

INTERACTIVE GENJAM: INTEGRATING REAL-TIME PERFORMANCE WITH A GENETIC ALGORITHM

John A. Biles
Rochester Institute of Technology
102 Lomb Memorial Drive
Rochester, NY 14623-5608
<http://www.it.rit.edu/~jab/>
jab@it.rit.edu

Abstract

This presentation will describe and demonstrate recent enhancements to GenJam, an interactive genetic algorithm jazz improviser. The most significant enhancement incorporates a pitch-to-MIDI capability, which allows GenJam to integrate human improvisations into its own improvisations. Specifically, when GenJam “trades fours” with a human, it listens to the human’s last four measures, maps what it hears to its chromosome representation, mutates the chromosomes, and plays the result as its next four. In other words, GenJam evolves what it “hears” into what it will play in real time. Other recent enhancements are also discussed.

1 Introduction and Background

GenJam is an interactive genetic algorithm that models a jazz improviser and performs as a featured soloist in the author’s Virtual Quintet. Previous papers [Biles 94, Biles 96] have described GenJam's hierarchically related populations of melodic ideas, its chromosome representations for those ideas, its genetic operators for evolving new ideas, and the training of new soloists. This training is done under the guidance of a human mentor, who listens to GenJam improvise and indicates "good" or "bad" whenever so moved. The mentor's feedback is used to increment or decrement the fitness of individual melodic ideas and serves as the environment in which those ideas survive or die off. New ideas evolve by selecting the “better” ideas to be parents, breeding children using single-point crossover and musically meaningful mutation, and replacing the “worse” ideas in the population with these new children.

This paper describes several recent enhancements, chief among them an interactive capability that allows GenJam to “trade fours” with a human improviser by listening to what a human plays and evolving what it thought it “heard” into what it plays in response. The subtitle of this paper, then, has two interpretations. The more obvious meaning is “the addition of real-time interactive performance to GenJam’s capabilities.” The more subtle meaning is “the use of genetic algorithms as a paradigm for accomplishing real-time interactivity.” This second interpretation refers to the approach taken in making GenJam interactive, which was to reuse as much of GenJam’s genetic infrastructure as possible. By using GenJam’s existing chromosome structures, mutation operators and harmonic knowledge, we were able to build a very robust system that successfully confronts the notorious pitch-tracking problems of off-the-shelf pitch-to-MIDI products.

2 Interactivity via Real-Time Evolution

Figure 1 shows the high-level algorithm GenJam executes as it trades fours with a human player. The target of the pitch-to-MIDI activity is GenJam’s chromosome structure for representing melodic ideas, which we term GenJam Normal Form (GJNF). Once the human’s four has been mapped to GJNF, it may be mutated using any of GenJam’s musically meaningful mutation operators, and the result is guaranteed to be playable over any four-bar chord sequence. By taking advantage of this highly robust representation, most of the typical pitch-to-MIDI problems evaporate, and the resulting system is highly fault-tolerant. This makes it feasible to use an off-the-shelf

Human plays four bars into cheap microphone plugged into Roland GI-10 pitch-to-MIDI converter GI-10 sends MIDI events to GenJam running on host computer <i>via</i> Yamaha MU80 tone generator GenJam listens to MIDI events and builds chromosomes for a phrase and its four measures in GJNF Chromosomes are mutated in last 1/32 note (roughly) of human’s four Mutated chromosomes are used to generate GenJam’s next four as if they were part of a stored soloist

Figure 1. GenJam algorithm for trading fours

pitch-to-MIDI converter, the Roland GI-10 in this case. Although the GI-10 was designed for use with a guitar, the inclusion of a microphone input allows its use with acoustic instruments like the author's trumpet.

The rest of this section will focus on an example to help illustrate how GenJam evolves a human's four into its response as it progresses through the steps in Figure 1. The example comes from measures 25-32 of Jerome Kern's *All the Things You Are*. In the example, the human played the first four bars of the Charlie Parker line *Prince Albert* over measures 25-28 (shown in Figure 2). Even though this four is not spontaneous, it was chosen because it should be familiar to jazz listeners, presents some interesting rhythmic challenges for the pitch tracker, and because GenJam evolved it into a nice four played over bars 29-32 of the tune. The four bars that the human played are notated in Figure 2 with the chords played by the rhythm section indicated for each measure. The notation in this and later figures does not try to capture a swing interpretation of the eighth notes.



Figure 2. *Prince Albert* quote played over measures 25-28 of *All the Things You Are*

Table 1 gives the chord-scale mappings used both during the listening phase and in the playing phase. GenJam builds these maps from the chord progression for the tune, which it reads from a data file. The maps are used as lookup tables, where an event in a measure chromosome is used as an index into the table, and the contents of the table is the corresponding pitch. During the listening phase, incoming MIDI pitches are searched for in the table, and the index of the closest match is used as the event in the chromosome. During the playing phase, the event from the chromosome is used as an index into the table, and the pitch found at that location is used for the MIDI note-on message sent to the tone generator.

Bar	Chord	Scale	Pitches for new-note events 1-14
25	Fm7	Hexatonic Minor (avoid 6th)	C Eb F G Ab Bb C Eb F G Ab Bb C Eb
26	Bbm7	Hexatonic Minor (avoid 6th)	C Db Eb F Ab Bb C Db Eb F Ab Bb C Db
27	Eb7	Hexatonic Mixolydian (avoid 4th)	C Db Eb F G Bb C Db Eb F G Bb C Db
28	AbMaj7	Hexatonic Major (avoid 4th)	C Eb F Gb Ab Bb C Eb F Gb Ab Bb C Eb
29	DbMaj7	Hexatonic Major (avoid 4th)	C Db Eb F Ab Bb C Db Eb F Ab Bb C Db
30	Gb13	Hexatonic Mixolydian (avoid 4th)	Db Eb Fb Gb Ab Bb Db Eb Fb Gb Ab Bb Db Eb
31	Cm7	Hexatonic Minor (avoid 6th)	C D Eb F G Bb C D Eb F G Bb C D
32	Bdim	Whole/Half Diminished	D E F G Ab Bb B Db D E F G Ab Bb

Table 1. Chord-scale mappings for measures 25-32 of *All the Things You Are*

When listening to a four, GenJam quantizes each measure into 8 eighth-note-length windows, one window for each locus in a measure chromosome. Figure 3 shows the chromosomes in GJNF resulting from GenJam's listening to the *Prince Albert* quote. The left column represents the phrase chromosome, which contains pointers to four measure chromosomes, in this case measures 0-3. The right column contains the four measure chromosomes (0-3), one measure per row, with eight loci corresponding to the eight windows in each row. Locus values of 0 are rest events, which are played by sending a note-off message to the tone generator. Locus values of 15 are hold events, which are played by sending no messages (holding the previous event). Locus values from 1-14 are new-note events, which are played by sending a note-off message followed by a note-on message, using the pitch found at the event's location in the appropriate scale, as described above. GJNF has the advantage of unifying pitch and rhythmic structures and results in four-bar phrases that can be played in any harmonic setting.

0	9	10	11	12	13	11	10	9
1	14	14	13	14	13	12	12	11
2	11	11	10	11	15	12	12	14
3	15	0	12	11	10	9	15	0

Figure 3. Chromosomes in GJNF built from listening to the *Prince Albert* quote

Figure 4 shows the chromosomes of Figure 3 played against the original chords of measures 25-29. Comparing Figures 2 and 4 reveals some of the inaccuracies in mapping the *Prince Albert* quote to GJNF. Some of these inaccuracies are due to the fact that GenJam can represent only eighth-note multiples, which is an obvious limitation in handling the triplets and sixteenth notes in the second measure of Figure 2.



Figure 4. Traditional notation of GJNF in Figure 3

Other inaccuracies are due to pitch-tracking and quantization errors. All the “correct” pitches were transmitted by the GI-10, but they were often obscured by spurious notes, some of which were repetitions of correct notes and some of which were “chromatic passing tones” likely due to slurred articulation. A full-blown analysis of the MIDI signals sent by the GI-10 reveals that the 32 notes played by the author resulted in 53 note-on/note-off pairs, with three windows, none of which were in the measure containing the triplets and sixteenths, receiving four pairs. The heuristic used by GenJam to cope with extra notes is simply to select the last note-on that occurs in a window and to ignore note-offs in a window once a note-on has been received. The only way, then, that a rest event (0) can occur is if only a single note-off event is received in its window. If no MIDI events at all are received in a window, its corresponding locus in the measure chromosome will remain a hold event (15), which is the initialized value.

In the last 30 milliseconds of the human’s four, GenJam stops listening and performs musically meaningful mutations on some of the chromosomes in preparation to play them back as its next four. The available mutations on measure chromosomes include (1) reverse - play the loci in reverse order, similar to retrograde, (2) rotate - rotate the loci a random amount from 1 to 7 positions to the right, (3) invert - subtract the locus value from 15 and rescale the result to the pitch range of the original measure, and (4) transpose - raise or lower the new note events by a random amount. The available mutations on phrases include (1) reverse - play the measures in reverse order, (2) rotate - rotate the measures a random amount from 1 to 3 positions to the right, (3) repeat - select a random measure and repeat it, replacing the measure that would have been played with the repetition, and (4) sequence phrase - build a special phrase beginning with the last measure of the human’s four, repeating that measure one or two more times and filling out the remainder of the phrase with other measures from the human’s four.

In the example, two random mutations were performed: (1) the phrase chromosome was rotated three positions to the right, and (2) measure 0 was transposed down two scale degrees. In addition, a heuristic was applied to the repeated 12 in measure 2, changing it to a 13 to make it a passing tone between its two neighbors, 12 and 14. The resulting chromosomes are summarized in Figure 5, where the left column is the mutated phrase chromosome and the right column is the reordered measures with the mutated loci in bold italic.

<i>I</i>	14	14	13	14	13	12	12	11
2	11	11	10	11	15	12	<i>I 3</i>	14
3	15	0	12	11	10	9	15	0
0	7	8	9	<i>10</i>	<i>11</i>	9	8	7

Figure 5. Mutated measure chromosomes in order used to generate GenJam’s four

Finally, the mutated chromosomes are played against the chords of measures 29-32, and the resulting four, in standard notation, is presented in Figure 6. Notice that the repeated 14’s and 12’s in measure 1 and the repeated 11’s in measure 2 became chromatic passing tones. The heuristic here is to look for eighth notes that repeat either the immediately preceding and/or succeeding pitch and alter them to be a half step above or below the next note.



Figure 6. GenJam’s four played over measures 29-32

3 Other Enhancements

In earlier versions of GenJam, the initial populations of measure and phrase chromosomes were generated using a simple uniform random number generator. This produced soloists that sounded pretty dismal in early generations, and detracted from audience-mediated performance situations where the audience serves as a collective mentor to train their own soloist using feedback paddles [Biles 95]. To confront this issue, I developed a simple fractal generator, which generates initial populations that statistically resemble mature, well trained populations. This fractal generator is similar to the dice-based fractal generator proposed by Martin Gardner [Gardner 78] and is used to generate new-note events for measure chromosomes.

The fractal generator uses two “dice,” one with 0-7 spots and the other with 1-7 spots. Summing the dots on these dice yields numbers in the range 1-14, which is the required range for new-note events. After the first number is generated by rolling both dice, successive numbers are generated by rolling only one of the dice and alternating which die is rolled from one number to the next. The frequency distribution of these numbers is a ramp function that peaks in the middle (7 and 8), and the interval distribution for successive numbers skews toward small intervals, with an average of around 2 and a maximum interval of 7. In contrast, a uniform generator yields a frequency distribution that is flat and an interval distribution that averages about 7. Musically, we can say that the fractal generator tends to pick notes in the middle of the instrument’s range, with intervals that average roughly a third and the maximum interval roughly an octave. This is typical for a mature, trained soloist, which means that the mentor(s) are more likely to hear rewardable moments in early generations, thereby speeding up the training process.

Another enhancement to GenJam has been the gradual addition of more chord families to GenJam’s harmonic knowledge base. As GenJam’s repertoire has grown to its current size of 120 tunes, new chords have been added to handle increasingly esoteric harmonic situations. GenJam currently recognizes 17 distinct chord families, which are listed in Table 3 with their associated scales, assuming a chord root of C natural. Since GenJam is a strictly vertical player and maps chords to scales in a simple, context-free way, the scales selected are “safe” in that they avoid notes that may be “inappropriate” in some contexts. For example, the use of a hexatonic major scale with no fourth for tonic major chords avoids the decision of whether to use a natural or a Lydian fourth.

Chord	Scale	Chord	Scale
CMaj7, C6, C	C D E G A B	C7#9	C Eb E G A Bb
C7, C9, C13	C D E G A Bb	C7b9	C Db E F G Bb
Cm7, Cm9, Cm11	C D Eb F G Bb	CmMaj7	C D Eb F G A B
Cm7b5	C Eb F Gb Ab Bb	Cm6	C D Eb F G A
Cdim	C D Eb F Gb G# A B	Cm7b9	C Db Eb F G A Bb
C#5	C D E F# G# A B	CMaj7#11	C D E F# G A B
C7#5	C D E F# G# A#	C7sus	C D E F G A Bb
C7#11	C D E F# G A Bb	CMaj7sus	C D E F G A B
C7alt	C Db D# E Gb G# Bb		

Table 3. Chord-scale mappings

In summary, the use of GJNF as the target representation for pitch-tracking has led to a very robust interactive system. Indeed, the inevitable errors made in pitch-tracking are desirable since they serve to “develop” rather than misrepresent what the human plays. Any time you can turn a bug into a feature, you’re on the right track!

References

- [Biles 94] John A. Biles. GenJam: A Genetic Algorithm for Generating Jazz Solos. In *Proceedings of the 1994 International Computer Music Conference, ICMA*, San Francisco, 1994.
- [Biles 95] John A. Biles and William Eign. GenJam Populi: Training an IGA via Audience-Mediated Performance. In *Proceedings of the 1995 International Computer Music Conference, ICMA*, San Francisco, 1995.
- [Biles 96] John A. Biles, Peter G. Anderson, Laura W. Loggi. Neural Network Fitness Functions for a Musical IGA. In *Proceedings of the International ICSC Symposium on Intelligent Industrial Automation (IIA’96) and Soft Computing (SOCO’96)*, March 26-28, Reading, UK, ICSC Academic Press, pp. B39-B44.
- [Gardner 78] Martin Gardner. White and brown music, fractal curves and one-over-f fluctuations. *Scientific American*, 238(4), pp. 16-27, 1978.