammar to "parse" a piece of music into an analysis. Most decisions made during the analysis process under the context-free paradigm have to be made somewhat "locally," and are unable to consider many important "global" properties that are critical to producing a highquality musical analysis. For instance, certain compositional techniques, such as identical repetitions of an arbitrarily melodic sequence, are impossible to describe satisfactorily with a context-free grammar [13]. Other common practices, such as the rules of musical form, manifest themselves as certain shapes and structures in the hierarchy [9, 10, 15], and it is unclear whether a purely context-free system could identify such structures. In this work, we show (a) evidence that context-free grammars can reproduce certain global structures in music analyses, and (b) similarities and differences in those global structures across various types of hierarchical music analysis.

2. REPRESENTATION OF HIERARCHICAL ANALYSES

For simplicity, we assume we are analyzing a monophonic sequence of notes. The two most common ways to do this in a hierarchical manner are (a) to create a hierarchy directly over the notes, and (b) to create a hierarchy over the melodic intervals *between* the notes. Each representation has distinct advantages and disadvantages [2, 12], but Schenkerian analysis is more easily represented by a hierarchy of melodic intervals. We explain this through the following example. Imagine we wish to analyze the five-note descending passage in Figure 1, which takes place over G-major harmony. Schenkerian analysis is guided by the concept of a *prolongation*: a situation where a note or pair of notes controls a musical passage even though the governing note or notes may not be sounding throughout the entire passage. For instance, in Figure 1, the sequence D–C–B contains the passing tone C, and we say the C prolongs the motion from the D to the B. The effect is similar for the notes B–A–G. However, there is another level of prolongation at work: the entire five-note span is governed by the beginning and ending notes D and G. This two-level structure can be represented by the binary tree shown in Figure 2(a), which illustrates this hierarchy of prolongations through the melodic intervals they encompass. A more succinct representation is shown in Figure 2(b): this structure is known as a *maximal outerplanar graph* or *MOP*, and illustrates the same hierarchy as the binary tree [14].

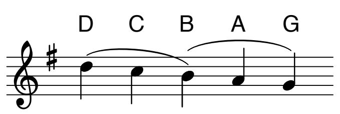

Figure 1. An arpeggiation of a G-major chord with passing tones. The slurs are a Schenkerian notation used to indicate the locations of prolongations.

A MOP is a graph representing a complete triangulation of a polygon. Like their binary tree equivalents, MOPs are rooted, but by an edge, rather than a vertex; this edge represents the most abstract level of the melodic hierarchy. Every triangle within a MOP corresponds to a hierarchical relationship among the three notes that form the triangle, with the middle note taking on a subservient role in relation

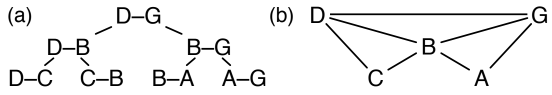

Figure 2. The prolongational hierarchy of a G-major chord with passing tones represented as (a) a tree of melodic intervals, and (b) a MOP.

to the left and right notes. It is equivalent to say that every triangle has a parent edge and two child edges, or has two parent vertices and a child vertex. Triangles closer to the root of the MOP express more abstract relationships than those farther away. Though originally developed to represent Schenkerian prolongations, we will use MOPs later in this paper to work with the GTTM reductional systems as well.

2.1 Large-Scale Organization of MOPs

Like binary trees, MOPs can be described by a variety of global attributes that are determined from their overall shape. We explore three such attributes and how common music composition and analysis practices affect these attributes.

Height: The *height* of a MOP is defined to be the number of triangles in the longest possible sequence of triangles from the root edge moving through subsequent child edges to the bottom of the MOP. It is analogous to the height of the equivalent binary tree. For a MOP with a fixed number of triangles, there are a certain range of possible heights; for instance, Figure 3 shows two MOPs with five triangles, one with a height of 5 and one with a height of 3. Because a MOP with n vertices will always contain $n - 2$ triangles, we can say the maximum height of such a MOP is $n - 2$, whereas the minimum height is $\lceil \log_2(n-1) \rceil$.

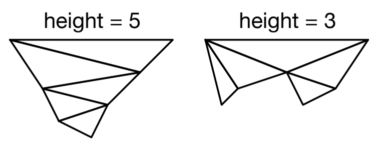

Figure 3. Two MOPs, each with five triangles, but different heights.

Investigating the heights of the MOPs that result from hierarchical analysis gives us insight into the compositional structure of the music from which the MOPs were created. MOPs with large heights result from situations where the notes of critical importance in a hierarchy are positioned towards the beginning and end of a musical passage, with importance decreasing monotonically as one moves towards the middle of the passage in question (as in the left MOP of Figure 3. In contrast, MOPs with small heights result from hierarchies where the structural importance does not increase or decrease monotonically over time during a passage, but rather rises and falls in a pattern similar to how strong and weak beats fall rhythmically within a piece.

Average Path Length: While height provides a coarse measure of the shape of a MOP, a different metric, *average path length*, provides finer-grained detail about the balancedness or skewedness of the structure [5]. Measuring the level of balancedness or unbalancedness in a tree gives an overall sense of whether the leaves of the tree are all at roughly the same level or not; when applied to a MOP, this determines whether the leaf edges are all roughly the same distance away from the root edge. Towards that end, the average (external) path length in a MOP is calculated by averaging the lengths of all the paths from the root edge to all leaf child edges. For a MOP with n vertices, the minimum average path length is very close to $\log_2 n + 1$, while the maximum average path length is $\frac{(n-1)(n+2)}{2n}$ [1, 8].

GTTM and related literature posits that music is constructed around structures that are largely balanced, that is, tending towards shorter average path lengths. One reason for this, from a prolongational standpoint, is that a melodic hierarchy expressed as a MOP illustrates patterns of tension and relaxation: each triangle represents tension rising from the left parent vertex to the child, then relaxing from the child to the right parent. Balanced MOP structures imply waves of tension and relaxation roughly on the same level, whereas unbalanced MOPs imply tension and relaxation patterns that may seem implausible to a listener: "it is most unlikely that a phrase or piece begins in utmost tension and proceeds more or less uniformly towards relaxation, or that it begins in relaxation and proceeds toward a conclusion of utmost tension." [9, 10].

Left-Right Skew: An important consideration not accounted for by the concepts of height or average path length is determining whether the MOP has more left-branching structures or more right-branching structures. To this end, we define a variant of path length by choosing to count a left-branch within a path as -1 and a right-branch as $+1$. It is clear that this metric assigns negative numbers to all MOP edges that lie on paths from the root edge with more left branches than right branches, and positive numbers to edges in the opposite situation. Using this measure of distance, we define the *left-right skew* of a MOP to be the sum of these numbers for all paths in a MOP from the root edge to a leaf edge, giving us an overall sense of whether the MOP is skewed to the left or to the right. Due to the organization of leaf edges within a MOP, a fully right-branching MOP with n vertices will achieve a left-right skew of

$$\sum_{i=-1}^{n-4} i + (n-2) = \frac{n^2}{2} - \frac{5n}{2} + 3$$

and a fully left-branching MOP will achieve a corresponding negative value.

3. FIRST EXPERIMENT

Our first experiment is intended to answer the question, "Can a fully-automated algorithm for music analysis based on context-free parsing techniques produce MOPs with global structural attributes matching those of ground-truth MOPs?" Note that we are assuming that the three global attributes — MOP height, average path length, and left-right skew — are not randomly distributed; this assumption is based on the previous work described earlier detailing that the overall shape of a hierarchical music analysis is most decidedly not random, but influenced by the way compositions are constructed and the manner in which listeners hear and interpret them.

We used the PARSEMOP system in concert with the SCHENKER41 data set to conduct this experiment. SCHENKER41 is a data set of 41 common practice period musical excerpts along with corresponding Schenkerian analyses in MOP form for each excerpt [6]. The excerpts are homogeneous: they are all in major keys, written or arranged for a keyboard instrument or voice with keyboard accompaniment, and do not modulate. The corresponding analyses of the excerpts are all derived from textbooks or other expert sources, and can be regarded as ground truth. SCHENKER41 serves as training data for the PARSEMOP machine-learning system, which learns the rules of Schenkerian analysis by inferring a probabilistic context-free grammar from patterns extracted from SCHENKER41 [7]. After training, PARSEMOP can produce MOP analyses for new, previously-unseen pieces of music.

There are three variants of PARSEMOP, which vary only in treatment of the *Urlinie*, a uniquely Schenkerian concept. According to Schenker, all tonal music compositions should have, at the most abstract level of the melodic hierarchy, one of three possible background structures. These three structures, representing Schenker's fundamental conception of melody, consist of a stepwise descent from the third, fifth, or eighth scale degree to the tonic below, and the entire melodic content of the piece serves as an elaborate prolongation of this descending melodic line. PARSEMOP-A, when trained on the SCHENKER41 corpus, does not have any *a priori* knowledge of the *Urlinie*: it does not even know that such a concept exists in music. Therefore, PARSEMOP-A produces output MOPs that usually do not contain an *Urlinie*, except if by chance. PARSEMOP-B, on the other hand, is given information about the *Urlinie* for the pieces of music it is analyzing. PARSEMOP-B produces output MOPs that always contain the correct *Urlinie* (the structure is copied from the input music). Clearly, PARSEMOP-A and PARSEMOP-B represent two opposite ends of the spectrum with regard to the *Urlinie*. PARSEMOP-C is a compromise between the two: it uses extra rules in the context-free grammar to guarantee that an *Urlinie* will be produced in the output MOP analyses, but it may not match the notes of the correct *Urlinie* exactly.

We ran the three PARSEMOP variants using leave-one-out cross-validation on each of the 41 excerpts in the SCHENKER41 corpus, leaving us with four sets of MOPs: one ground-truth, and three algorithmically produced. Because the minimum and maximum values for each of the three global MOP attributes are dependent on the number of vertices in a MOP (corresponding to the length of the musical excerpt in question), we normalized the values for the attributes as follows. MOP height was scaled to always occur between 0 and 1, with 0 corresponding to a MOP of minimum height for a given piece, and 1 corresponding to the maximum height. Average path length was similarly scaled to be between 0 and 1. Left-right skew was scaled to be between -1 and $+1$, with 0 corresponding to a perfectly left-right balanced MOP, and $-1/+1$ corresponding to maximally left- or right-branching MOPs. Finally, we compared the distribution of the attributes obtained for each PARSEMOP variant against corresponding attribute distributions calculated from the ground-truth MOPs; histograms can be seen in Figure 4.

The data illustrate a number of phenomena. Firstly, the histograms for height and average path length suggest that all the data sets, both ground-truth and algorithmically-produced, show a preference for MOPs that are a compromise between balanced and unbalanced, but tending toward balanced: very deep MOPs or MOPs with long path lengths (values close to 1) are avoided in all data sets. PARSEMOP-B and -C have higher average heights and average path lengths than PARSEMOP-A due to the presence of an *Urlinie*, which is always triangulated in a deep, unbalanced fashion. This is most evident in the PARSEMOP-B and -C plots for left-right skew: these show a dramatic left-branching structure that is almost certainly due to the *Urlinie*, especially when compared to the histogram for PARSEMOP-A.

Secondly, it is clear that in some situations, the context-free grammar formalism does a remarkably good job at preserving the overall shape of the distribution of the attributes, at least through visual inspection of the histograms. We can confirm this by computing Spearman's rank correlation coefficient ρ for each pair of data — this calculates the correlation between a list of the 41 pieces sorted by a ground-truth metric and the same metric after having been run through PARSEMOP. All pairs show positive correlation coefficients, with eight of the nine being statistically significant at the $\alpha = 0.005$ level (obtained via the Šidák correction on 0.05). In particular, rank correlation coefficients for all three PARSEMOP-B comparisons are all greater than 0.9, indicating a strong correlation.

However, a high Spearman ρ coefficient does not necessarily imply the distributions are identical. We ran two-sample two-tailed t-tests on each pair of data to determine if the means of the two data sets in question were different (the null hypothesis being that the means of each data set within a pair were identical). Two cases resulted in p-values significant at the $\alpha = 0.005$ level, indicating rejection of the null hypothesis: the height comparisons for PARSEMOP-B, and the left-right skew comparisons for PARSEMOP-C. This implies a situation where PARSEMOP-B is apparently very good preserving relative ranks of MOP heights in the data set (this rank correlation between algorithmic MOP heights and ground-truth MOP heights was revealed above), but there is also a statistically significant, though small, difference in the means of the distribution of these heights.

4. SECOND EXPERIMENT

Our first experiment suggests that there may be some bias in our calculations being introduced by the *Urlinie*, namely because it has a particular structure that is always present in the resulting analyses. This situation is further complicated by the presence of a number of short pieces of music in the SCHENKER41 data set, where, for instance, the music may consist of ten notes, five of which constitute the *Urlinie*. In a situation like this, the MOP structure is already likely determined by the locations of the notes of the *Urlinie*, and so the PARSEMOP algorithm has very little effect on the final shape of the MOP analysis. In short, we suspect that these two factors may be artificially increasing the height and average path length of the algorithmically-produced MOPs.

To address this, we replicated the first experiment but only calculated the global MOP attributes for pieces with at least 18 notes (leaving 23 pieces out of 41), hypothesizing that having more notes in the music would outweigh the effects of the *Urlinie*. The leave-one-out cross-validation step was not altered (this still used all the data). Figure 5 illustrates the new histograms compiled for this experiment. In short, these new data support our hypothesis: removing short pieces largely eliminates very deep MOPs and those with very long average path lengths.

We can again address similarities and differences using tests involving Spearman's correlation coefficient ρ and paired t-tests. All of the statistically significant results for ρ relating to PARSEMOP-B still remain: all three global MOP structure attributes calculated on the PARSEMOP-B MOPs are strongly positively rank-correlated ($\rho > 0.8$) with the ground-truth MOP attributes. There is a weaker rank correlation ($\rho \approx 0.583$) between the left-right skew attribute calculated on the PARSEMOP-C data and its ground-truth that is also statistically significant ($p < 0.005$). In contrast, the two statistical significances identified via the t-tests in the first experiment both disappear when run on only the pieces of at least 18 notes, suggesting that these association may have spurious, caused by noise in the shorter pieces.

5. THIRD EXPERIMENT

In our third experiment, we branched out from Schenkerian analysis to explore *A Generative Theory of Tonal Music*'s time-span and prolongational reductions. These are two forms of music analysis that, like Schenkerian analysis, are designed to illustrate a hierarchy among the notes of a musical composition.

Time-span reduction is introduced in GTTM as grounded in the concept of pitch stability: listeners construct pitch hierarchies based primarily on the relative consonance or dissonance of a pitch as determined by the principles of Western tonal music theory. However, pitch stability is not a sufficient criteria upon which to found a reductional system, because pitches do not occur in a vacuum, but take place over time: there are temporal and rhythmic considerations that are required. Lerdahl and

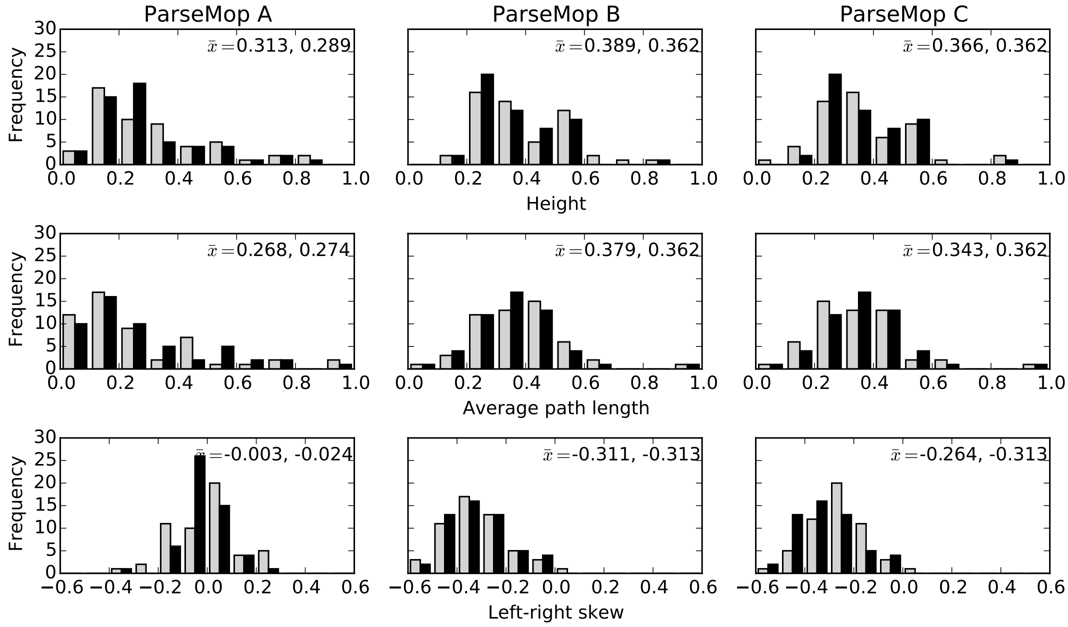

Figure 4. Histograms displaying the distribution of global MOP attributes comparing algorithmically-generated MOPs (grey bars) and ground-truth MOPs (black bars). Sample means are shown for algorithmic and truth MOPs, respectively. The ground-truth bars for PARSEMOP-A are different from -B and -C because PARSEMOP-A has no conception of the *Urlinie*, and therefore the *Urlinie* in the ground-truth MOPs is triangulated slightly differently in the training data for PARSEMOP-A.

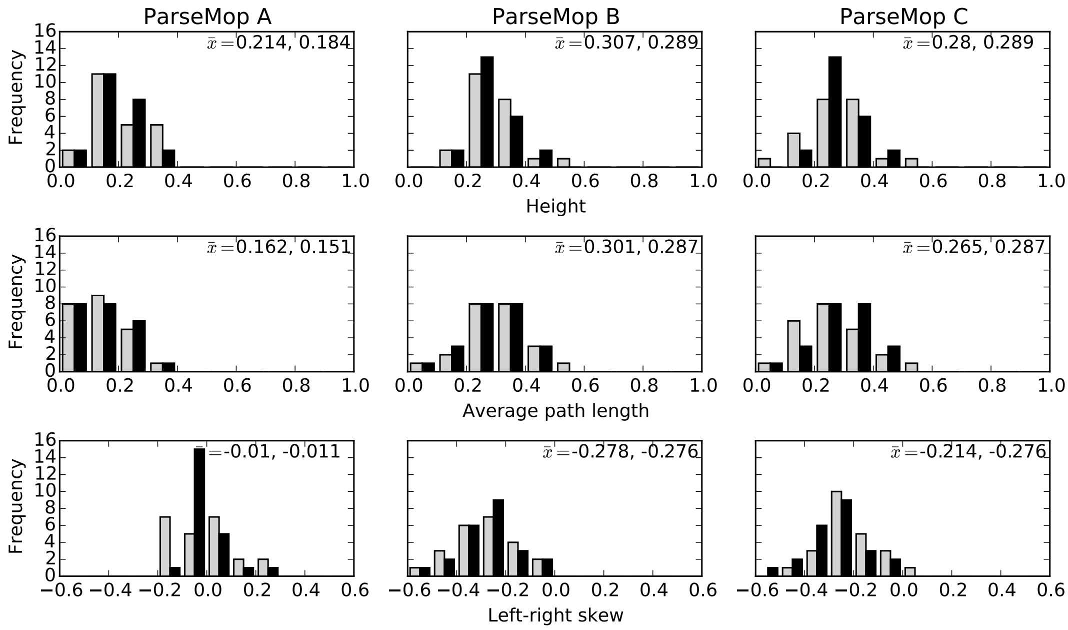

Figure 5. Histograms of the same variety as in Figure 4, but only for excerpts containing 18 or more notes.

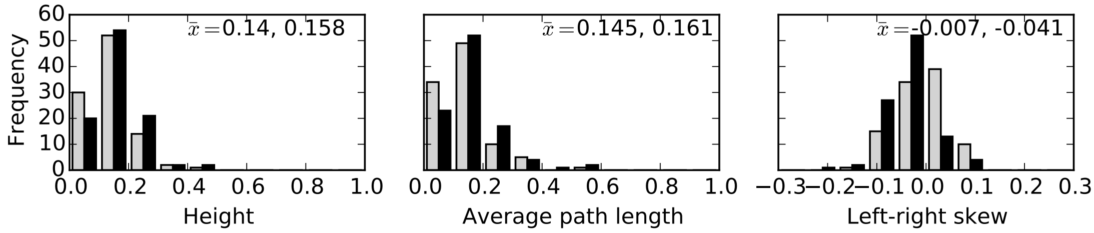

Figure 6. Histograms of the global MOP structure attributes comparing prolongational reductions (grey bars) and time-span reductions (black bars).

Jackendoff address this by basing time-span reductions on their conceptualizations of metrical and grouping structure, where metrical structure is determined from analyzing the strong and weak beats of a composition, while the grouping structure comes from a listener perceiving notes grouped into motives, phrases, themes, and larger sections.

Though similar to time-span reduction, prolongational reduction adds the concepts of tension and relaxation to the criteria that are used to form a musical hierarchy. The motivation for the need for two types of reduction is that time-span reductions cannot express some structural relationships that take place across grouping boundaries, which determine the overall form of a time-span analysis. In contrast, prolongational reductions are not tied to grouping boundaries, and therefore can represent rising and falling tension across such boundaries. The two types of trees produced by the reductional systems are often similar in branching structure at the background levels, but become more dissimilar at lower levels of the hierarchies [10].

Because time-span and prolongation reductions seem similar on the surface, it is appropriate to address their similarities and differences through a study of the three global attributes calculated for MOPs in the previous section. We perform this study by using the GTTM database developed by Hamanaka et al. [3], through their research on the feasibility of automating the analytical methods described in GTTM [4]. This database consists of 300 eight-bar excerpts of music from the common practice period, along with time-span reductions and prolongational reductions for certain subsets of the pieces. Specifically, there are 99 excerpts that include both time-span and prolongational reductions. Note that these reductions in the database were produced by a human expert, not an algorithm.

Our first task was to convert the time-span and prolongation reductions into MOPs. This is necessary because although time-span and prolongational reductions are expressed through binary trees (which are structurally equivalent to MOPs), the GTTM reductions use binary trees created over the notes of a composition, whereas MOPs are equivalent to binary trees created over melodic intervals between notes, as shown earlier in Figure 2. Therefore, we require an algorithm to convert between these two fundamentally different representations.

Time-span and prolongational reductions are represented by trees with primary and secondary branching, like that of Figure 7(a). Phase one of the conversion algorithm converts these trees into an intermediate representation: a multi-way branching tree where all children of a note are represented at the same level, as in Figure 7(b). Phase two converts this intermediate representation to a MOP by adding edges in appropriate places, as in Figure 7(c). This conversion algorithm is guaranteed to preserve all hierarchical parent-child relationships present in the original time-span or prolongational tree. It may introduce other relationships through adding additional edges, however.

Once all the time-span and prolongational reductions were converted into MOPs, we computed histograms of the MOP height, average path length, and left-right skew for both the time-span reductions and prolongational MOPs. These are shown in Figure 6. Spearman's rank coefficient test reveals positive rank-correlations between MOP height ($\rho = 0.660$), average path length ($\rho = 0.762$), and left-right skew ($\rho = 0.300$) calculated from time-span analyses and the corresponding attribute for prolongational analyses. At the same time, paired t-tests suggest that the sample means have statistically significant differences for all three attributes as well, when comparing time-span and prolongational reductions. Lastly, though the paired histograms for height and average path length may appear similar, the left-right skew paired histograms seem more visually different. This is confirmed via a two-sample Kolmogorov-Smirnov test, which indicates the left-right skew values for time-span versus prolongational reductions are drawn from different distributions. All of these statistical significances account for multiple comparisons using the Šidák correction ($\alpha = 0.05 \rightarrow \alpha = 0.017$).

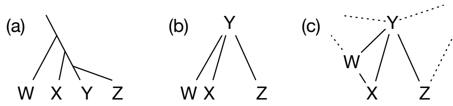

Figure 7. Illustration of a time-span/prolongational tree structure converted into a MOP.

6. CONCLUSIONS

The data presented here suggest a number of conclusions. The first two experiments involving ParseMop imply that when ParseMop makes mistakes in analyzing music, the mistakes do not drastically change the overall shape or structure of the corresponding ground-truth analysis. This information is challenging to reconcile with the fact that ParseMop, like any music analysis algorithm derived from context-free parsing techniques, does no global calculations related to shape or structure during the analytical process. One explanation is that the notes of a music composition imply an overall shape and structure that the analytical process simply reveals, in that the shape is inherently present in the music and does not have to be given explicitly to the grammar. If this were true, then using a formal grammar class higher in the Chomsky hierarchy (e.g., a grammar with some amount of context-sensitivity) may not be necessary to create algorithms that can analyze music satisfactorily.

The third experiment comparing time-span and prolongational analyses reveals fundamental differences and similarities in the overall structure of the two analytical forms. For instance, it is clear that both types of reductional system strongly prefer balanced, shallow trees, as is clear from the histograms on height, average path length, and left-right skew. Also, both analysis varieties produce trees that slightly skew to the left. However, our statistical tests also strongly suggest that the underlying distributions of the global MOP structure attributes are different, even if the differences in means happen to be small.

7. REFERENCES

[1] Helen Cameron and Derick Wood. Maximal path length of binary trees. *Discrete Applied Mathematics*, 55(1):15–35, 1994.

[2] Édouard Gilbert and Darrell Conklin. A probabilistic context-free grammar for melodic reduction. In *Proceedings of the International Workshop on Artificial Intelligence and Music, 20th International Joint Conference on Artificial Intelligence*, pages 83–94, Hyderabad, India, 2007.

[3] Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo. Musical structural analysis database based on GTTM. In *Proceedings of the 15th International Society for Music Information Retrieval Conference*, pages 325–330, 2015.

[4] Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo. σGTTM III: Learning based time-span tree generator based on PCFG. In *Proceedings of the 11th International Symposium on Computer Music Multidisciplinary Research*, pages 303–317, 2015.

[5] Maarten Keijzer and James Foster. Crossover bias in genetic programming. In *Proceedings of the European Conference on Genetic Programming*, pages 33–44, 2007.

[6] Phillip B. Kirlin. A data set for computational studies of Schenkerian analysis. In *Proceedings of the 15th International Society for Music Information Retrieval Conference*, pages 213–218, 2014.

[7] Phillip B. Kirlin and David D. Jensen. Using supervised learning to uncover deep musical structure. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, pages 1770–1776, 2015.

[8] Donald E. Knuth. *The Art of Computer Programming, Volume 3: Sorting and Searching*. Addison Wesley Longman, Redwood City, CA, second edition, 1998.

[9] Fred Lerdahl. *Tonal Pitch Space*. Oxford University Press, Oxford, 2001.

[10] Fred Lerdahl and Ray Jackendoff. *A Generative Theory of Tonal Music*. MIT Press, Cambridge, MA, 1983.

[11] Alan Marsden. Schenkerian analysis by computer: A proof of concept. *Journal of New Music Research*, 39(3):269–289, 2010.

[12] Panayotis Mavromatis and Matthew Brown. Parsing context-free grammars for music: A computational model of Schenkerian analysis. In *Proceedings of the 8th International Conference on Music Perception & Cognition*, pages 414–415, 2004.

[13] Mariusz Rybnik, Wladyslaw Homenda, and Tomasz Sitarek. *Foundations of Intelligent Systems: 20th International Symposium, ISMIS 2012*, chapter Advanced Searching in Spaces of Music Information, pages 218–227. 2012.

[14] Jason Yust. *Formal Models of Prolongation*. PhD thesis, University of Washington, 2006.

[15] Jason Yust. *Organized Time*, chapter Structural Networks and the Experience of Musical Time. 2015. Unpublished manuscript.