ABSTRACT

This work is an effort towards the development of a system for the automation of traditional harmonic analysis of polyphonic scores in symbolic format. A number of stages have been designed in this procedure: melodic analysis of harmonic and non-harmonic tones, vertical harmonic analysis, tonality, and tonal functions. All these informations are represented as a weighted directed acyclic graph. The best possible analysis is the path that maximizes the sum of weights in the graph, obtained through a dynamic programming algorithm. The feasibility of this approach has been tested on six J.S. Bach’s harmonized chorales.

1. INTRODUCTION

The musical analysis is a mean to better understand the thought of the composer when creating a piece. A musician must perform a good musical analysis to execute a correct interpretation of a work. The melodic, harmonic, and tonal function analyses are the basic elements in order to achieve an optimal musical analysis. Besides the interpretation, there are many applications of automatic analysis to diverse areas of music: education, score reduction, pitch spelling, harmonic comparison of works, etc.

The automatic tonal analysis has been tackled under different approaches and objectives. Some works use grammars to solve the problem [6]. There are probabilistic models like that in [3] and others based on preference rules or scoring techniques [5]. Taube [4] solves the problem by means of model matching. A more comprehensive review of these works can be found in [1].

Our objective is not only to obtain a high percentage of correct analyses, but also to describe in a human-readable way the reasons why the system has chosen an analysis.

The system described in this paper uses both a rule engine approach and a ranking method, along with a dynamic programming algorithm, to describe analytically a musical piece. The complexity and nature of the variables that must be utilized to perform a musical analysis are different among genres. However, the general scholastic method can be explained using the harmonized chorals of Bach. The presented work is devoted to test the ability of the proposed approach and has been founded on the baroque music rules and tested using some Bach chorals.

2. METHODOLOGY

In order to analyze a musical piece the system performs the following steps.

Firstly, a melodic analysis is performed. The notes are tagged either as harmonic tones for those belonging to the current harmony at each time, or as non-harmonic tones for those ornamental notes. There can be notes with different possible analysis that will be disambiguated later.

Secondly, a vertical analysis of the piece is performed: after segmenting each bar into a number of time windows, all possible chords are obtained from the notes in each individual window.

The third step chooses which of the 24 different tonalities are candidates to be the central key in each window, given the accidentals of the notes involved.

The fourth step has all the possible chords and tonalities for each window as input. From these data, a weighted acyclic directed graph (wDAG) organized by layers is built. Each layer represents a window. The nodes of the graph correspond to chords with tonal functions in a tonality. The edges of the graph contain the valid progressions between the nodes in successive layers, weighted according to the importance of those progressions in order to establish a tonality.

The final step uses a dynamic-programming approach to compute the best path along the graph, discovering the best tonality and tonal function sequence. The output of this step is the Roman numeral analysis along with a tonality segmentation. Finally, those notes still having a multiple melodic analysis are disambiguated.

Next, each of these steps are described more indepth.

2.1. Melodic analysis

The melodic analysis tags each note as harmonic or non-harmonic-tone (NHT) based only on the melodic information, leaving aside any harmonic vertical information. Since a given note can have multiple different analysis, a confidence value is also computed. The NHT tags are given based on the type of ornament: neighbor, etc.

The JBoss Drools rule engine framework has been used to implement a series of rules written by the authors based on the music theory applicable to the baroque period. Those rules are based on meter, pulse, duration, and pitch interval information.

Before detailing the rules some terms must be defined.

---

1 Supported by the projects GV06/166 and CICYT TIN2006-14932-C02, partially funded by EU ERDF

2 http://labs.jboss.com/jbossrules
Definition 2.1 \( rd(n_i) = \frac{duration(n_i)}{duration(beat)} \)

The relative duration function determines the ratio between the duration of a note, \( n_i \), and the duration of a beat.

Definition 2.2 \( ratio(n_i) = \frac{rd(n_i)}{rd(n_{i-1})} \times \frac{rd(n_{i+1})}{rd(n_i)} \)

The ratio function is used to compare the relative duration of a note with its next and previous notes.

Definition 2.3 \( pitchName(n_i) \)

It defines the position of the name of the note \( n_i \) in the ordered set \( \{C, D, E, F, G, A, B\} \).

Definition 2.4 \( pitchClass(n_i) \)

Order of the note name including its accidental in within the octave. (e.g. \( pitch(E♭4) = 3 \).)

Definition 2.5 \( pitchIntv(n_i, n_{i+1}) = d.s \)

It computes the pitch interval between two notes \( n_i \) and \( n_{i+1} \), where \( d = pitchName(n_{i+1}) - pitchName(n_i) + 1 \) and \( s = pitchClass(n_{i+1}) - pitchClass(n_i) \) specified with a resolution of two decimals. A positive value indicates an ascending interval. E.g., the unison interval \( pitchIntv(n_i, n_i) = 1.00 \) and the tritone is 4.06.

Definition 2.6 \( prevI(n_i) \)

The previous interval is the interval between a note \( n_i \) and its predecessor. \( prevI(n_i) = pitchIntv(n_{i-1}, n_i) \)

Definition 2.7 \( nextI(n_i) \)

The next interval is the interval between a note \( n_i \) and its successor. \( nextI(n_i) = pitchIntv(n_i, n_{i+1}) \)

Definition 2.8 \( subbeat(n_i) \)

It is a boolean function that is true when the note onset does not match the exact position of the beat.

Definition 2.9 \( strong(n_i) \)

For quaternary meters a note is strong when its onset is located in the first or third beat of the measure. In ternary meters, it is strong if and only if it onsets in the first beat. For the compound meters, this function can be computed from these two situations.

Definition 2.10 \( tied(n_i) \)

It is a boolean function that is true whenever the note \( n_i \) is tied to the previous note using a prolongation tie.

A NHT is tagged that way depending on its location in strong or weak beats, its duration, and the intervals it defines from its previous note and to the next one. There are 38 different rules that can be summarized in a simplified version in Table 1.

Each rule in the whole set assigns a confidence to the harmonic tag selected from the following catalog: ‘s’ for sure, ‘h’ for high, ‘m’ for medium, and ‘l’ for low. This confidence value will be useful in next steps. The whole set of complete rules has been stored in http://grifia.dlisi.ens-lyon.fr/.

<table>
<thead>
<tr>
<th>NHT kind</th>
<th>rules</th>
</tr>
</thead>
</table>
| appoggiatura | strong \( ^\text{\&} \) prevI \( = 1.00 \)
| suspension | strong \( ^\text{\&} \) prevI \( = 1.00 \) \& nextI \( \in \{-2.01, -2.02, +2.01\} \) |
| passing tone | \(~\text{strong} \) \& (ratio \( \leq 1 \)) \& (prevI \( \in \{-2.01, -2.02\} \)) \& (nextI \( \in \{+2.01, +2.02\} \)) \\
| neighbor tone | \(~\text{strong} \) \& (ratio \( \leq 1 \)) \& (prevI \( \in \{+2.01, +2.02\} \)) \& (nextI \( \in \{-2.01, -2.02\} \)) |

Table 1: Summary of melodic rules for a note \( n_i \). The parameter \( n_i \) of the functions has been omitted to reduce space. Actually, a passing tone starting in a weak beat or sub-beat is allowed using a low confidence.

2.2. Chord extraction

For each of the bars, the duration of the shortest figure or rest is selected as the time resolution for that bar. Then, the bar is divided into windows of that duration for all the voices in the score. This way, the musical piece is sliced into a sequence \( W \) of windows. For each window \( w \in W \), a set \( S_w \) of all notes sounding in it and a set \( C_w \) of chords from all combinations of notes in \( S_w \) are built.

Two adjacent notes of a chord \( c_i \in C_w \) must hold (1):

\[
\forall \text{ note}_{j-1}. \text{ note}_{j} \in c_i, \quad \text{pitchIntv}(\text{note}_{j-1}, \text{note}_{j}) \in \{3.s, 5.s\}.
\]

A backtracking scheme computes all combinations of note names without octave in \( S_w \) leaving out those orderings containing adjacent notes not holding that condition, and storing in \( C_w \) the set of valid chords.

Note that the backtracking process reorders all notes in the chords: even if the original positions of the notes suggest an inversion, this process deletes it and returns the root note of the chord to the lowest note position.

2.3. Accidental analysis

The accidental analysis filters out from the 24 possible tonalities those that are not feasible given the accidentals of a set of notes. This filtering is performed for each window. The following definitions will be used to describe how this step of the analysis works.

Definition 2.11 \( \text{mode}(k) \in \{\text{Maj}, \text{min}\} \)

The mode of a tonality can be major or minor.

Definition 2.12 \( \text{expectSem}(g, m) \)

The values of this function are detailed in table 2 as the set of valid semitone intervals from the first degree of a scale to the specified degree \( g \) given the mode \( m \) of a tonality.

Definition 2.13 \( \text{actualSem}(n_i, k) \)

It its value corresponds to the actual number of semitones from the tonic of \( k \) to \( n_i \).
Definition 2.14 \( \text{deg}(n_i, k) \)
It is defined as the degree of a note \( n_i \) given a tonality \( k \).

Definition 2.15 \( \text{isDiatonic}(n_i, k) \)
This boolean function is computed as
\[
\text{isDiatonic}(n_i, k) = \text{actualSem}(n_i, k) \in \text{expectSem}(\text{deg}(n_i, k), \text{mode}(k))
\]
We introduce here the set \( K_w \) of all valid tonalities for the set of notes in a given window, \( S_w \). A tonality \( k \in K_w \)
if \( \text{isDiatonic}(n_i, k) \) is true \( \forall n_i \in S_w \).

<table>
<thead>
<tr>
<th>\text{expectSem}</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maj</td>
<td>0</td>
<td>2,(1)</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>min</td>
<td>0</td>
<td>2,(1)</td>
<td>3,(4)</td>
<td>5</td>
<td>7</td>
<td>9.8</td>
<td>11.10</td>
</tr>
</tbody>
</table>

Table 2: Diatonic scales for the definition (2.12). The cells contain all the valid semitone differences from the tonic of the specified degree, the (1) value represents the Neapolitan, the (4) represents the Picardy ending.

2.4. Tonal functions
Music is based on a continuous flow of tensions and relaxations. In tonal music, this is performed by adding dissonances to the tensions that are resolved to successive chords without dissonances (i.e., chords with tonic function). When passing from musical instability to stability, a transitory function named subdominant is included. Under our point of view, all notes, all chords, all music fulfill a transitory function named subdominant is included. In our knowledge, this treatment is not found in the music theory literature, however, we attribute this function to it because we consider the tonal functions as levels of tension and musical stability. From this point of view, and considering three levels of tension, when the third degree is located between a tonic and a dominant, it should be given a subdominant transitory function.

VI chord
As stated above, the VI chord can be assigned two tonal functions depending on its harmonic surroundings. The VI degree will be a subdominant function when the previous chord is a tonic or a subdominant. In this case, the next chord has no impact in this decision. Otherwise, if the previous chord is a dominant function, this VI chord is considered as tonic.

2.5. Weighted Acyclic Directed Graph
As outlined in section 2.2, the piece is segmented in \( |W| \) time windows and, from each window, \( w \in W \), the sets of chords \( C_w \) and valid tonalities \( K_w \) are built. Now, following the principles exposed in 2.4, the set \( F_{w,k}(c_i) \) of all feasible tonal functions \( f \) for all \( k \in K_w \) and all \( c_i \in C_w \) is computed. The wDAG is constructed using these data.

2.5.1. Graph construction
Let’s define the wDAG as \( G = (V, E, D) \), where \( V \) is the set of nodes and each \( v \in V \) is labeled with the values provided by \( F_{w,k}(c_i) \). Thus, we will use the labels (tonal functions) \( f \) to represent the nodes, instead of \( v \). \( E \subseteq V \times V \) is the set of edges, and \( D \) is the set of weights that are computed by a weight function \( d : E \rightarrow \mathbb{R} \). The vertices are partitioned into \( |W| \) disjoint subsets \( V_i, 1 \leq i \leq |W| \), in a way that, if \( \{f_a, f_b\} \in E \), then \( f_a \in V_i \) and \( f_b \in V_{i+1} \). Each subset \( V_i \) includes the nodes from a given window \( w \). This way, the graph is structured as a sequence of layers, \( V_i \), representing the course of time along the score. The definition of the weight function \( d \) is explained below in the cadences section.

Cadenas
Cadences are points of relaxation that have the function of reaffirming the tonality. A cadence takes place when it concludes with the tonic function, and a half cadence when the target is a dominant or subdominant. Cadences and half cadences are summarized in Table 4. The sequence of cadences and half cadences is a hint towards the recognition of a tonality in a given point. A sequence of chords analyzed from two different tonalities yields two different cadence / half-cadence successions. Weight information is introduced in the graph for the edge linking nodes \( f_a \) and \( f_b \) in adjacent layers by means of the
weight function \(d(f_a, f_b)\). The more conclusive a cadence is, the higher the weight is assigned to it. Each sequence of tonal functions will produce a different weight value. The path with the highest sum of weights will be selected as the most suitable one. Those weights have been established empirically. A negative value is assigned to reflect the tonal regression \(D \rightarrow S\). Tonality changes are allowed only when a cadence in a new tonality is found. Not feasible progressions (those specified in section 2.4) are weighted with \(-\infty\).

<table>
<thead>
<tr>
<th>Cadence</th>
<th>Tonal funct.</th>
<th>Degrees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect authentic</td>
<td>(D \rightarrow T)</td>
<td>(V \rightarrow {I, I})</td>
</tr>
<tr>
<td>Imperfect</td>
<td>(D \rightarrow T)</td>
<td>(vii \rightarrow I, vi_d \rightarrow i)</td>
</tr>
<tr>
<td>Interrupted</td>
<td>(D \rightarrow T)</td>
<td>(V \rightarrow {vi, VI})</td>
</tr>
<tr>
<td>Plagal</td>
<td>(S \rightarrow T)</td>
<td>({IV, iv, VI, vi, ii_d, ii})</td>
</tr>
<tr>
<td>Dominant (Hc)</td>
<td>(T \rightarrow D, S \rightarrow D)</td>
<td>({I, I, III, iv, IV, iv, VI, VI})</td>
</tr>
<tr>
<td>Subdominant (Hc)</td>
<td>(T \rightarrow S, S \rightarrow S)</td>
<td>({I, I, III, iii, IV, iv, vi, VI})</td>
</tr>
</tbody>
</table>

Table 4: Cadences. ‘\(Hc\)’ stands for half cadence. The degrees in lower case represent minor chords. Subindex \(d\) is applied to diminished chords.

2.5.2. Best path computation

Once the graph is constructed, the selection of the best path is reduced to the problem of the computation of the best path in a graph using dynamic programming [2]. The nodes visited in this path are taken for the best analysis.

2.6. Melodic analysis disambiguation

After the best path computation, the system has selected which chord, tonal function, and tonality are the best for each window. The first step of the analysis (the melodic analysis step) left some notes with conflicting rules unable to classify them in harmonic or non-harmonic tones. Now, having solved the harmony of the whole piece, those notes are tagged as NHT when they do not belong to the current chord at each moment.

3. EXPERIMENTS

As mentioned above, to test the system the harmonized chorals from J.S.Bach have been used. The input of the system was transcriptions of the chorals in MusicXML format found at the Humdrum Kern Scores Site. Due to the lack of a digitally annotated corpus of these chorals, we focused the analysis of results on the chorals: choral #25, BWV-0269, BWV-0367, BWV-0400, BWV 2-6. We did so in order to be able to compare our results to those of a relevant work named MTW (Music Theory Workbench) [4]. A full report of the obtained results can be downloaded from our website quoted below.

Since two different analyses sometimes can be both valid we cannot give success percentages in order to compare to those reported by MTW. For our point of view, the MTW fails in some tonal function progressions and seems to make mistakes when analyzing alternative tonalities by not solving the chord cadence (e.g. BWV 2-6 at bar 3, beats 2-4). Our system corrects those errors by means of the cadence scoring. However, we must correct some errors the MTW does not make. Anyway, our system has failed only in two chords in the analysis of choral #25.

4. DISCUSSION AND CONCLUSIONS

This paper has described a system to analyze automatically a score from the melodic and harmonic points of view, providing the tagging of each note as harmonic or not harmonic, the tonality changes and the Roman numeral analysis of each chord along with its tonal function. The presented system is comparable in results to those reported by MTW [4], but it has the advantage to be ready to work with monodic melodies only adding more possibilities of analysis at each layer of the graph.

We are working on it to compact the output so that the chords that have been split by the windowing process can be merged again. Currently the system does not show the inversion of chords, the implementation of this feature is straightforward and is being developed. We have defined the conditions for the melodic tags: double neighbor tone, cambiatta, escape tone, and fux, and currently we are implementing them. That may correct some of the erroneous analysis the system outputs nowadays. Other important line of work is to improve the merging of the melodic analysis with the harmonic steps of the process.

Modulation points are currently displaced in a ± one chord distance. This is due to the dynamic programming algorithm. Hopefully, this problem will be solved by the use of a harmonic rhythm and a modulation subsystem.

Other current work is the creation of a larger corpus of tagged pieces to be able to learn the scoring of the tonal function progressions from that corpus. The next process to be tackled is the redefinition of rules for other genres.

References