A Music Composition Model with Genetic Programming –A Case Study of Chord Progression and Bassline-

Kanae Kunimatsu, Yu Ishikawa, Masami Takata, and Kazuki Joe
Graduate School of Humanities and Science, Nara Women's University, Nara-City, Nara, Japan

Abstract - *In this paper, we present the improvement of our automatic music composition model using genetic programming, which is constructed with two independent but cooperative models for chord progression and melody. The improvements are the chord progression generation model and the bassline generation model. The chord progression generation model, considering tension notes, generates chord progression suitable for jazz blues. In the music theory, tension notes that are granted for the corresponding chords are determined for some cases but it is not easy to automatically determine what to apply from such theoretical codes. In this model, by focusing on the relation between the chord's top notes with tension notes and melody tones, we attempt to automatically generate chords. Further, by adding the bassline generating part, this system is applied to Jazz improvisation.*

Keywords: Genetic Programming, chord progression, bassline

1 Introduction

In recent years, as the spread of ICTs, a large number of users want to transmit their music contents by utilizing the CGM (Consumer Generated Media). There are composer support systems with a variety of DTM (Desk Top Music) software. The composer support systems generate music with just giving initial parameters such as keys, tempo and moods while others generate music with giving interactive selections of user's favorites during the music generation processes.

Since good music compositions require professional knowledge and experience, it is difficult for ordinal persons to compose music. To make them easily compose music, automatic music composition systems have been proposed. As methods of automatic music composition, a probabilistic method and an evolutionary computation method have been proposed. In this paper, we develop an automatic music generation system with the evolutionary computation that does not require professional knowledge or experience. Evolutionary computation is an engineering model with the mechanism of biological evolution to be applied to combinatorial optimization problem, machine learning and system analysis. Considering the automatic composition is basically an optimization problem of searching an optimal combination from the combinations of large number of musical notes, the evolutionary computation is suitable for the automatic composition.

As an automatic composition system using evolutionary computation, we have proposed the automatic music composition system with genetic programming (GP) [1]. It utilizes the genetic structure of GP, namely a tree structure, to associate the tree depth level with the length of a music note. In this framework, any length of music notes can be expressed for melody line and chord progression. In [1], we narrow down the music genre to blues and adopt a seventh chord in a chord progression part. However, the generated chords do not contain tension notes, which are absolutely imperative for expressing various sounds such as Jazz. Actual Jazz improvisation is performed with voicing, namely omitting chord tones or adding tension notes, to generate more various sounds. In this paper, we improve the chord progression generation model by adding chords that include tension notes and voicing the chords. Thereby, we improve the chord progression model so that the generated chords include tension chords and look like a real Jazz performance. Furthermore, in order to apply this system to Jazz improvisation, we propose the third model to generate bassline.

This rest of the paper is constructed as follows. In section 2, we explain tension notes and voicing. In section 3, we introduce the music composition model with genetic programming. In section 4, we improve the chord progression model and propose the bassline model. In section 5, we present some experiments to validate the model.

2 Characteristics of Blues Music

In this section, first we explain chord classification and blues chord progression. Next, we describe tension chord and voicing.

2.1 Chord classification

A chord is categorized into three functions.

- Tonic (T) : the basic note: beginning and ending
- Dominant (D) : notes for returning to Tonic
- Subdominant (SD) : notes for procession and connection: easy to transfer to Dominant

As above described, there is the basic chord progression of T-SD-D-T. T-D-T and T-SD-T are also natural progressions. In

Table 1: A musical scale plotted with I to VII.

	T	**D**	**SD**
Primary triad	I	V(7)	IV
Secondary triad	VI	VII	II

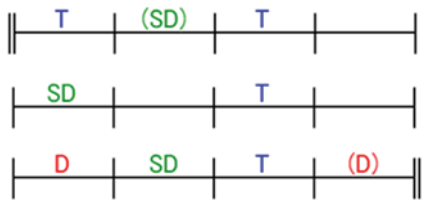

Fig. 1: Blues chord progression

Tab.1, a musical scale is plotted with I to VII. III is not presented because the property of III is not included.

2.2 Blues chord progression

Blues is a music format that consists of 12-bars, and is played with adding blue notes to the major scale to fit to the fixed chord progression, using only three of the four types of Seventh chord (dominant seventh, major seventh, minor seventh, minor major seventh). Figure.1 represents a blues chord progression. In harmonics of traditional western music, the chord progression of D-SD is not used because it is considered as unstable sound. But it is included between the 9^{th} and 10^{th} bars in the blues chord progression. When the 11^{th} and 12^{th} bars are T, it means that the song is finished.

2.3 Adding tension notes

A tension note is a note of non-harmonic tones to give a sense of tension in the chord sound without interfering the chord progression. A chord with tension notes is called a tension chord. In order to provide effective tension notes, voicing should be considered at the same time as the melody. The applicable tension notes to a chord are decided by music theory but it is difficult to make a decision which tension notes are actually used. Dissonant halftone collision against melody sound is an improper harmony to be avoided. In our model, the code generation is performed so that the relation between the top note of the tension embedded chord and the melody sound is considered.

2.4 Voicing

When a chord is given, the operation of arranging the composition tones of the chord in each voice part is called voicing that presents very colorful musical expressions. Especially in popular music such as jazz, the use of tension notes gives a tension to the chord sound to perform the effect of voicing. Adding a tension note to chord tones, while keeping the harmonic function of the chord, it is possible to express a colorful sound. There are two methods for voicing: 1) close voicing, which takes a dense placement of the chord tones within an octave and 2) open voicing, which take a wide placement beyond an octave. Any number of voicing patterns for a single chord is possible and the number of voicing patterns with tension notes increases explosively. In this paper, we just focus on a voicing technique called spread voicing, which is a kind of open voicing, for chord generation.

2.5 Spread Voicing

The spread voicing takes placement of the chord tones beyond an octave to keep the root of the chord as the bottom note. Generally used in jazz piano, since the harmonic tones are arranged in a wide range, it presents more effective and richer harmony tones to the melody. As the chord progression generation part, we adopt the chord generation using a 4-way spread voicing method.

3 A Music Composition Model with Genetic Programming

In this section, we explain the music composition model with genetic programming [1]. The composition model consists of two parts: chord progression and melody. In general there are two methods for composing music. One method is melody first to apply chords in accord with the melody while the other method is chord first to create melody using the chord. The composition model adopts the latter method provided that tempo and accent are not changeable.

To generate good music that is not confined by existing ones, both parts make use of genetic operators. To represent any length of notes such as a triplet and dotted one within each individual, we adopt GP to express complicated data such as tree structures rather than GA that is based on array structures. Figure 2 shows an example of genetic individuals. The depth level of the tree structure represents the length of the notes. The deeper the depth level is, the finer note it represents. Figure 3 gives examples of a binary tree and a triplet tree. In the case of the binary tree, the note length in the upper level is as twice as in the lower level. In the case of the triplet tree, the note length in the upper level is as 3 times as in the lower level. Until it meets certain criteria, genetic operators such as crossover or mutation are repeated to develop next populations. Each individual gene represents a chord progression and a melody line with the length of a fixed bar. Each node has the number of branches in the case of non-terminal nodes or in the case of terminal nodes the type of chord (chord progression) or the sound information (melody) such as sound, keyboard continuation symbol and rest. Using depth-first search, just the terminal nodes that contains the sound information are detected from left to right and it is possible to replace old gene individuals of chord progression and melody with new ones. Thus, it generates a melody line based on the chord progression by bar unit.

Although the use of GP eliminates the limit of the note length such a triplet or dotted note, it is difficult to maintain the similarity and the continuity of the new melody bar with the previous bar. When music is generated, we use some restrictions to retain the music similarity and the continuity from the past by utilizing the characteristics of the chord progression. In this model, we squeeze the music genre blues, and evaluate the generated music in terms of the blues characteristics. The number of bars of a song is set to 12 for one individual suitable for the blues. Based on the first 12 bars of a blues song, a partial chord progression is generated by selecting a good individual in the chord progression generation model. Then, a partial melody line is generated by selecting a good individual in the melody generation model.

The partial chord progression i is evaluated with the fitness using the following items.

I. Percentage that satisfies the chord progression pattern of blues

II. Integrity with the melody notes in each bar of the partial melody *i-1*

Item I is used for adapting to the blues chord progression pattern shown in Fig.1. Item II is used for imitating the i^{th} 12 bars with the $i\text{-}1^{th}$ 12 bars. The first note of the melody in each bar among partial melody *i-1* and the chord of each leaf node among partial chord progression *i* are examined to calculate the rate if the first note is included in the chord.

The partial melody *i* is evaluated with the fitness using the following items.

1. Integrity with the partial chord progression *i*

2. Comparison of the music entropy function with the partial melody *i-1*

3. Comparison of the rhythm patterns with the partial melody *i-1*

Item 1 is used for generating the partial melody *i* suitable for a given partial chord progression *i*. The notes of each bar in partial melody *i* are examined to calculate the rate if they are included in the chords of each bar in the partial chord progression *i*. Item 2 and 3 are used so that the impression between the partial melody *i* and the partial melody *i-1* does not change abruptly. Each comparison is performed by bar unit. In item 2, we propose the concept of impression as music entropy to compare the difference of impression as numeric values. Using the transition probability of the partial melody *i-1*, the music entropy is calculated by bar unit to quantify the occurrence degree for each note. We assumed that the impression is similar when the music entropy values are close. In item 3, the average of note values in a bar is calculated to compare the rhythm intervals.

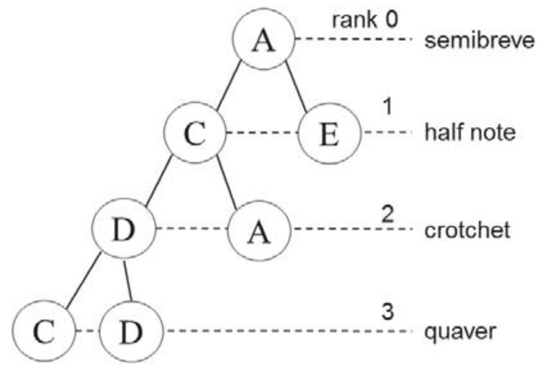

Fig. 2: An example of genetic individuals.

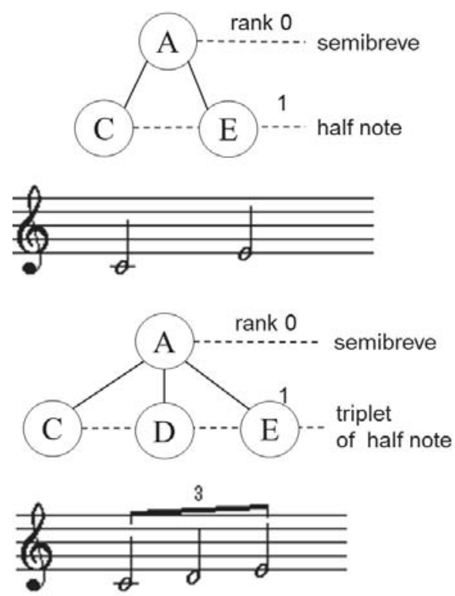

Fig. 3: Difference of expression by the number of branches.

Through the above fitness evaluations, the partial chord progression *i* and the partial melody *i* get involved in the partial melody *i-1* and the partial chord progression *i-1*, respectively. Since we focus on blues in this paper, each partial chord progression and melody *i* consist of 12 bars as explained in 2.2. Note that this kind of interaction is found in real Jazz blues improvisation.

4 Improvement of the music composition model

In this section, we explain the improvement of the chord progression generation part and the newly developed bassline generation part. Figure 4 shows the overview of the whole model. GP is applied in each part to generate melody, chord progression and bassline in this order.

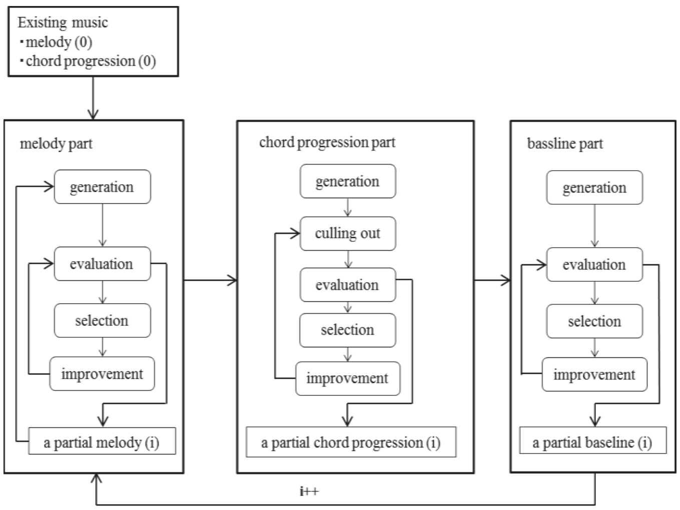

Fig. 4: Overview of the model.

Table 2: An example of 4-way spread voicing

	pattern1	pattern2
Top note	A tension note	A chord tone
Middle notes	7th	3rd
	3rd	7th
Bottom note	Root	

4.1 Chord progression part

In the chord progression generation model, a voicing is applied to the three chords used in the blues form to produce various types of chords. In order to generate the top note of the chord so that it is suitable for the melody, this model generates the melody first.

4.1.1 Information of individuals

In this model we perform a 4-way spread voicing: the root as the bottom note, the combination of 3^{th} and 7^{th} as the middle two notes and a tension note or a chord tone as the top note. Table 2 shows an example of a 4-way spread voicing. A candidate for the tension note includes minor 9^{th}, 9^{th}, major 9^{th}, 11^{th}, major 11^{th}, minor 13^{th} and 13^{th}. The notes that constitute the minor 9^{th} upwardly from each note of chord tones are called avoid note and often avoided from the use of tension notes. However it is sometimes intended to be used depending on the type of the target chord. For example, in the case of dominant seventh chords, although the note of 11^{th} should be avoided as an avoid note, other notes (minor 9^{th}, 9^{th}, major 9^{th}, major 11^{th}, minor 13^{th}, 13^{th}) can be used as a tension note because they do not interfere in the function of chords. In this chord progression generation model, voicing is applied to the three chords used in existing music to get candidates for terminal nodes, and 22 (11 (tension notes or chord tones) × 2 (Voicing patterns)) types available for a single chord are prepared to perform a 4-way spread voicing as shown in Table 2. In general, typical chord progression changes in a half note or longer when the theme is presented. However, in the case of improvisation, the chord progression becomes often-complicated and finer. In this paper, we limit the chord progression change by a half note.

4.1.2 Fitness evaluation

To improve the chord progression generation part explained in the previous section, we add the following two items to calculate the fitness for the partial chord progression *i*.

III. Comparison of the top notes of the chords and the melody in each bar of the partial melody *i*

IV. Comparison of the entropy function with the partial chord progression *i-1*

Item III is used for the comparison of the top notes, including the tension notes of chords for each bar, with the melody line. The aim of the evaluation is to avoid the dissonant semitone collision with tension notes and the melody as much as possible. When the first note of the partial melody *i* and the top note of the chord are in contact with major 2^{nd} and minor 2^{nd} at a chord change time, the fitness value is decreased. When an avoid note is selected for the top note, the fitness value is decreased, too.

Item IV is used for preventing abrupt changes in the impression between the partial chord progression *i* and the partial chord progression *i-1*. In chord voicing, the top note should be selected so that the continuous musical progression is observed. In the chord progression generation of this model, we compare the music entropy values for the top notes of the chords to avoid abrupt change of the music impression, and make the successive bars enough similar.

Let a set S be $\{X_1, X_2, \cdots, X_{14}\}$ that represent 12 kinds of top notes of a chord, continued and rests, namely {*C*, *Cs*, *D*, *Ds*, *E*, *F*, *Fs*, *G*, *Gs*, *A*, *As*, *B*, ***, *~*}. In addition, let the probability of the occurrence of an event X_k under the condition in event X_l occurrence be Y_{kl}. The music entropy is defined as follows.

$$\text{Music Entropy} = -\sum_{kl=2}^{nN} Y_{kl} \times \log Y_{kl} \qquad (1)$$

As the characteristic of the blues form, the tonic chord in the first four bars out of the 12 bar (1 chorus) is highlighted, then the next four bars are transited from the subdominant chord to the dominant chord, and finally the last four bars are transited from dominant chord to the tonic chord. Therefore, it has a structure with three parts, each of which contains 4 bars. From the blues structure, the value of the music entropy in the chord progression generation part is calculated every 4 bars to compare the value with the corresponding 4 bars section in the previous partial chord progression. Note that the partial chord progression with music entropy 0 means that the chord progression does not change from the previous chord progression at all. On the other hand, the partial chord progression with music entropy 1 means that the chord progression consists of completely random notes; it is more free than “free jazz”.

From a given original blues song, the partial chord progression *0* is generated without any tension notes and each bar contains a single chord. In the partial chord progression *1*, the subdivision of each chord is performed to select the individual such that the total of the transition probabilities Y_{kl} between the chords included in the 4 bars becomes larger and the top notes of the chords walk smoothly. At the same time the music entropy of each 4 bars is calculated to compare the music entropy of the next partial chord progression.

4.2 Bassline part

In the bassline generate model, the 4-beat bassline often used in jazz is generated. To generate a bassline suitable for the chord progression, the bassline is generated right after the chord progression generation.

4.2.1 Information of individuals

Since we require the 4-beat bassline for jazz improvisation, generated bassline individuals are fixed to crotchet. Therefore the number of branches of a non-terminal node in the depth level below 0 is just 2 as in the same case of the chord progression generation part. The individual of terminal nodes includes the notes of the key range, hold (~) and rests (*).

4.2.2 Fitness evaluation

The partial bassline *i* is evaluated with the fitness using the following items.

A. Integrity with the partial chord progression *i*

B. Comparison of the music entropy function with the partial melody *i-1*

Item A is used for generating the partial bassline *i* suitable for a given partial chord progression *i*. The notes of each bar in partial bassline *i* are examined to calculate the rate if they are included in the chords of each bar in the partial chord progression *i*. In each bar, when the first beat and the third beat match the chord tones, the fitness value is highly increased. The fitness value is decreased when a note of the

Fig. 5: Original melody

bassline is in contact with a corresponding chord tone by a semitone. Item B is used so that the impression between the partial bassline *i* and the partial bassline *i-1* does not change abruptly. This evaluation is performed every 4 bars unit.

In the partial bassline *i*, individuals are selected so that the total of the transition probabilities between the bassline notes included in the 4 bars becomes larger and the notes walk smoothly. At the same time the music entropy of each 4 bars is calculated to compare the music entropy of the next partial bassline.

5 Experiments

In this section, we report the experimental results about chord progression generation and baseline generation.

5.1 Experimental environments

In our model, we perform a generation of music derived from the previous 12 bars every 12 bars, to be close to the method of blues improvisation. Therefore, we use an existing music as partial melody 0 and partial chord progression 0. In this experiment, we use a famous standard jazz number with the blues format, Billie's Bounce by Charlie Parker. Figure 5 shows the first 12 bars of Billie's Bounce. Each part of melody, chord progression and bassline is represented as a gene individual with 12 bars; the number of non-terminal nodes in the depth level (*level*) 0 is fixed to 3 to generate three blocks of 4 bars in *level* 1. Thus, it is easier to compare the partial melody, the partial chord progression and the partial bassline during fitness evaluation. Terminal nodes are set to be generated in *level* 4 or deeper to generate the 12 bars.

For terminal nodes of chord progression, we prepare voicing chords by adding tension notes or chord tones to existing three chords (C7, F7 and G7). Since typical chord progressions vary from two beats to one bar, the early genes individual depth level is set to 4; the early genes individual consist of half notes in this experiment. The parameters we used for the experiments are the number of individuals (*nI*), crossover probability (*c*), mutation rate (*m*), and the maximum generation number (*Generation*). These parameters are not theoretically decided but empirically found. In [1], the preliminary experiments for determining the numbers of average individuals and average generations are reported that good gene individuals among 100 individuals are mainly selected by the 50th generation. The crossover probability and the mutation rate are set to the same values in [1]. Thus, we adopt the following parameter choices for the chord progression generation part.

- *nI*=100
- *c*=0.8
- *m*=0.1
- *Generation*=50

For terminal nodes of the bassline generation part, we prepare 17 kinds of nodes that consist of 15 notes for the target (2E - 3G) range, a hold and a rest. The initial gene individuals are adjusted to the depth level of 5. In addition good gene individuals are selected by the 200th generation from our preliminary experiments for bassline. Thus, we adopt the following parameter choices for the bassline generation part.

- *nI*=100
- *c*=0.8
- *m*=0.1
- *Generation*=200

5.2 Experiment results

Figure 6 shows an example of the partial score 1 (the partial melody 1, the partial chord progression 1 and the partial bassline 1). Figure.7 presents an example of the partial score 2. The first stage shows the melody score, the second and third stages show the chord progression score, and the fourth stage shows the bassline score.

In the chord progression generation part, all of generated chords consist of half notes that are the same to the initial individual because the chord progression generation part does not perform note length comparison like the melody part. As for top notes, the chord progression includes chords with tension notes. Comparing the top notes of the generated chords and the melody tones in each bar, the notes in contact with minor 2nd or major 2nd are relatively selected out. However several sound conflicts between tension notes of the chords and the melody tones are observed. So several improper chords are generated. A possible reason is that the target melody tone to be compared with the chord top notes is just focused on a melody tone when the chord progression changes. In the partial chord progression 1, the top notes are selected such that they do not comparatively create a large movement, and any unnatural leap is not observed. In the partial chord progression 2, since it performs a comparison with the music entropy calculated in the partial chord progression 1, the generated top notes are in a natural flow.

Fig. 6: An example of partial score 1.

Fig. 7: An example of partial score 2.

The bassline generates only crotchets and four beat based individuals are generated. The first beat and the third beat are generated from the notes of the corresponding chord tones, and in particular the first beat is mainly generated from the individuals with the root of the chord. The second beat and the fourth beat are often generated with tension notes to collide in a semitone with chord tones and generate dissonant sounds.

From the experiments, in the case of the chord progression we observe that tension chords are generated by selecting out the sounds with semitone collisions and comparing the music entropy for the chord top notes at the fitness evaluation. In the case of the baseline generation, we observe that sounds suitable for chord tones are successfully generated. However, each evaluation formula interferes with each other. Therefore, any tone column to satisfy all the evaluation formula is not generated. In particular, since the chord progression and the bassline tend to generate the sounds causing semitone collision that are not completely selected out, it is necessary to improve the evaluation formula.

6 Conclusions

In this paper, we proposed a chord progression generation model considering tension notes and a baseline generation model using GPs. Since target music is blues, each part for individuals is fixed to 12 bars.

The chord progression generation model generates chords including tension notes suitable for the given melody with applying selection methods based on the specific chord progression blues, and select out the sounds where the chord top notes and melody tones become dissonant. The bassline generation model generates 4^{th} note based bassline suitable for the chord progression by fixing the rhythm to crotchet, and evaluates the integrity with the corresponding chord tones. In addition, both the chord progression model and the bassline model adopt an evaluation method for music impression by music entropy reported in [1]. Thus, we prevent abrupt changes during the top notes of each chord and the baseline. In the generation of the partial chord progression 1 and the partial baseline 1, we select the individuals with high transition probabilities between chords or baseline that is defined in the music entropy. In the partial score 1, thereby individuals such as top notes of chords and the bassline show natural flows.

Our future work includes the improvement of evaluation formula and multipurpose GPs because it is not possible to completely select out the sounds that cause semitone collision in our current models. Furthermore, we aim to create a system to perform a jazz session with human in real-time. Thus, we would like to apply our future system to real jazz improvisation.

7 References

[1] K.Komatsu, T.Yamanaka, M.Takata and K.Joe, “A Music Composition Model with Genetic Programming,” in *Proc. PDPTA’10*, II, pp.686-692, 2010.

[2] J.A.Biles, “GenJam: A Genetic Algorithm for Generating Jazz Solos,” In Proceedings of the 1994 International Computer Music Conference, ICMA, SanFrancisco, pp.3-4, 1994.

[3] T.Kitahara, M.Katsura, H.Katayose and N.Nagata, “Automatic Chord Voicing System Using Bayesian Network,” IPSJ journal, Vol.50, No.3, pp.1067-1078, 2009.

[4] N.Emura, et al., “A modular system generating Jazz-style arrangement for a given set of a melody and its chord name sequence,” Acoust.Sci.&Tech., Vol.29, No.1, pp.51-57, 2008

[5] Noteflight-Online Music Notation , http://www.noteflight.com/.