SCATLAVA: Software for Computer-Assisted Transcription Learning through Algorithmic Variation and Analysis

David Su
Inspiro
104 rue d'Aubervilliers, 75019 Paris
david.d.su@gmail.com

ABSTRACT

Transcribing music is an essential part of studying jazz. This paper introduces SCATLAVA, a software framework that analyzes a transcription for difficulty and algorithmically generates variations in an adaptive learning manner in order to aid students in their assimilation and understanding of the musical material and vocabulary, with an emphasis on rhythmic properties to assist jazz drummers and percussionists. The key characteristics examined by the software are onset density, syncopation measure, and limb interdependence (also known as coordination), the last of which introduces the concept of and presents an equation for calculating contextual note interdependence difficulty (CNID). Algorithmic methods for analyzing and modifying each of those properties are described in detail; adjustments are made in accordance with user input at each time step in order to adapt to students' learning needs. Finally, a demonstration of the SCATLAVA software is provided, using Elvin Jones' drum solo from "Black Nile" as the input transcription.

1. INTRODUCTION

Transcription is a fundamental part of the jazz education process, and strengthens multiple facets of musicianship such as ear training, technique, history, and analysis. While the process of learning a jazz transcription is similar to that of learning to play a piece for a classical performance, the end goals vary, as a jazz musician is rarely called upon to recreate a prior performance note-for-note. Instead, the jazz musician aims to assimilate the vocabulary of the performance into his or her own improvisational method [1]. Software tools have proven to be useful for other facets of jazz education [2, 3], but the potential for aiding transcription studies is relatively untapped. This paper presents ***s**oftware for **c**omputer-**a**ssisted **t**ranscription **l**earning through **a**lgorithmic **v**ariation and **a**nalysis* (SCATLAVA), a program that aids with the assimilation of rhythmic material in jazz transcriptions. It uses algorithmic composition and computational analysis to help musicians more efficiently internalize the vocabulary of a transcription as well as learn the music itself with more ease. Given the author's background as a jazz percussionist, the analytical components of the software currently focus on rhythmic properties as applied to drum set performance, although the software can be easily extended to incorporate melodic and harmonic material as well as different target instruments and genre-specific parameters.

2. TECHNICAL OVERVIEW

Figure 1 details the input-process-output (IPO) model used for the SCATLAVA program.[1] The transcribed input data is represented using the platform-agnostic MusicXML format for maximum compatibility across computer systems. Conversion to and from MusicXML is supported by most major notation software such as Sibelius and Finale as well as modern web browsers with libraries such as VexFlow.[2] In addition, the flexible XML tree structure allows other visual elements and metadata, such as titles and annotations, to remain untouched by the parsing process.

Upon initialization of the program, the user can specify b, which represents the number of "bins", or beat windows, of primary or strong beats, that a measure is divided into for analysis purposes. A higher value of b corresponds to increased granularity.

3. PARAMETERS FOR ANALYSIS

Many different methods have been proposed for determining the difficulty of a piece of music, particularly within the realm of music information retrieval [4]. The primary properties of a musical passage currently examined in SCATLAVA are onset density, syncopation measure, and degree of coordination required, denoted by d, s, and c respectively. An onset refers to a non-rest note with nonzero duration, and d can be represented as the number of onsets per strong beat divided by the user-specified granularity. To calculate s, we use a variation on Keith's measure [5], adapted such that strong beats reference the first beat of a bin, and normalized to fit within the framework of variable granularity.

[1] The full source code for SCATLAVA can be found at https://github.com/usdivad/SCATLAVA.
[2] http://www.vexflow.com/

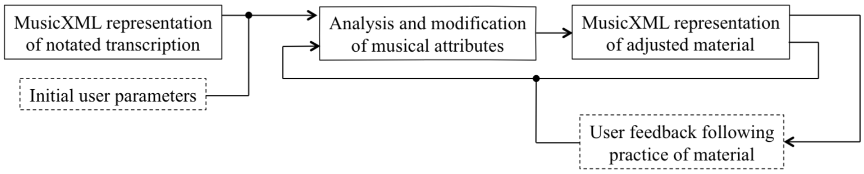

Figure 1: The IPO model for the SCATLAVA program.

Coordination between the limbs, also known as interdependence, refers to when “each limb knows exactly what the others are doing and how they work together, not independently” [6]. Here we present a method for quantifying and calculating c, the degree of difficulty in terms of coordination for a given beat window, resulting in a *contextual note interdependence difficulty* (CNID) value. The equation for calculating the CNID for a beat window is illustrated below:

$$CNID = \begin{cases} 0 + \dfrac{\alpha_i}{L} & \text{if } n_i = n_{i-1} \text{ and } n_i = n_{i+1} \\ 1 + \dfrac{\alpha_i}{L} & \text{if } n_i = n_{i-1} \text{ and } n_i \neq n_{i+1} \\ & \text{or } n_i \neq n_{i-1} \text{ and } n_i = n_{i+1} \\ 2 + \dfrac{\alpha_i}{L} & \text{if } n_i \neq n_{i-1} \text{ and } n_i \neq n_{i+1} \end{cases} \quad (1)$$

where n represents the note at subdivision index i of the beat window, α represents the number of simultaneous onsets associated with the note, and L represents the maximum number of simultaneous limbs. By default L is set to 4, representing the use of the left hand, right hand, left foot, and right foot on a typical drum set.

Once d, s, and c have been calculated for a beat window, a weighted average of the three values can be computed to yield a difficulty value D for that period. The precise values of the weights given to each input variable can be adjusted by the user; the program's default weights, denoted by w_p for parameter p, are $w_d = 0.33$, $w_s = 0.33$, and $w_c = 0.34$.

While the value of D for a single beat window can be useful for analyzing that bin itself, the difficulty of an entire measure cannot always be accurately expressed as the mean of its constituent bins’ D values. We can see from Table 1 that increasing b yields decreasing values of both s and c but not d; this is due to the fact that onset density is already expressed as a function of b, whereas both the adapted version of Keith’s measure and CNID calculation depend on inter-window note onsets. As such, the program utilizes $b = 1$ for analysis purposes in order to provide the most comprehensive calculations for s, c, and D. However, as Section 6 details, changing the value of b affects the modifications made to the phrase, and generally values of $b > 1$ yield more musically useful results. Thus, by default $b = 4$ is used when performing adjustments.

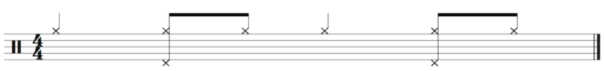

Figure 2: Drum set notation for a basic swing pattern commonly used in jazz music.

4. ADAPTIVE LEARNING

With the computed values of D, we can then begin applying modifications and creating variations on the original transcription in an adaptive learning manner. Adaptive learning refers to a method by which the educator adjusts material presented to the student based on certain properties of how the student is learning [7]. SCATLAVA implements a variant of the method proposed in [8], utilizing user self-assessments to drive its adjustments in order to improve retention of material [9] as well as provide flexibility for the user.

At each time step t, representing the generation, practice, and evaluation of a new score, the user can manually adjust w_p as well as u_p, which denotes the user's confidence value, for each parameter p in [d, s, c]. Each value of u_p is then converted to a gradient, denoted by g_p, which determines the degree to which each difficulty parameter should be adjusted for a given generation. With each successive exercise, the system adapts to the user's learning goals, represented by adjustments in w, and outcomes, represented by values of u.

5. ADJUSTMENT ALGORITHMS

Variations and modifications are made using the following adjustments to yield different values for each parameter: rhythmic expansion or contraction for d, rhythmic transposition [6] for s, and drum set orchestration decisions, such as adding or removing voices and increasing or decreasing repetition of voices, for c. Each of these

b	d	s	c
1	0.25	0.833	0.375
2	0.25	0.667	0.375
4	0.25	0.278	0.167
8	0.25	0.0	0.125

Table 1: Differences in means of d, s, and c, corresponding to a change in b for the measure in Figure 2.

variations can either be created by using the original transcription as input or by a feedback mechanism in which the output of the variation process at step t is then used as input at step $t + 1$. The process of adjustment, which is applied on the scale of each individual bin, is continued until either a target difficulty τ has been reached or a certain number of time steps, adjustable by the user, has gone by without any change in D. The detailed adjustment processes for a single time step are as follows:

5.1 Density of Onsets (*d*)

A single onset from the bin, chosen at random and excluding the first onset in the bin, is removed if the bin contains more than one onset. Figure 3 depicts an example, with the original bin on the left and the two possible outcomes of adjusting *d* on the right.[3]

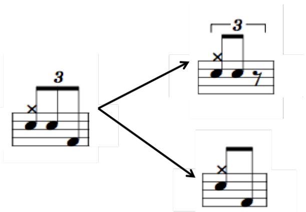

Figure 3: Possible results of adjusting for *d* in a bin.

5.2 Syncopation Value (*s*)

The first onset in the bin is shifted to the beginning of the bin and thus falls on a strong beat according to our syncopation measure. As a result, surrounding syncopations become anticipations, surrounding anticipations become hesitations, and surrounding hesitations are no longer syncopated at all, as seen in Figure 4.

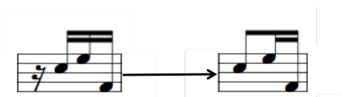

Figure 4: Example of adjusting for *s* in a bin.

5.3 Coordination and Interdependence (*c*)

An onset is chosen such that at least one of the following conditions is true:

I. At least one neighbor of the onset is played on a different surface (i.e. has a different pitch or notehead).

II. The onset is performed with one or more simultaneous onsets.

[3] Note that the bin on the bottom right has been converted to straight eighths in accordance with the jazz notation convention that triplets without the middle note are notated as straight eighths with the understanding that they should be performed swung.

If the onset satisfies condition I, then it is to be played on the surface of the differing neighbor. If both neighbors are different, one of the two is chosen at random. If the onset satisfies condition II, a random simultaneous onset is removed from the bin. If the onset satisfies both conditions, then one of the corresponding actions is chosen at random and applied to the onset. Figure 5 demonstrates the modification possibilities for an example bin when the second note of the bin is selected for adjustment.

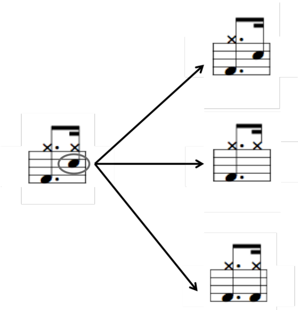

Figure 5: Possible results of adjusting for *c* in a bin, given the selected note (circled above).

For all parameters, if there is no possible adjustment that can be made to a bin, then that bin is returned without any modification. In addition, if the user passes in custom values for each stochastic modifier f_p, then each time an adjustment process is called, the program will use f_p to determine whether the process will actually be executed. The frequency of adjustment for a parameter p is a linear function of g_p.

6. EXAMPLE USING "BLACK NILE"

Elvin Jones' drum solo on the composition "Black Nile" from Wayne Shorter's 1964 album *Night Dreamer* [10] is a popular transcription choice for jazz drummers,[4] especially following educator John Riley's publication of his transcription of the solo [11]. Here the drum solo is used as an example input to SCATLAVA in order to demonstrate the musical output that the program generates. The following examples are selected outputs generated by the program and thus represent a subset of possible outputs given the input parameters used.

In this section SCATLAVA operates on the 4-bar excerpt of the drum solo, transcribed by the author, shown in Figure 6. Upon initial analysis, the phrase yields a dif-

Figure 6: First four bars of Elvin Jones' "Black Nile" solo.

[4] A YouTube search for "elvin jones black nile" yields over 800 results, with videos of other drummers performing transcriptions of Jones' solo comprising 10 of the 14 results on the first page.

Figure 7: SCATLAVA outputs (using default weights, gradients, and bin divisions unless noted): **(a)** $\tau = 0.2$, **(b)** $\tau = 0.2$, $b = 2$, **(c)** $\tau = 0.2$, $w_d = 0.1$, $w_s = 0.1$, $w_c = 0.8$, $g_d = 0.1$, $g_s = 0.1$, $g_c = 0.8$, **(d)** $\tau = 0.5$, **(e)** $\tau = 0.8$.

ficulty of $D = 0.573$. Default values of $w_d = 0.33$, $w_s = 0.33$, $w_c = 0.34$, and $b = 4$ are used, and the program is initially run with target difficulty $\tau = 0.2$. The resulting output can be seen in Figure 7a. While the contour of Jones' phrasing remains clear, the adjustments render the passage easier to interpret and perform. For example, the first three beats of measure 1 demonstrate the simplification that results in reducing onset density, while the fourth beat exemplifies the reduction in difficulty of both syncopation (the note has been moved from the last eighth note of the measure to the last quarter) and coordination (the simultaneous crash cymbal is omitted).

Figure 7b and 7c demonstrate the versatility of the SCATLAVA software. Perhaps the user would like to see the outline of the phrase on a higher level; by setting $b = 2$ instead of $b = 4$, the resulting output is even less dense than before, though it still maintains the motivic contour of the original phrase. Similarly, it is possible for a student to have little difficulty with onset density and syncopation but to struggle with interdependence. All the user has to do is enter his or her confidence values to reflect that, and the computed gradients will allow the appropriate adjustments to be made; Figure 7c shows the output for parameters $g_d = 0.1$, $g_s = 0.1$, $g_c = 0.8$, with b set to 4 once more. The weights of the parameters have also been adjusted to reflect the gradients. The resulting phrase bears more resemblance to the original passage, with fewer adjustments to density and syncopation, yet the coordination elements are clearly less challenging, and thus present a much lower difficulty to said user.

In general, by increasing the target difficulty we approximate the original transcription more and more closely. Figure 7d shows an example output for $\tau = 0.5$; the most noticeable difference between Figure 7d and Figure 7a is that the former has a higher density of notes. Similarly, Figure 7e shows the output for $\tau = 0.8$, which introduces even more activity across all parameters.

7. CONCLUSION

This paper documented the SCATLAVA software as a model for an adaptive learning environment for the education of jazz transcriptions through algorithmic variation and analysis, with an emphasis on rhythmic material. Together with a variety of user inputs, the software uses onset density, syncopation measure, and limb interdependence using CNID. The primary drawback of the software in its current state is the lack of support for melodic, harmonic, and timbral characteristics. Additional future improvements include more sophisticated machine learning methods to better infer and adapt to users' needs and skills, as well as a streamlined interface that implements methods for learning and playing by ear. Arrangements with jazz educators have been made to begin testing the system with students; this will yield user feedback as well as further results beyond the examples in the paper. With such insights and improvements, it is the author's hope that SCATLAVA will become a powerful yet intuitive software platform for augmenting and extending the tradition of studying transcriptions in jazz.

8. REFERENCES

[1] I. Sandoval Campillo, "In Your Own Sweet Way: A Study of Effective Habits of Practice for Jazz Pianists with Application to All Musicians," Universidad Autònoma de Barcelona, 2013.

[2] C. W. J. Chen, "Mobile learning: the effectiveness of using the application *iReal Pro* to practise jazz improvisation for undergraduate students in Hong Kong," Hong Kong Institute of Education, 2014.

[3] B. Keller, S. Jones, B. Thom, and A. Wolin, "An Interactive Tool for Learning Improvisation Through Composition," *Technical Report HMC-CS-2005-02*, Harvey Mudd College, 2005.

[4] V. Sébastien, H. Ralambondrainy, Olivier Sébastien, and N. Conruyt, "Score analyzer: automatically determining scores difficulty level for instrumental e-learning," *Proceedings of the 13th International Society for Music Information Retrieval Conference, ISMIR 2012*, Porto, 2012, pp. 571-576.

[5] M. Keith, *From Polychords to Polya: Adventures in Music Combinatorics*, Vinculum Press, Princeton, 1991.

[6] J. Riley, *The Art of Bop Drumming*, Alfred Music Publishing, pp. 17-21, 1994.

[7] P. Brusilovsky and C. Peylo, "Adaptive and Intelligent Web-based Educational Systems," *International Journal of Artificial Intelligence in Education*, vol. 13, no. 2-4, pp. 159–172, 2003.

[8] M. Nour, E. Abed, and N. Hegazi, "A Proposed Student Model Algorithm for an Intelligent Tutoring System," *IEEE Proceedings of the 34th SICE Annual Conference*, Hokkaido, 1995, pp. 1327-1333.

[9] P. Sadler and E. Good, "The Impact of Self- and Peer-grading on Student Learning," *Educational Assessment*, vol. 11, no. 1, pp. 1-31, 2006.

[10] W. Shorter, "Black Nile," *Night Dreamer*, Blue Note, 5:02-5:38, 1964.

[11] J. Riley, *The Jazz Drummer's Workshop*, Modern Drummer Publications, pp.43-55, 2005.