Allison Tetreault and Thomas Van Duyne

Mark LeBlanc

COMP 131: Computing for Poets

30 April 2013

Project Parse Parker: A Computational Analysis of Charlie Parker’s Most Common Jazz Riffs

**Summary Statement**

Charlie Parker is a saxophonist extolled for his influence on improvisation in jazz music. This research, which studies ten of his most performed songs, ultimately provides evidence for which riffs, or combinations of note intervals independent of key signature, Parker used the most in improvisational settings.

**Introduction**

During his lifetime, Charlie Parker was famous for his jazz techniques and improvisational methods. Many scholars assert that elements of his style were widely copied by later jazz musicians, as his solos shaped the future of jazz improvisation. However, in order to make this claim, these scholars must first know what riffs, or motives of intervals independent of key signature, Charlie Parker played the most in his solos. These motives were the building blocks for his improvised melodies.

Thomas Owens is an accomplished jazz scholar from University of California Los Angeles. For his dissertation, “Charlie Parker: Techniques of Improvisation,” he collected records, listened, transcribed, and analyzed Parker’s solo techniques. His research became a two-volume work on Parker’s principal motives. He found that Parker used “descending scale passages as a basis for improvisation” (Owens 1974). After painstaking research of his entire discography, he found that “Parkers melodic vocabulary is based on the major-minor modal system, embellished by a liberal amount of chromaticism stemming from passing and neighboring motion, implied secondary dominants, and modal borrowing” (Owens, 1974). In other words, Parker used descending melodic lines combined with chromaticism (represented in the data as either M1 or m1) to solo over established chords.

Other scholars have also developed methods for analyzing jazz improvisation. Richard Hermann, renowned pedagogue in jazz studies, said, “Jazz pedagogy for improvisation emphasizes three constructs: motivic manipulation, ‘scale running,’ and ‘running changes’ (arpeggiation of the chords)” (Hermann, 2004). He asserts that Parker’s solos are thematic depending on song, but does not analyze his entire discography as Owens does.

The problem that this program addresses is the difficulty of analyzing jazz improvisation on a “big data” scale. There is room for human error when copying, listening, transcribing and analyzing jazz improvisation by hand. Solos were not meant to be written out; they’re spontaneous, improvised events in music. A program would be able to take these events and calculate the most used motives and phrases in a much more accurate way, essentially creating a fingerprint of that musician. Therefore, this software uses the concept of nGrams, treating intervals as words, in order to find the most four-gram, five-gram, and n-gram motives used by Charlie Parker.

**Methods**

The first step of this project was inputting 10 already transcribed songs from a Charlie Parker Omnibook in Eb PDF into a Sibelius music file manually. There were ~ 4800 notes to input. Then, the file was exported to a MIDI file, which was then converted to a CSV file with a program called MidiOX (midiox.com) executed with the command line function “midicsv [filename.mid] [desiredfilename.csv]”. The MIDI file included many unneeded words and letters, so parseParker.py scrapes only the useful data: the midi note value represented by a number between 1 and 127. Then, parseParker.py takes that file and converts the notes to intervals. Thus, notes are independent of key signature. After that, a number of functions are called, including:

(All function definitions can be found in the program as well.)

* **getNotes** – This function dives into the CSV files in the input folder (which will default to ‘CSVs’) and grabs the notes (which are currently midi values from 1-127). It then stores these values in a list called noteList.
* **getIntervals –** This function continues on from getNotes(). It takes the list of notes (noteList) and creates a list of intervals. The list of intervals is key independent, creating useful data for analysis.
* **n\_gram\_creator –** This function takes in a list of intervals and, based on a user’s smallest and largest interval sets, creates ngrams of length start\_spot to length end\_spot and prints them out into documents in the ngrams folder, adding the ngrams to a list that can be used in later functions to determine most common ngrams.
* **crunchTheNumbers –** This function accepts a dictionary and a list as inputs. The list is a list of ngrams. These ngrams are then counted and stored into the passed dictionary. The ngram itself is the key and its frequency is the value.
* **printDictionary –** This function compares the frequency of the ngrams between the lengths given by the user and then prints them out for the user in one file dynamically named based on the length of the ngrams being compared. For example, if the user was to enter 3 for the smallest interval set and 7 for the largest interval set, the file created would compare all interval sets from 3-7 and create a file called results3\_7.txt in the folder ‘outputs’ to print the results.
* **printNgramList –** This functionprints the ngram out in a file for the user called results.txt. This file can hopefully be used by the user to create interesting representations of the data. It contains all the ngrams in all the CSV files of the lengths specified by the user

The process of the overall project is:

The process of the program ParseParker.py is:

**Results**

The following is a table of the most used motives between 3- and 7-interval phrases. The letter “m” denotes a descending interval, while the letter “M” denotes an ascending interval.

***Table 1: The Top 10 3-Interval to 7-Interval Sets***

|  |  |  |
| --- | --- | --- |
| **Interval Set Frequency** | | **Interval Set** |
| 100 | | m1, m1, m1 |
| 84 | | m2, m2, m1 |
| 84 | | m2, m1, m2 |
| 79 | | m1, m2, m2 |
| 73 | | m1, m2, m1 |
| 58 | | m2, m1, M3 |
| 46 | | M2, m2, m1 |
| 44 | | m1, m1, m1, m1 |
| 40 | m2, m1, m2, m2 | |
| 39 | m1, m1, m2 | |

**The frequency an interval between the nGram 3 and 7 appears and the interval itself**

The above graph shows that 30/32 or 94% of intervals were completely comprised of descending (lowercase “m”) intervals. This is a stark contrast to the 2/32 or 6% of ascending intervals. This ties back to Owens’ claim that a majority of Charlie Parker’s licks are descending.

***Table 2: The Top Played Interval Sets from 1-10***

|  |  |  |
| --- | --- | --- |
| **Interval Set Length** | **Interval Set Frequency** | **Interval Set** |
| 1 | 800 | m2 |
| 2 | 278 | m2, m1 |
| 3 | 100 | m1, m1, m1 |
| 4 | 44 | m1, m1, m1, m1 |
| 5 | 19 | m2, m1, M3, M3, M4 |
| 6 | 12 | m2, m2, m1, M3, M3, M4 |
| 7 | 8 | m1, m2, m2, m1, M3, M3, M4 |
| 8 | 5 | M1, M2, m3, M1, M2, m3, M1, M2 |
| 9 | 5 | m3, M1, M2, m3, M1, M2, m3, M1, M2 |
| 10 | 5 | m3, M1, M2, m3, M1, M2, m3, M1, M2, M5 |

**The frequency an interval set appears based on nGrams 1-10**

Table 2 looks at the interval set length, the interval set frequency, and the interval set itself. The pyramid diagram that results offers interesting data to analyze. Firstly, the four most used phrases between set lengths 1 and 4 are descending half steps or whole steps (1s or 2s respectively). This provides evidence for the conjecture that Parker improvises heavily with descending lines.

**Concluding Statement**

This experiment provided empirical evidence using a sample of Parker’s music for Owens’ claim that Parker’s solos consist of mostly descending interval phrases. In the future, one could add more songs to this corpus and then compare Parker to other famous jazz improvisation musicians such as Thelonious Monk, Dizzy Gillespie, John Coltrane, Cannonball Adderley, or Ella Fitzgerald.

References

Owens, T. (1974). *Charlie Parker: Techniques of Improvisation* (Unpublished doctoral dissertation). University of California, Los Angeles.

Hermann, Richard. “Charlie Parker’s Solo to ‘Ornithology’: Facets of Counterpoint, Analysis, and Pedagogy.” *Perspectives of New Music* 42.2 (2004): 222-262.