An Evolutionary Approach Applied to Algorithmic Composition

Jônatas Manzolli – jonatas@nics.unicamp.br
University of Campinas – Interdisciplinary Nucleus of Sound Communication – UNICAMP/NICS
Artemis Moroni – artemis@ia.cti.br
Technological Center for Informatics – The Automation Institute - CTI/IA

Fernando Von Zuben – vonzuben@dca.fee.unicamp.br
University of Campinas – Electric Engineer Faculty - UNICAMP/FEE

Ricardo Gudwin – gudwin@dca.fee.unicamp.br
University of Campinas – Electric Engineer Faculty - UNICAMP/FEE

Abstract

This paper presents an end- user interface that allows real time parametric control of sound events resulting in an interactive environment, in which Evolutionary Computation is applied to Algorithmic Composition. The resulting system, Vox Populi, uses genetic algorithms to generate and evaluate a sequence of chords played as MIDI data. Harmonic, tonal and voice range fitness are used to control musical features. Based on the ordering of consonance of musical intervals, the notion of approximating a sequence of notes to its harmonically compatible note or tonal centre is used. This method employs fuzzy formalism and is posed as an optimisation approach based on factors relevant to hearing music.

1 Introduction

Evolutionary Computation has been successfully applied to control music processes. An application of genetic algorithms to generate Jazz solos is described in (Biles, 1994) and this technique has also been studied as a way of controlling rhythmic structures (Horovitz, 1994).

This paper describes a procedure for algorithmic composition based on the controlled production of chord cadences: a population of chords is properly codified according to the MIDI protocol, and is submitted to evolution by the application of genetic algorithms. The research presented here was already discussed in (Moroni et all, 1999). A fitness criterion is defined to indicate the best chord at each generation, and this chord is selected as the next element in the sequence to be played. The produced real time sound output stream allows this system to be used in live electronic music as a computer performance instrument.

In what follows, a general description of the main components of the computational environment are presented and a musical fitness based on melodic, harmonic and voice range criteria is defined. Finally, the end-user interface is described in detail.

2 DEFINITIONS

Some basic concepts are fundamental to the understanding of the main results.

3 Individuals and Populations

Individuals of the population are defined here as chords of four notes and they are potential solutions for the selection process. Initially, the chord's notes are randomly generated on the interval [0...127], which corresponds to MIDI note number. In each generation, a population of 30 chords is produced and evaluated. The chords are internally represented as a chromosome with 28 bits, composed of 4 words with 7 bits (Fig. 1).

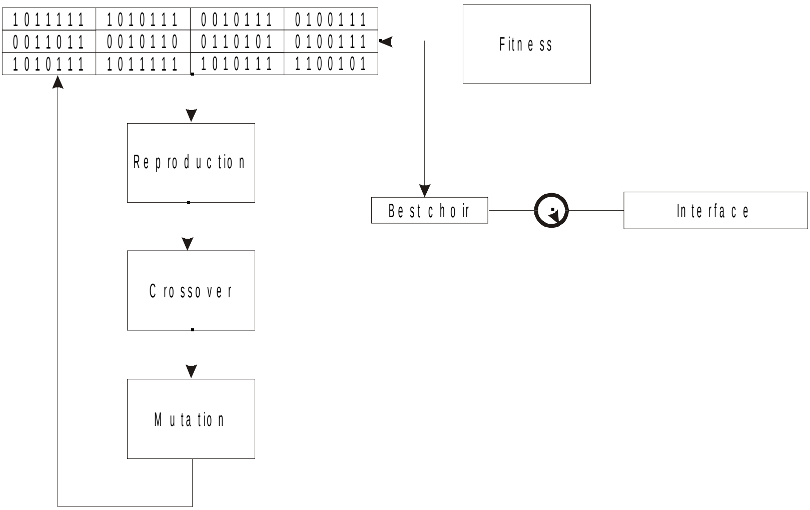
1011111	1010111	0010111	0100111

Figure 1 - The structure of a MIDI chromosome

4 Rhythmic Genetic Cycle

The general architecture of the rhythmic genetic cycle is depicted in **Fig. 2**. It is possible to see on the diagram two co-operative processes in the genetic cycle: one producing notes and the other (the interface) consuming notes. As described in 2.1, once the initial population of individuals is created, the fitness of each chord can be evaluated. The fitness function is defined as a composition of three terms: voice range fitness, vertical consonance or harmonic fitness, horizontal consonance or tonal fitness.

After the music fitness evaluation, typical genetic operators of crossover and mutation are applied to the individuals (Michalewicz, 1996). Once the best chord is selected, it is put available to be played. The interface, which is looking for new notes, sends them to the MIDI port.



1011111	1010111	0010111	0100111
0011011	0010110	0110101	0100111
1010111	1011111	1010111	1100101

Figure. 2 – The rhythmic genetic cycle.

The following steps are realized in the genetic cycle (Pedrycz & Gomide, 1998):

<u>Step 1</u>: Create an initial population randomly;

<u>Step 2:</u> While not stopped (by the user), perform the following:

- Evaluate the musical fitness of each individual in the population;
- Apply the genetic operators to the population of MIDI chromosomes (groups of voices), taken into account based on the musical fitness, to create a new population. That is:

Reproduction: Copy existing individual strings to a new population;

Crossover: Create two new chromosomes by crossing over randomly chosen sublists (substrings) from two existing chromosomes;

Mutation: Create new chromosomes from an existing one by randomly mutating the character in the list;

Step 3: Find the best individual in the new population and play it as a MIDI event. Go to **Step 2**.

The steps above stress the existence of many operations executed in each cycle. The time interval between the selection of the best chords in two successive cycles may be different. On the other hand, the interface is regularly "asking for new notes". Despite the fact that there is an average time cycle to designate the best individual in each generation, small variations in each time cycle determine the *genetic rhythm*. Different times for the notes being played are perceived as a rhythmic profile of the pitch sequence generated by the genetic cycle.

5 FITNESS EVALUATION

This section presents the criteria used to evaluate the system musical fitness associated to each chord. Mathematical formulation is omitted, detailed description is found in (Moroni et all, 1999).

6 Voices Range Criterion

The chord's notes are related to voices that are associated to the linguistic terms: bass, tenor, contralto, soprano and nh (no human). The related fuzzy sets are shown in **Fig. 3**. Each voice is assigned to a membership value associated with each linguistic term in the set {NH, B, T, C, S}. For the classification of each voice, the membership function is evaluated for each set and the maximum value is taken. In the case of coincidence, the distance to the centre of the fuzzy set is considered. The interval of voices reached by the human voices is assumed to be in the interval H = [39..84], given in MIDI note values.

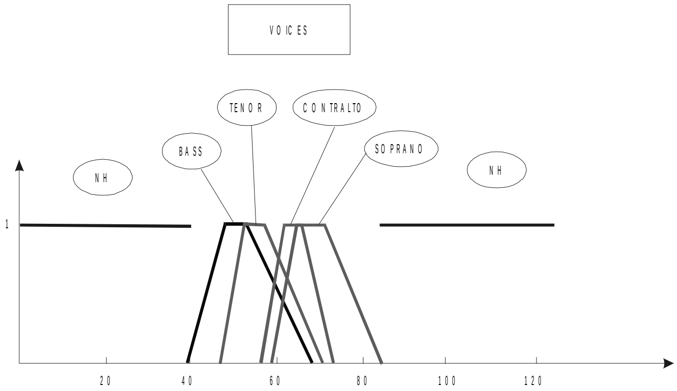


Figure. 3 – The linguistic values associated with the voices

Once the voices of each chord are evaluated according to its distribution in the interval of voices, the voice range criterion returns a value in the set {NH, W, M, G, E}. These linguistic values are associated to the concepts *No Human, Weak, Medium, Good and Excellent*. The optimal case – Excellent - is considered when the chord contains the voices Bass, Tenor, Contralto and Soprano. In this case, Nvalues = 4. The absence of these voices returns NH; the presence of one of them returns W; two returns M; three returns G; with Nvalues = 0, 1, 2, 3, respectively. Therefore, the voice range fitness is evaluated as:

$$O = NValues/4$$

7 The Consonance Criterion

The consonance among the four voices is evaluated as a function of the voice attributes. Consonance is defined as a function of the overlap between the harmonic series of two given notes (Vidyamurthy, 1992). This overlap measurement is then scaled to a value between 0 and 1, with 1 denoting complete overlap (i. e., the two notes being the same), and 0 denoting no overlap at all. This notion of overlap can be succinctly captured in the fuzzy set formalism further described.

The harmonic series derived from a given note is a set tone consisting of its fundamental tone and upper harmonic tones. In **Fig. 4** the weighting of the harmonic series versus the relative pitch is represented, and the sum of the weights is normalised. Note that n denotes the nth key on the piano, and that (n + k) denotes the key k semitones above key n. In **Fig. 4** , the upper tones of notes 60 (C) and 64 (E) are presented. Following, the resultant overlap or consonance is evaluated.

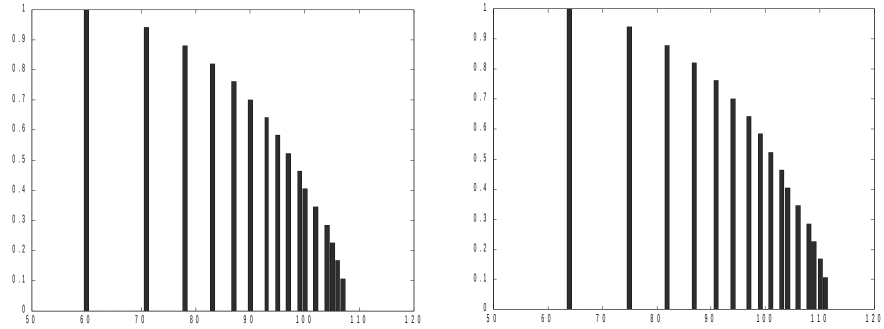


Figure 4 –The weighting of notes 60 (C) and 64 (E)

Formally defining, each note is a fuzzy set on a countable universe of discourse. The consonance or overlap between the notes is defined as the sum of the intersection of the harmonic series weights (Vidyamurthy & Chakrapani, 1992), and results in a value in [0..1]. **Fig. 5** cosonance between the pitches showed in **Fig. 4**.

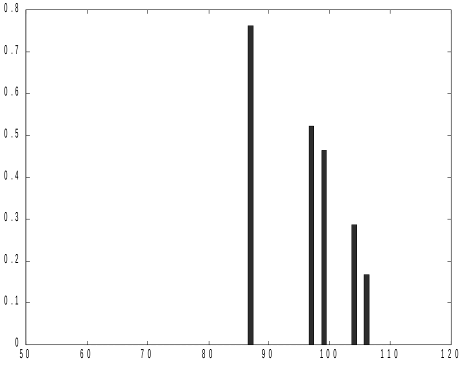


Figure 5 – The consonance or overlap between note 60 (C) and note 64 (E).

3.2.1 Vertical Consonance or Harmonic Fitness

Given a four note chord, the harmonic fitness is defined as the sum of the consonance or overlap of the harmonic series of the four notes present in the chord (as in **Fig. 5** is presented for two notes).

3.2.2 Horizontal Consonance or Tonal Fitness

Given a four note chord and a tonal centre named **Id**, the tonal fitness is defined as the maximal value of the consonance between each note present in the chord and the tonal centre **Id**.

8 Music Fitness

The resulting Musical Fitness is a conjunction of the previous functions and is defined as:

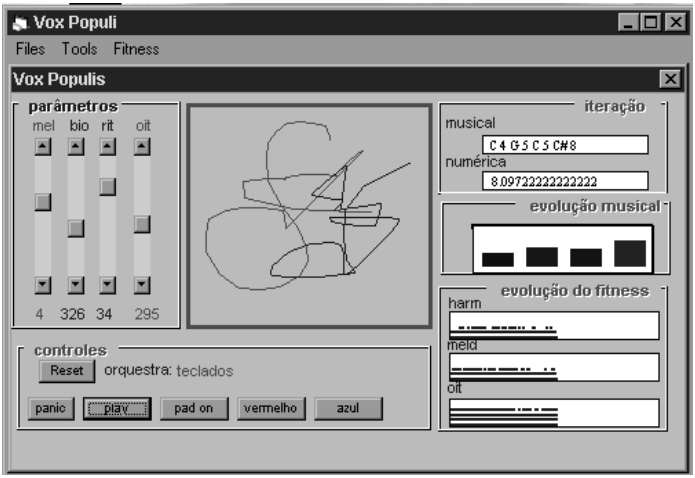
Music Fitness = Voice Range Fitness + Harmonic Fitness + Tonal Fitness

9 Vox Populi System

The system was implemented to perform a series of sound experiment and eventually to be used as a tool for Algorithmic Composition.

10 Interface and Parametric Control

The user may interfere in the fitness function through five interface controls: 1) Tonal Centre Control, 2) Biological Control, 3) Rhythmic Control, 4) Voice Range Control and 5) Orchestra Control.



mel bio rit oit
4 326 34 295

orquestra: teclados

musical: C4 G5 C5 C#8
numérica: 8.09722222222222

Figure 7 – Vox Populi Graphic Interface

It follows a short description of the controls available to user interaction with Vox Populi.

4.1.1 Tonal Center Control

The Mel scroll allows modifying the value **Id,** which is the tonal centre in the evaluation of the tonal fitness.

4.1.2 Biological Control

The Bio scroll allows interfering in the duration of the genetic cycle, modifying the time between genetic iterations. Since the music is being generated in real time, this artifice was necessary to synchronise the different process that is running. This value determines the slice of time necessary to apply the genetic operators, such as crossover and mutation, and can also be interpreted as the reproduction time for each generation.

4.1.3 The Rhythmic Control

The Rit scroll changes the time between evaluations of the musical fitness. It determines the “time to produce a new music generation” or the slice of time necessary to evaluate the musical fitness of the population. It interferes directly in the music rhythm. Changes on this control make the rhythm faster or slower.

4.1.4 The Voice Range Control

The Oct scroll allows enlarging or diminishing the interval of voices considered in the voices range criterion.

4.1.5 The Orchestra Control

Six MIDI orchestras are used to play the selected chords: 1) keyboards; 2) strings and brasses; 3) keyboards, strings and percussion; 4) percussion; 5) sound effects and 6) random orchestra, that takes any instrument from the General MIDI list.

11 Interactive Pad Control

The Pad On button enables and disables the pad change on four controls defined above. They are coupled in two pairs that are interpreted as variables of a two- dimensional phase space. This allows a user to draw an oriented curve to determine the music evolution. There are two curves, associated with different colours.

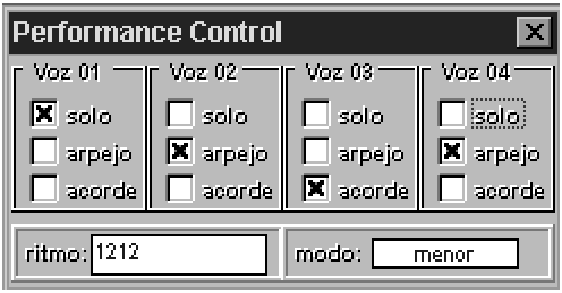
a) **RED CURVE:** describes a phase space of the *tonal* and *voice range* control variables.

b) **BLUE CURVE:** describes a phase space of the *biological* and *rhythmic* control variables.

The pad may be musically interpreted as an elementary tool that allows a “master gesture” to conduct the music.

12 Performance Control and Music examples

The Performance Control window presented in **Fig. 8** allows the user to shape the music output in real time. Each voice can be played as: a) *Solo*, b) *Arpeggio* and b) *Chord*. The Solo option enables direct output to the MIDI port. The Arpeggio option uses the last four notes calculate by the algorithm as MIDI output. The Chord option plays the last four notes as a block. A musician can change pitch material as a Chromatic, Major or Minor mode using the *Modo Option*. *Rhythmic Control* presented in 4.1.3 can be shaped by a rhythmic pattern input by the user as a string of small integers numbers. Vox Populi has also a Write MIDI option to produce MIDI files to be processed later in a sequencer program. Vox Populi demo can be downloaded from http://www.ia.cti.br/~artemis/voxpopuli.


Performance Control

Voz 01: ☒ solo ☐ arpejo ☐ acorde

Voz 02: ☐ solo ☒ arpejo ☐ acorde

Voz 03: ☐ solo ☐ arpejo ☒ acorde

Voz 04: ☐ solo ☒ arpejo ☐ acorde

ritmo: 1212

modo: menor

Figure 8 – Performance Control options

13 CONCLUSION

The resultant music moves from very pontilistic sounds to sustained chords. It depends upon the duration of the genetic cycle and the number of individuals on the original population. The octave fitness forces the notes to be in the range **H,** assumed to be the range reached by the human voices and associated with the central region of the notes in the piano. But since that several orchestras of instruments are used, this range is too limited for some of them. The original decision to restrict the generated voices to specific ranges was just to resemble human’s voices; nevertheless a user can enlarge this ranges using the Octave Control. The interface was designed to be flexible for the user to modify the music being generated, and can be thought as a prototype environment for algorithmic composition. The system has been proved to be flexible to receive new features. Most of the controls did not exist in its primer version, and were added as new interesting aspects that were identified to be extended. It has been revealed as a laboratory of sound experiments in which lots of possibilities may be explored.

Further, we are developing an integration of this system with Gesture Interfaces, as a glove to enhance the man/machine interaction with the purpose of allowing a human gesture to be the real time controller. This is a natural extension of the Pad Control.

This approach was previously applied in the project ActContAct (Manzolli et al. 1998), where an electronic shoe was used in music generation.

14 References

Biles, J. A. "GenJam: A Genetic Algorithm for Generation Jazz Solos". Proceedings of Computer Music Conference (ICMC '94), pp. 131-137, 1994.

Horovitz, D. "Generating Rhythms with Genetic Algorithms". Proceedings of Computer Music Conference (ICMC '94), pp. 142- 143, 1994.

Michalewicz, Z., *"Genetic Algorithms + Data Structures = Evolution Programs"*, Springer- Verlag, pp. 283 – 287, 1996.

Manzolli, J., A. Moroni and C. Matallo. "AtoConAto: new media performance for video and interactive tap shoes music", 6Th ACM International Multimedia Conference, pp. 3, 1998.

Moroni, A., J. Manzolli, F. Von Zuben and R. Gudwin. "Vox Populi: Evolutionary Computation for Music Evolution**".** Proceedings of the AISB'99 Symposium on Creative Evolutionary Systems, ISBN 1902956036, Published by AISB, pp. 82- 85, 1999.

Pedrycz, W. and Gomide, F., *"An Introduction to Fuzzy Sets Analysis and Design", The MIT Press, 1998.*

Vidyamurthy, G. and Chakrapani, J., "Cognition of Tonal Centers: A Fuzzy Approach", Computer Music Journal, The MIT Press, 16:2, 1992.

Acknowledgements

We would like to thank our fellow student Leonardo N. S. Pereira for developing the routines to evaluate the consonance criterion. This paper was supported by FAPESP (São Paulo State Research Foundation - process no. 95/8479- 3) and CTI (Technological Center for Informatics), and CNPq (process no. 300910/96- 7).