TWO WEB APPLICATIONS FOR EXPLORING MELODIC PATTERNS IN JAZZ SOLOS

Klaus Frieler[1*] **Frank Höger**[1] **Martin Pfleiderer**[1] **Simon Dixon**[2]

[1] Institute for Musicology, University of Music "Franz Liszt" Weimar, Germany

[2] Center for Digital Music, Queen Mary University of London, UK

*Please direct correspondence to: klaus.frieler@hfm-weimar.de

ABSTRACT

This paper presents two novel user interfaces for investigating the pattern content in monophonic jazz solos and exemplifies how these interfaces could be used for research on jazz improvisation. In jazz improvisation, patterns are of particular interest for the analysis of improvisation styles, the oral transmission of musical language, the practice of improvisation, and the psychology of creative processes. The ongoing project "Dig That Lick" is devoted to addressing these questions with the help of a large database of jazz solo transcriptions generated by automated melody extraction algorithms. To expose these transcriptions to jazz researchers, two prototypes of user interfaces were designed that work currently with the 456 manually transcribed jazz solos of the Weimar Jazz Database. The first one is a Shiny application that allows exploring a set of 653 of the most common patterns by eminent players. The second one is a web interface for a general two-staged pattern search in the Weimar Jazz Database featuring regular expressions. These applications aim on the one hand at an expert audience of jazz researchers to facilitate generating and testing hypotheses about patterns in jazz improvisation, and on the other hand at a wider audience of jazz teachers, students, and fans.

1. INTRODUCTION

Music Information Retrieval offers exciting options for musicological research, particularly for methodologies which are hard (or impossible) to carry out manually, e. g., large corpus studies and measuring acoustical properties. One such field of application is the mining of patterns. Patterns – and repetitions in general – play an important role in nearly all music styles [10] and are thus of interest for many sub-fields of musicology. In particular, they form a crucial component of jazz improvisation [1, 13, 14]. The concept of 'pattern' can be defined in different ways and appears under different names in the literature. The more formal definition as 'repeated sub-sequences' (over a suitable sequence space) contrasts with more specific usages in jazz theory and practice, where patterns are often called 'formulas', 'licks', 'stock-phrases', and 'riffs'. The main differences between these terms lie in their supposed origin, their function, and their musical characteristics .

A formula is mostly understood as a rather short pattern, which is well-rehearsed by an improviser. A formula is generally not musically autonomous, i. e., it can be rhythmized differently or embedded in other formulas to make longer phrases (e. g., John Coltrane's solo on "Giant Steps" [9]). The term 'lick' (or stock-phrase) usually refers to a melodic unit with a distinctive recognizable character. In some cases, licks trace their origin to an individual performer or even to a single solo. For example, Charlie Parker created many licks that were used by other jazz musicians [14]. In most cases, however, licks cannot be attributed to a single originator. Thus, they form specific music vocabularies of smaller and wider scope. A riff can be regarded as a lick which is constantly repeated as an accompaniment and has thus a different musical function than a normal lick [12]. Other special cases are short quotations of popular tunes which are often used to humorous effect or for the cultural practice of inter-textuality ('intermusicality'), or 'signifyin' [11].

Given the significance of patterns and licks in jazz, several research questions arise. Some concern historical issues, e. g., the oral tradition of licks and the development of a typical jazz language; some are of a more systematic nature, e. g., the psychology of the creative process, where patterns and formulas can be regarded as necessary to accomplish the highly virtuoso feat of modern jazz improvisation. Some of these research questions are:

- To what extent are patterns and licks used to shape an improvisation?
- When and by whom are patterns and licks created and how are they transmitted between players over time (*pattern archeology*)?
- Does pattern usage change with time and styles?
- Is there an influence of jazz education on pattern usage (e. g., by published pattern collections)?
- How are patterns used to build phrases, e. g., to construct a typical bebop line?

- Which role do external musical influences such as quotes and signifying references play in jazz improvisation?

In this paper, two web tools which could help to address these questions are introduced. First, the research background and some similar tools are discussed (Sect. 2), before we describe the tools in detail in Sect. 3, including two use case examples. The web applications are still prototypes under active development, but they are already helpful to make some interesting observations which will be reported in Sect. 4. Thoughts on future prospects of these tools conclude the paper (Sect. 5).

2. BACKGROUND

The project "Dig That Lick: Analysing Large-Scale Data for Melodic Patterns in Jazz Performances" (DTL) is a two-year project within the fourth "T-AP Digging Into Data Challenge". [1] It sets out to investigate some of the aforementioned research questions using an interdisciplinary approach combining musicology, computer science, MIR, and jazz research. The project aims, first, at developing tools for pattern mining on symbolic as well as audio data, and, second, at understanding psychological and social aspects of patterns and licks in jazz. The development of appropriate tools consists of three main pillars:

1. Automatic transcription of jazz solo improvisations from audio informed by discographic metadata.
2. Pattern mining and search in the melody transcriptions.
3. Development of suitable user interfaces.

Since the project is still in its initial phase, we will focus in this paper on the second and third issues by describing two prototypes of user interfaces for pattern mining in the Weimar Jazz Database [18] which was created by the Jazzomat Research Project [19] [2] .

In recent years, several scientific or commercial web-based melody search engines with interfaces for different databases of different provenance and quality have been implemented, all of which are scored-based. A web search showed that many of these projects are now defunct or discontinued. To name a few: C-Brahms (defunct), Midomi (discontinued, but seemingly functional), Hymnar (active, only hymns), Mutopia, Music N-gram Viewer (discontinued, but functional), Best Classical Tunes (outdated, but functional), and Melody Search (discontinued, functionality unclear due to Flash player issues). The large number of abandoned sites suggests that melody search is not very popular with a general audience. Some more recent sites, though, aim at musicological experts, e. g., the Troubadour Melodies Database, Global Chant, and Cantus Manuscript Database, which provide simple but efficient search interfaces to specialized corpora.

[1] http://dig-that-lick.eecs.qmul.ac.uk/

[2] https://jazzomat.hfm-weimar.de

An older but still functional melody search engine is *Themefinder* [3] , which interfaces with some large databases of folk and classical music in **kern format. The search works well and is fast, though it does not offer metadata filters, regular expressions, or a multi-staged search.

Musipedia [4] is branded as a "Wikipedia for Music" and is based on a user-generated database of melodies. Search queries are given in Lilypond format, but a piano-like user interface to enter queries with the help of the mouse is also provided. The search is based on similarity matching [22] and, hence, always fuzzy; it is not possible to enforce only exact matches. The result set always comprises full melodies without indicating the matching location for the query. The underlying corpus is not clearly specified, but it seems that some well-known databases such as the Essen Folk Song Collection [21] are incorporated.

Furthermore, Gulati [7] developed a system for melodic pattern discovery in Indian Art Music based on automatically extracted pitch contours and a large set of specialized methods. A demo for browsing patterns in a large audio corpus and a visualization of pattern networks including audio snippets can be found on the accompanying web site [5] .

The investigation of patterns in jazz has rather different requirements compared to those melody search engines. Particularly, the hybrid format of the Weimar Jazz Database, which combines transcriptions with audio, as well as the greater length of jazz solos (as compared to, for instance, incipits, and folk songs) demands fine grained and controlled access to pattern instances. Furthermore, to assist users during exploration, providing scores and audio snippets along with more abstract representations is important in order to connect to established methodological standards in jazz research and practice.

3. TWO PATTERN MINING APPLICATIONS

3.1 The Pattern History Explorer

The main goal of the Pattern History Explorer [6] , an interactive Shiny web application [2], is to enable the exploration of interval patterns in jazz solos by providing information from many different angles. It provides an overview of interval patterns frequently used by a selected subset of performers and traces their usage in the Weimar Jazz Database, allowing for the discovery of cross-artist and cross-temporal relationships.

Currently, 653 interval patterns with 11,630 instances are included. The pattern corpus was created by mining for interval patterns in solos of eminent performers using the partition mode of the `melpat` module from the MeloSpySuite [6]. Subsequently, all instances of these patterns were searched for in the Weimar Jazz Database, and the results were included in the application. Since the interval distributions of jazz solos are dominated by step-wise motion,

[3] http://www.themefinder.org/

[4] https://www.musipedia.org/

[5] http://compmusic.upf.edu/node/304

[6] https://jazzomat.hfm-weimar.de/pattern_history/

a certain number of patterns can be expected by chance alone, even more so by assuming Markov processes of first or higher order [4]. Therefore, the following restriction to a minimum number of instances and sources was imposed to ensure significance of the patterns: interval patterns were limited to be of no fewer than six elements occurring in at least three different solos of at least one musician. According to previous investigations [4], this length seems to be a critical point for pattern distributions. The number of instances of each pattern depends partly on the amount of musical material available. For Bob Berg the criterion was relaxed to patterns of at least seven elements occurring at least twice in two different solos, since from a former study [5] it was already known that many interesting and highly peculiar patterns occur only twice in the subcorpus of Berg's solos in the Weimar Jazz Database. Another exception was Charlie Parker, for whom the source patterns were extracted not from the Weimar Jazz Database but from the Omnibook, a collection of 56 transcriptions of his solos. In order to find only Parker's more eminent patterns, a criterion of at least six elements occurring in at least ten different solos was applied.

In general, the user of the Pattern History Explorer selects a certain interval pattern from the overall set of 653 patterns. Several options are available in order to filter the pattern set or to change the ordering of the patterns according to several criteria (e. g., filtering by performer, length, intrinsic characteristics such as Huron contour [8] or tonality type, or content). For the selected pattern, various kinds of information can be accessed in the following tabs:

- **Listen & See**. A sortable list of all instances of the pattern in the Weimar Jazz Database is displayed. It includes metadata such as name of the performer, title of the solo, year of recording, metrical and start position. In one column the tonal context is displayed as a combination of chord context and extended chordal pitch class values (CDPCX, cf. [6]). Most importantly, score snippets and audio links allow for visual and aural inspection. Here, and in the following tabs, cross-links to the Pattern Search web application (see below) is provided to allow more refined searches.

- **Instances**. Further information about the instances can be found here. A rhythmic encoding based on absolute inter-onset interval (IOI) class (very short, short, medium, long, very long), the starting pitch, the chord context, the CDPCX value and a binary vector indicating the position of metrical accents in the pattern are displayed. Additionally, some of the metadata from the Listen & See tab are repeated. The final column indicates whether the instance is contained in one single phrase. This table is also sortable.

- **Stats**. Musical characteristics and statistical information of the pattern are compiled. Due to space limitations we cover just some of these here, and refer the interested reader to the online documentation for full details. One important feature is *log excess probability*, which is defined as $log_{10}\frac{p_o}{p_e}$, where p_o is the observed frequency of a certain pattern within the database and p_e is its expected frequency according to a Markov model of zeroth order taken from the global interval distribution. Not surprisingly, log excess probability is strongly correlated with pattern length ($r = .82$, $p < .001$). Another interesting feature is the number of different starting pitches, as this indicates whether an interval pattern is indeed a pitch pattern. A pitch pattern with a fixed set of pitches might be based on a physiological motor pattern, i. e., tied to a specific fingering on the instrument. Other interesting information provided here is which soloists favor this pattern and who played it first. A list of instances by performer can be found at the bottom of the page.

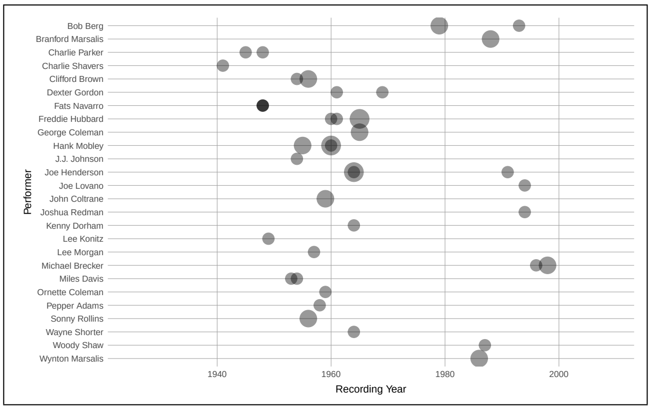

Figure 1. History plot of the pattern [-2, -2, -1, -2, -2, -1, -1]. The first instance can be found in a solo by Charlie Shavers on "Limehouse Blues" from 1947. A cluster can be identified in the years around 1960. The circle radius represents the frequencies of patterns in each solo in that year.

- **Timeline**. This page contains a visualization of the distribution of pattern instances over time and performer. See Fig. 1 for a sample plot for the pattern [-2, -2, -1, -2, -2, -1, -1] (a descending diatonic line with a chromatic ending). The first instance can be found in a solo by Charlie Shavers from 1947 and it is favored by Hank Mobley with six instances in three different solos, mostly starting on the supertonic (second scale step) over a ii^7-V^7 transition. This tonal interpretation is preferred in many other instances. Other players who use this pattern are Joe Henderson and Freddie Hubbard with five instances each. As the plot shows, this pattern is spread widely over jazz history since its first occurrence.

Besides the tabs related to patterns, the tab 'Info' provides general information on the Pattern History Explorer and the 'Help' tab contains detailed documentation. The tab 'General Stats' collects plots pertaining to statistics of all included patterns and instances (e. g., Fig. 3).

3.2 Pattern Search

While pre-computing a set of patterns is helpful in regard to the exploratory approach of the Pattern History Explorer, searching for instances of arbitrary patterns of any length and frequency of occurrence within a database requires a different approach. Although it is possible to search the Weimar Jazz Database with its accompanying software MeloSpySuite and MeloSpyGUI, our web-based pattern search interface [7] provides most of the functionality of the `melpat` search module while also extending it with audio and score snippets (both as isolated patterns and within their melodic context) for visual and aural inspection.

To execute a basic search, the user has to enter a pattern as a space or comma separated list of elements and choose a corresponding transformation, that is, a mathematical mapping of the basic melodic representation, also known as viewpoints [3]. Currently, ten pitch-related transformations for primary search are offered (e. g., MIDI pitch, semitone intervals, CDPCX). An additional 18 transformations, such as duration, IOI classes and various structural markers, are supplied for the optional secondary search. Additionally, the search space can be constrained by seven metadata categories, like performer, style, or recording year. Search patterns can be regular expressions (in a specific hybrid syntax depending on the selected transformation) [8] which allows searches for variants in a single run. The secondary search can be used to refine the result space, e. g., by filtering out certain rhythmic or metrical configurations or by confining instances to a single phrase. Since the last constraint is used frequently, there is also a shortcut checkbox that fills in the correct secondary search pattern (which is based on phrase boundary markers). The user also has the option to request inclusion of up to 20 tones before and after the actual pattern instance in both score and audio files. In order to generate these, the corresponding checkboxes have to be selected, which is the default. There might be cases though, where the result set is very large (e. g., searching for very short intervals), and one wants to avoid generating all audio and score files as it would take a considerable amount of time. If both checkboxes are disabled, instances will appear in the result table after a few seconds which allows for a first examination of the results. The generation of audio and score files is also suppressed (with a warning being shown) when the result set exceeds the (current) limit of 100 instances. Finally, by clicking the 'Display whole phrase' button, the score of the whole phrase containing the pattern is displayed.

The underlying search algorithm is built upon the basic Python regular expression module, which is fed with virtual Unicode strings constructed from the different melodic representations (transformations) with different alphabets. Scores are generated with the help of Lilypond, while audio snippets are directly extracted from the solo audio files. On average, a full search including audio and score generation takes several minutes, depending strongly on the size of the result set. If the results can be found in the cache, they are returned in a few seconds.

Finally, a documentation page [9] is included as well as a search history which gives a (browser-local) overview of all distinct search requests which have been submitted by the user. For each specific search, a comment can be added and by clicking on the 'Restore Search' link the result of the corresponding search is displayed.

3.3 Example 1: A typical Parker lick

Assume that we want to find all occurrences of interval pattern [-1, -2, -1, -9, 3, 3, -1, -2] which was often played by Charlie Parker within various recordings [14]. For this, we enter it into the primary pattern field and select 'Semitone intervals' as the primary transformation. By executing the search we get eight instances, six by Charlie Parker, and one each by Dexter Gordon and Sonny Stitt, who are both known to be strongly influenced by Charlie Parker [15].

3.4 Example 2: Hunting for Coltrane's "Giant Steps" pattern

In his improvisations on "Giant Steps" recorded in 1959, seminal tenor saxophonist John Coltrane repeatedly uses a four-tone pattern that consists of the root, the supertonic, the third and the fifth of the underlying chord [9, 20]. It would be interesting to know whether other jazz musicians have used this simple pattern, and if Coltrane used it on other recordings as well. Additionally, it is interesting to know whether the pattern is only played over major chords or also with minor chords. While the Pattern History Explorer currently covers only patterns with seven tones (six intervals) or more, the Pattern Search application offers several options to search for arbitrary patterns using various transformations and filters.

The pattern can be expressed, for example, with the transformation 'Chordal Pitch Class' (CPC), which maps pitches to pitch classes starting with the root of the underlying chord. Since the third of a chord can be either major (CPC = 4) or minor (CPC = 3), we use regular expression syntax to reflect this and the search pattern would be

```
0, 2, "[", 3, 4, "]", 7
```

This translates as "the root (0) followed by the supertonic (2) followed by either the minor (3) or the major third (4) followed by the fifth of the chord (7)" [10]. Note that special symbols of regular expressions, here the square brackets as character set indicators, must be quoted, because of the hybrid syntax, where chordal pitch classes are expressed as integers (not characters).

Hitting the search button yields 323 instances. Because the limit of 100 instances is exceeded, no score or audio file are generated and only the result list together with a

[7] https://dig-that-lick.hfm-weimar.de/pattern_search

[8] https://jazzomat.hfm-weimar.de/commandline_tools/melpat/melpat.html#search-pattern-syntax

[9] https://dig-that-lick.hfm-weimar.de/pattern_search/documentation

[10] Alternatively, the same search could be carried out using the pattern 1235 over the extended chordal diatonic pitch class (CDPCX) representation.

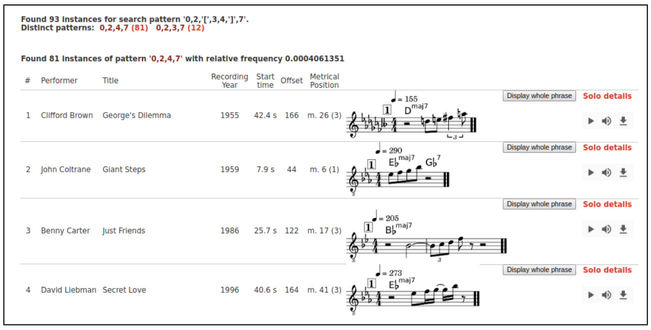

Figure 2. The first four search results for a pattern with chordal pitch class transformation using regular expression `0, 2, "[", 3, 4, "]", 7` and secondary search pattern `1, 0, ".", 0` over primary and secondary metrical accents.

warning is displayed. To prune the result set, one can tick the 'Within Single Phrase' box, which results in 297 instances. Characteristic of Coltrane's use of the pattern is that it very often starts on a strong beat (first or third beat in 4/4 time). To express this constraint with a secondary search, we use the transformation 'Primary and secondary metrical accent', which takes on the values '1' for an event on a beat position and '0' otherwise, and the search pattern

```
1,".{3}"
```

with the operation 'Match'. The dot stands for any symbol (here only '0' or '1'), while the number 3 in braces is a quantifier meaning "exactly three repetitions". Again, the quotes are necessary here because of the hybrid syntax. This leads to all CPC patterns that start on a strong beat, whereby the last three tones can lie on arbitrary metrical positions. This results in 93 instances (by 37 players in 58 different solos), which are displayed as a complete list (Fig. 2). There are 12 instances with a minor third and 83 instances with major. Interestingly, John Coltrane never uses the minor version. 26 instances originate from him, from which 18 can be found in two recordings of "Giant Steps". Michael Brecker, who is said to be heavily influenced by John Coltrane, accounts for nine instances, all of which are also major.

Visual inspection of the results shows, however, that some of the instances are either not over one single chord or do not follow an ascending step-wise motion. Similarly, even though most of the instances feature a plain motion in eighth notes, there are a few instances with rather different rhythms which still satisfy our metrical constraint. To filter these cases, one could use other transformations, e. g., searching for duration patterns. Thus, here and in other situations, a tertiary or even quaternary search stage would be needed, which is not available yet but planned for a future version of the Pattern Search application.

4. SOME OBSERVATIONS AND A HYPOTHESIS

In order to demonstrate the usefulness of the two systems in the context of jazz research, we like to report some first observations.

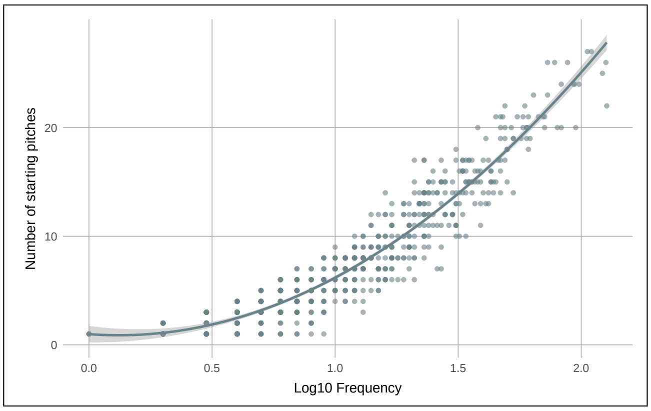

Figure 3. Number of different starting pitches versus $\log_{10}$ of pattern frequency. The correlation is $r = -.92$. The fit is a quadratic polynomial.

Observation 1 *Pattern usage varies with performers and appeared with bebop.*

Looking at the distributions of patterns in the Pattern History Explorer with respect to performer, it seems clear that different jazz improvisers have different levels of pattern usage. For example, players from earlier styles (e. g., New Orleans, Swing) have far fewer long interval patterns in their repertoire than later players. This might be partly due to the fact that these performers are seldom playing long lines – a practice that only became widespread with the advent of bebop, and which probably made the usage of patterns a necessity.

Observation 2 *The more frequent an interval pattern, the more tonally flexible it is.*

As shown in Fig. 3, the number of distinct starting pitches of a certain pattern generally increases with its frequency of occurrence. The relationship is approximately logarithmic in the frequency $N_p \propto \log f$ with a strong correlation of $r = .92$ ($p < .001$). In other words, the more frequent (and shorter) a pattern, the more tonally flexible it is. This seems to reflect typical rehearsal routines, where shorter patterns are more likely to be practiced in all keys whereas long and very long patterns are designed to fit in only one or two tonal contexts, e. g., to chord changes of specific songs. This has to be tested on a larger database, taking into account that not all keys, chords, and chord combinations are equally likely to occur in jazz.

Observation 3 *Patterns are mostly simple and reflect common rehearsal routines.*

Diatonic patterns, i. e., step-wise motions, are by far the most frequent pattern type followed by chromatic patterns. Together they account for about 78 % of all patterns included in the Pattern History Explorer. This implies that the main share of patterns is musically rather simple, i. e., built from diatonic scales, chromatic runs, and arpeggios. One can conjecture that this is a result of practice traditions in which scales and arpeggios are rehearsed for technical

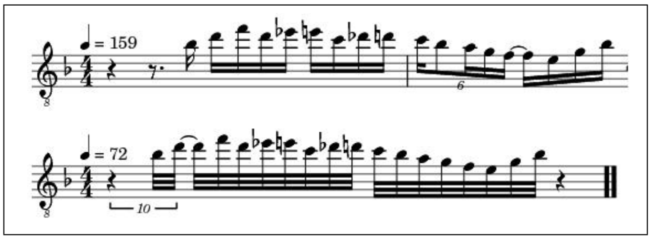

Figure 4. The interval pattern [4, 3, -3, 1, 1, -4, 1, 1, -2, -2, -1, -2, -2, -1, 3, 3] of length 16 as found in two solos by Charlie Parker.

fluency, but might end up 'to lie in the fingers', i. e., as motor patterns.

Hypothesis 1 *Jazz solos are hierarchically composed of adaptive chunks.*

A general problem of pattern mining in jazz solos and other melody corpora is, however, to distinguish truly meaningful (i. e., intended) patterns from randomly occurring patterns [7]. This is closely connected to the problem of finding adequate models for the underlying (random) processes, for which some evidence can be extracted already from the data.

Some very long patterns can be found, e. g., a pattern of 16 tones by Charlie Parker with two instances in two different solos which are also tonally and rhythmically very similar (Fig. 4). The *a priori* probability under any Markov model for this is practically zero. This clearly shows that this pattern was preconceived and then reproduced as a single chunk. In general, it seems very doubtful that jazz solos could be successfully modeled with Markov chains (of any order) on the level of single tones or intervals [16, 17].

This is also corroborated by the existence of (non-trivial) trill-like patterns (see Fig. 5 for an example), which are amongst the longest patterns that can be found. These patterns are somewhat trivial, as they are "oscillations" of a repeated shorter pattern, but they are also examples of "super-patterns", i. e., long patterns containing shorter sub-patterns. This hints at Markov models not working on the note event level but hierarchically on a chunk level.

To sum up, these first observations suggest that jazz solos do not follow Markov models of any order on a tone-level, but are rather created by hierarchical processes with interspersed "islands of high probability", where patterns are reproduced as complete chunks while being adapted rhythmically and tonally to the current context. This strategy could be dubbed "mutatis mutandis", which seems to be basic not only for improvisation but music creation in general.

5. CONCLUSION & OUTLOOK

Both applications presented in this paper are already usable interfaces for the Weimar Jazz Database and serve as prototypes for applications to explore large databases, which are going to be automatically extracted from large collections of jazz recordings. Both tools can be primarily viewed as bespoke interfaces for the specific needs of jazz researchers, but they could also be of interest to jazz teachers, students and fans, as well as for training courses in computational music analysis. Compared to the possibilities of the `melpat` module of the MeloSpySuite, they provide superior presentation of results, particularly in their provision of audio and score excerpts.

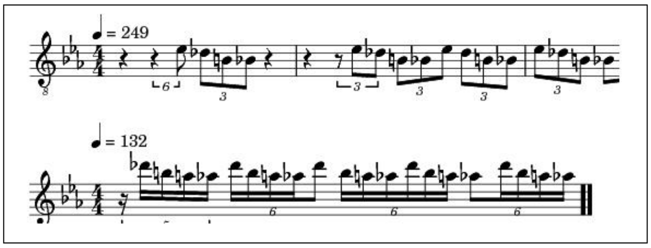

Figure 5. The interval pattern [-2, -2, -1, 5, -2, -2, -1, 5, -2, -2, -1, 5, -2, -2, -1, 5, -2, -2, -1] of length 19 as found in two solos by Bob Berg on the album 'Enter the Spirit' from 1993.

The development of both applications is still ongoing. The Pattern History Explorer will be augmented by more patterns in the future. For the Pattern Search application, removing the limit of 100 instances for full searches and speed improvements are already under construction. Another significant extension would be the implementation of arbitrarily many search stages, since the current status of only two stages is often too restrictive. The incorporation of other databases, such as the Essen Folk Song Collection, would also be feasible without major modifications.

The current system is based on the Python regular expression module and requires all melodies to be loaded into memory at once. This is sufficiently fast for the Weimar Jazz Database with 200,000 events. For searching larger data sets, however, a more advanced retrieval technology is needed, e. g., distributed NoSQL databases and sophisticated search algorithms. Moreover, the generation of score files could be optimized by switching from Lilypond to VexFlow [11], not only to speed up score rendering but also to allow for score customization.

Some features in the applications are only possible due to the high-quality manual transcriptions in the Weimar Jazz Database with its comprehensive set of annotations. In the scenario of automatically extracted transcriptions, several curtailments can be expected, e. g., transformations that depend on such annotations might not be usable. However, it is always possible to use pitch and interval transformations and to extract audio snippets for aural control, providing also feedback for the transcription algorithm.

Finally, it is planned to conduct user studies to gather feedback for further improvements of the interface.

6. ACKNOWLEDGEMENT

The "Dig that Lick" project is supported by the Deutsche Forschungsgemeinschaft (DFG, PF669/9-1) and the UK Economic and Social Research Council (ES/R004005/1).

[11] `http://www.vexflow.com/`

7. REFERENCES

[1] Paul F. Berliner. *Thinking in jazz. The infinite art of improvisation*. University of Chicago Press, Chicago, 1994.

[2] Winston Chang, Joe Cheng, J. J. Allaire, Yihui Xie, and Jonathan McPherson. *shiny: Web Application Framework for R*. 2017.

[3] Darrell Conklin and Christina Anagnostopoulou. Representation and discovery of multiple viewpoint patterns. In *Proceedings of the 2001 International Computer Music Conference*, San Francisco, 2001. ICMA.

[4] Klaus Frieler. Pattern usage in monophonic jazz solos, 2014. Talk given at the International Jazzomat Research Workshop, September, 26–27, 2014.

[5] Klaus Frieler. Bob Berg's Solo on "Angles". In Martin Pfleiderer, Klaus Frieler, Jakob Abeßer, Wolf-Georg Zaddach, and Benjamin Burkhart, editors, *Inside the Jazzomat. New Perspectives for Jazz Research.*, pages 41–84. Schott-Campus, Mainz, 2017.

[6] Klaus Frieler. Computational melody analysis. In Martin Pfleiderer, Klaus Frieler, Jakob Abeßer, Wolf-Georg Zaddach, and Benjamin Burkhart, editors, *Inside the Jazzomat. New Perspectives for Jazz Research.*, pages 41–84. Schott-Campus, Mainz, 2017.

[7] Sankalp Gulati. *Computational Approaches for Melodic Description in Indian Art Music Corpora*. PhD thesis, Universitat Pompeu Fabra, Barcelona, 2016.

[8] David Huron. The melodic arch in western folksongs. *Computing in Musicology*, 10:3–23, 1996.

[9] A. C. Lehmann and S. Goldhahn. Duration of playing bursts and redundancy of melodic jazz improvisation in John Coltranes "Giant Steps". *Musicae Scientiae*, 20(3):345–360, September 2016.

[10] Elizabeth Hellmuth Margulis. *On repeat: how music plays the mind*. Oxford University Press, New York, NY, 2014.

[11] Ingrid Monson. *Saying something. Jazz improvisation and interaction*. University of Chicago Press, Chicago, 1996.

[12] Ingrid Monson. Riffs, Repetition, and Theories of Globalization. *Ethnomusicology*, 43:31–65, 1999.

[13] Martin Norgaard. How jazz musicians improvise. The central role of auditory and motor patterns. *Music Perception: An Interdisciplinary Journal*, 31(3):271–287, 2014.

[14] Thomas Owens. *Charlie Parker. Techniques of improvisation*. PhD thesis, University of California, Los Angeles, 1974.

[15] Thomas Owens. *Bebop. The music and its players*. Oxford University Press, New York, 1995.

[16] Francois Pachet. The continuator: Musical interaction with style. *Journal of New Music Research*, 32(3):333–341, 2003.

[17] François Pachet and Pierre Roy. Markov constraints: steerable generation of Markov sequences. *Constraints*, 16(2):148–172, April 2011.

[18] Martin Pfleiderer. The Weimar Jazz Database. In Martin Pfleiderer, Klaus Frieler, Jakob Abeßer, Wolf-Georg Zaddach, and Benjamin Burkhart, editors, *Inside the Jazzomat. New Perspectives for Jazz Research*. Schott-Campus, Mainz, 2017.

[19] Martin Pfleiderer, Klaus Frieler, Jakob Abeßer, Wolf-Georg Zaddach, and Benjamin Burkhart, editors. *Inside the Jazzomat: New Perspectives for Jazz Research*. Schott Music GmbH & Co. KG, Mainz, 1 edition, 2017. OCLC: 1015349144.

[20] Lewis Porter. *John Coltrane, his life and music*. University of Michigan Press, Ann Arbor, 1998.

[21] H. Schaffrath. The Essen folksong collection. In D. Huron, editor, *Database containing 6,255 folksong transcriptions in the Kern format and a 34-page research guide*. CCARH, Menlo Park, CA, 1995.

[22] Frans Wiering, Rainer Typke, and Remco C. Veltkamp. Transportation Distances and their Application in Music-Notation Retrieval. In *Music Query: Methods, Strategies, and User Studies*, volume 13 of *Computing in Musicology*, pages 113–128. 2004.