In search of musical fitness on consonance

Eleonora Bilotta (eleb@abramo.it), Pietro Pantano (piepa@abramo.it)
Centro Interdipartimentale della Comunicazione
Università degli Studi della Calabria,
87036, Arcavacata di Rende (CS), Italy

Abstract. After the generation of a sample of musical compositions, the rules which gave birth to these groups of compositions have been codified into a genetic code and these families, which had the capability of adapting to the environment and to reproduce themselves have been selected. The fitness the authors were *in search for* was about musical consonance. In this paper we present some results which could prove interesting in defining consonance/dissonance, operated automatically on musical compositions.

Key words: Evolutionary Music, Cellular Automata, Genetic Algorithms and Fitness.

Introduction

The common thesis in Artificial Life models is the idea that complex structure can emerge from the repeated application of a limited set of simple operations and give birth to emergent properties (Langton, 1997). When applied in a musical context, one of the major goal of these models is to find the optimum method for structuring musical compositions (Miranda, 1999). This relationship, between music and some AL models, furnishes a *semiotic triangle of signification (or musification)*. In this triangle, mathematical structures, the codification system we can use and various kinds of representations let us obtain artificial artefacts of a different kind, according to the codification systems we have chosen (Bilotta et al., 2000). There are two important directions in which we are going in exploring the musical context. The first is a scientific experimentation on the kind of artefacts we can obtain in order to detect a musical theory that can be utilised to write new kind of compositions. In fact, while some parts of the crafts of counterpoint, orchestration and construction of melodies are often codified into explicit rules, others are not. Erickson (1982) has argued that: "We need a music theory that is not style bound, that is valid for world music, not its European, Chinese or Indonesia varieties". This lack is especially evident in the context of evolutionary music. Composers are used to search the space of musical constructs to find just the right choice for a particular moment in a piece or they rely on some historical style in order to obtain a particular musical effect or *taste*. This research is multifold: we are exploring the mathematical, psychological and semiotic aspects of music.
The second direction concerns musical fitness, which can give us some global characteristics of a melody. In evolutionary music, musical fitness (like visual one, Sims, 1993) is generally operated intuitively by listeners. It reflects an aesthetic judgement according to which some pieces, in a population of compositions, are better or worse, based on subjective, emotional or perceptual criteria, expressed by a listener's sample. Some authors have detected it as the source for a digital instrument (Takala, 1993), generating a single rhythm measure (Horowitz, 1994), evolving an "ear" module (Jacob, 1995), building up a knowledge base of melodic ideas for use in improvising jazz solos (Biles, 1994), training a neural net to discriminate fitness (Biles et al., 1996).
We have used a genetic algorithm to investigate musical fitness on consonance. We have generated a sample of musical compositions. The rules which give birth to these groups of

compositions have been codified into a genetic code and we have selected those families, which had the capability of adapting to the environment and of reproducing themselves (fitness) (Mitchell, 1996). The fitness we were *in search of* was about musical consonance. In fact, we computed the consonance and dissonance values between two notes in a melody (forgetting for a moment the discussion between tonal and a-tonal music). The offspring generations, created by some genes random mutation, have been selected according to their fitness and the process has been repeated many times. After many generations, it has been possible to observe an empowerment of some populations' fitness and of emergent properties in melodic organisations. In this paper we present some results which could be of interest in defining consonance/dissonance, operated automatically on musical compositions.

The opposition consonance/dissonance

What sort of information does sound convey? Human mental representations of acoustic events can be elaborate in a way that the word sound doesn't offer. Sound refers indifferently to the physical sound in the world and to the mental experience of it, so that consonance/dissonance is related to physical-mathematical theories, to psychological experiments on sound perception and to the musical organisation of sound. The problem of consonance/dissonance in music began with the Pythagoreans in the fifth century B. C. They discovered that two notes produced by strings were consonant (which means they sounded smooth together) when the ratios of the lengths of the strings were formed from low integers, such as 2:1 (the octave), 3:2 (the fifth) and 4:3 (the fourth). In modern Physics, the rule has been replaced to state that the most tonal combination involve simple ratio relationships between frequencies. Von Helmoltz, a German physiologist and physicist in the XIXth century, elaborated the best-accepted theory of consonance, which blames the roughness of dissonant tones on the perception of beats. Helmoltz argued that the auditory system resolves complex tones or mixture of tones into individual spectral components (partials). When two partials are too close in frequency for the auditory system to resolve them, it hears the beats (periodic fluctuations of intensity) created by their summation. To this theory has been added the more recent theory of *critical band*, which is defined as a certain frequency separation, within which partials are not fully resolved by the auditory system and then interact to produce phenomena such as masking or beats. The opposition consonance/dissonance is still a problem, since the theories of atonal music and computer music have opened a broad range of possibilities in creating musical compositions. While some authors consider consonance as lacking in meaning (Schoenberg, 1984), some experiments show that the adult judgement about consonance is well consolidated (Huron, 1991; Shellemberg and Trehub, 1994).
How can an auditory event be translated into a perceptual representation of it? When listeners create a mental representation of the auditory input, they must employ rules about what happens. Gestalt's principles of grouping were evolved by a group of German Psychologists in the early part of the XXth century to explain why elements in sensorial experience seemed highly connected to one another. The word Gestalt means "pattern" and the theory described how the brain created mental patterns by forming connections between the elements. Gestalt theorists argued that there was always competition between "the forces of attraction" of elements, so that the perceptual organisation that emerges from this conflict would be a consequence of the distribution of forces across the whole perceptual "field" and not of the properties of individual parts taken in isolation. The Gestalt psychologists' view was that the tendency to form perceptual organisations was innate and occurred automatically whenever we perceived anything. It was impossible, they claimed, to perceive sensory elements without their forming an organised whole. They argued that this organisational tendency was an automatic tendency of the brain. In fact, some recent research shows that there is a biological

basis for consonance and that results show clearly that consonance and dissonance are combinations of frequencies which produce different stimuli configurations in the neural net disposed for the aural perception (Zentner and Kagan, 1996, 1998).
What is the role of primitive organisation in music? Music builds elaborate structures of sounds, but its aesthetic and perceptual comprehension is not the raw properties of the individual sounds, as Gestalt Psychologists pointed out. It also builds upon structural concepts such as scales, model and key relations and a variety of transformations, which include transportation and repetition. Experienced listeners make use of a large number of musical concepts or schemas in listening to music. Traditionally, music is thought of as having a horizontal and a vertical dimension. This derives from musical notation in which the horizontal dimension stands for time and the succession of sounds that forms the melody, and a vertical one which depicts pitch relations, or the simultaneous sounds that form harmonies. Usually, musicians speak of musical texture in referring to how these types of threads go together, how strong the vertical dimensions are in comparison with the horizontal ones, how the horizontal ones change over time, and so on.
Since we want to verify some of the emergent properties of AL models in the musical context, we choose to work on consonance as a means of fitness. We are following the hypothesis that evolution is present in music and it is possible to detect in this evolution emergent properties and higher order organisation, which in some way resemble the historical evolution that has produced musical systems, as we actually know it. What happens in evolutionary music?

A genetic algorithm for automatically searching for musical fitness

Let us consider a one-dimensional network of cellular automata. The automaton *i* can assume value $x_i = 0,1...N-1$. The evolution of the automaton's state in time *t+1* depends on the automaton’s and its neighbourhoods states at the time *t*. This is a deterministic system since the evolution is fully predicted. For this reason, knowing the system's state at its start time, we'll be able to calculate the system's state at every instant of time. For a one-dimensional network of cellular automata, knowing a sequence of numbers $x_i(0) = c_i$ for $i = 1,2,...n$, which represents the automata values at the initial state, we can generate a numerical matrix, that gives us the automata values at every instant of time:

(1) $$A = \begin{pmatrix} x_1(0) & x_2(0) & \ldots\ldots & x_n(0) \\ x_1(1) & x_2(1) & \ldots\ldots & x_n(1) \\ \ldots\ldots & \ldots\ldots. & \ldots\ldots. & \ldots\ldots. \\ x_1(m) & x_2(m) & \ldots\ldots. & x_n(m) \end{pmatrix}$$

If we consider $P = \{0,1,..,N-1\}$ as the space of all possible automation values, a transition rule becomes an application from P^h to P, where h is the automation’s neighbourhood dimension. By supposing h=3, the single rule can be expressed as:

(2) $s_{ijk} = s_{ijk}(p_i, p_j, p_k)$,

with $p_x \in P$.

The evolution rules can be represented by the following sequence:

(3) $y = \langle s_{000}, s_{100}, s_{200}, ... s_{k00}, ... s_{kbc}, ... s_{kkk} \rangle$

We can consider the system's evolution rules as the network's genome.
Since, in the musical context, fitness is in correlation with musical composition, at this stage it is necessary to speak about the musification process we have utilised. There are many

possibilities. The first is to correlate every column of the matrix to a note, choosing which of them will be played simultaneously, and in which sequence. This could be done in the following ways: to play just a note, in the horizontal (from left to right) and in the vertical (from top to bottom) dimensions of the matrix; to play a triad; to simultaneously play every note in a row. In the first condition, the automata' values can coincide with the duration of the note, while in the third condition we have chosen to realise a correspondence between the matrix values of a row with the time at which a note has be played. The musification mechanism that we have chosen, functions in this way: consider a row of the matrix and select the automatons which have the same value; to these automatons will correspond some notes which will be played contemporaneously (Bilotta et al., 2000). After having chosen the musification mechanism, we can use the temperate diatonic scale to value the consonance relations among the notes. If we approximate the relations among the frequencies to these of the natural scale, we can construct a succession of numbers among 1 and 30 whose extremes represent the values of minimum and maximum consonance between two notes. The other values fluctuate between those two extremes. It's very important to note that the associated value doesn't represent a consonance "measure" but rather that of the relative position of one relation in respect of the other. Therefore we have constructed the following matrix:

(4)

	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
C	30	16	22	7	26	28	17	29	18	27	5	20
C#	16	30	14	22	25	13	28	11	29	24	27	23
D	22	14	30	16	21	12	6	28	17	10	4	27
D#	7	22	16	30	19	1	12	13	28	8	10	24
E	26	25	21	19	30	19	21	25	26	28	17	29
F	28	13	12	1	19	30	16	22	7	26	3	9
F#	17	28	6	12	21	16	30	14	22	25	26	15
G	19	11	28	13	25	22	14	30	16	27	2	26
G#	18	29	17	28	26	7	22	16	30	19	21	25
A	27	24	10	8	28	26	25	21	19	30	16	22
A#	5	27	4	10	17	3	26	2	21	16	30	14
B	20	23	27	24	29	9	15	26	25	22	14	30

We have picked out three classes in which it is possible to associate the pairs of notes to be played together.
A: fifth major, fourth major and third major;
B: second major and seventh major;
C: every other we can realise;
D: unison.
We can evaluate fitness function stressing the consonance of class A and B, instead of those which occur in the class C or D. For example, if some notes are played simultaneously, fitness could be defined as follows:

(5) $$F = \sum_{i,j=1}^{m,n} \sum_{k=1}^{n} r_{ijk} (c_{ik} + a_{ik})$$

where n is the column number of the matrix (1), and m is the number of the row; a_{ij} is an element of the matrix (1), c_{ij} is the weight to assign into the consonance's values of matrix (4), according to the class they belong to; $r_{ijk} = 1$ if $x_i(j) \neq 0$ and $x_k(j) \neq 0$, if not it is equal to 0. If notes are played one at a time, fitness function is a little more complex than (5), since one note is coupled only with that one that follows it, in the some row and not with all the others. In the same way, if we consider a triad, every single note of every row will be coupled only with the other two notes which follow it.

Some results

Let us suppose that the weights assigned to each family are:

(6)

Family	A	B	C	D
Weight	1000	500	-50	0

If we assign to random mutations a probability of 3% and consider many times the selection's process, starting from different initial states, the process of evolution that goes on is represented in the following diagram (figure 1). Let us analyse the diagram. Fitness curves grow quickly, until they arrive at high values, where they become more stable. Only one curve of fitness grows to lower values. In this case, it seems to us the system has arrived at a local maximum.

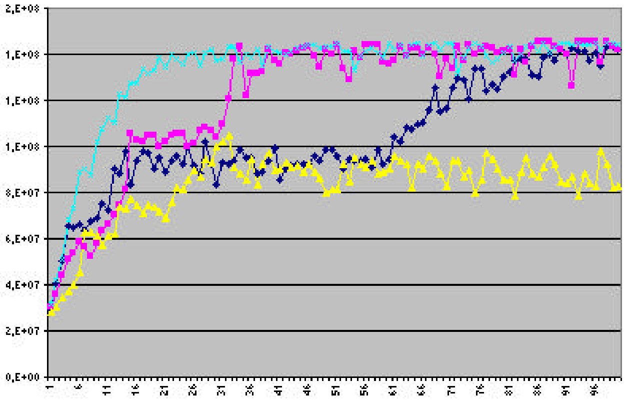

Figure 1. Fitness functions for mutation rate equal to 3%

On the contrary, the system percentage of success, in arriving at the absolute maximum, seems very good. In this case, considering a longer evolutive process doesn't give better results. On the auditory level, if we analyse as results musical compositions, we can listen to recursive melodies, based on a small quantity of notes (even if the musification process we used is the result of the fusion and superimposition of more than one note). The fitness function we have used lets the system continue in a recursive manner, and even from a very simple system we can obtain complex compositions, where some emergent properties are exhibited. In fact, the phenomenon has already been present for the first ten generations. There are little variations in a composition that has a higher fitness (of the last generations) and one of the first ten generations (with a lower range of fitness). The sharing characteristics amongst different generations are genetically stronger and, for this reason, they are inherited by the last generations. On the musical side, we are going to carry out experiments comparing expert and non-expert listeners, to confirm the emergence of these proprieties.

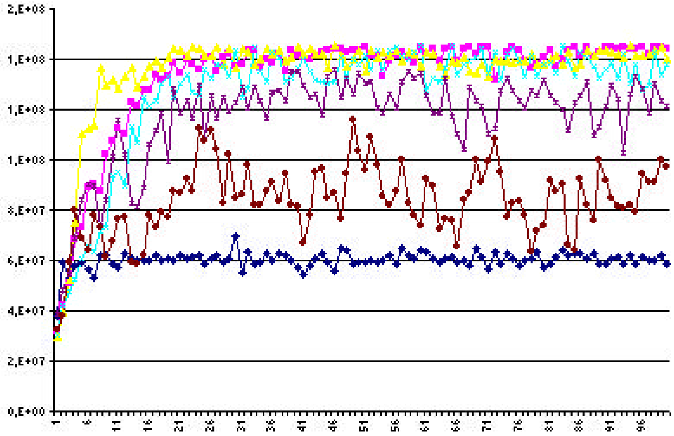

Figure 2. Fitness functions for different percentage of mutation

If we augment the system's percentage of mutation from 1% to 20 %, we can see that at 1% the fitness function doesn't grow. The system arrives at the highest values around 3%; while, for higher percentage (4%, 5%, 10%, and 20%), results aren't better, since oscillations from local maximum become greater (see Figure 2).

These results don't change in a relevant manner even if we make a crossover at a point of the genome, because the network is able, using the fitness function we have created, to arrive at its best efficiency. To force the network means to let it leave its point of stability. The network doesn't grow in efficiency with neighbourhood 5 or more than 5. On the contrary, it seems unable to arrive at the absolute maximum, getting only relative values. The network remains on lower fitness values if we change the weights (6), varying them as follows:

Family	A	B	C	D
Weight	500	100	-50	0

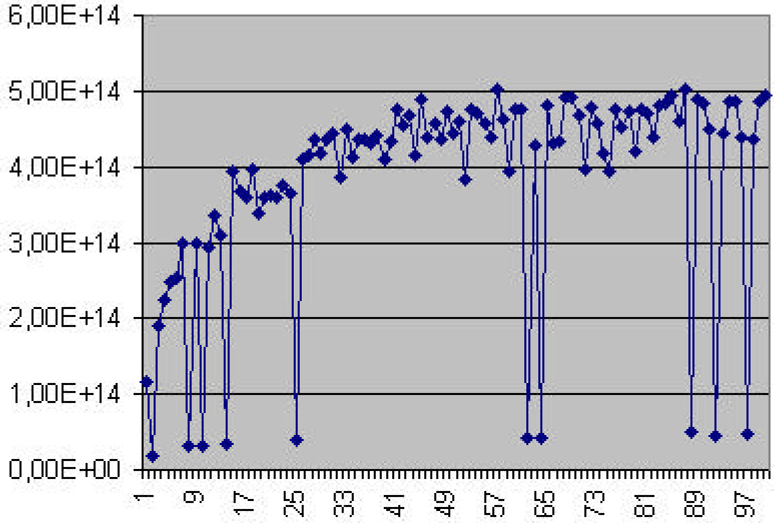

Figure 3. Fitness function with different weights.

We obtain the results, represented in Figure 3. Fitness grows relatively, the presence of relative and absolute maximum continues to emerge, while the oscillations increase because the system doesn't reinforce fitness as in the preceding situation.

Bibliography

Biles J. A (1994). GenJam: A Genetic Algorithm for Generating Jazz Solos. In *Proceedings of the 1994 International Computer Music Conference*, ICMA, San Francisco, 1994.

Biles J. A., P.G. Anderson, Loggi L. W. (1996). Neural Network Fitness Functions for a Musical IGA. At the following Internet address: http://www.it.rit.edu/~jab/SOCO96/SOCO.html , 04/02/2000.

Bilotta E., Pantano P., Talarico V. (2000). Synthetic Harmonies: An approach to Musical Semiosis by means of Cellular Automata. In *Artificial Life VII: Proceedings of the Seventh International Conference, edited by Mark A. Bedau, John S. McCaskill, Norman H. Packard and Steen Rasmussen*, MIT Press.

Erickson R. (1982). New Music and Psychology. In Deutsch D. (Ed.) *The Psychology of Music*, New York, Academic Press.

Horowitz D. (1994). Generating Rhythms with Genetic Algorithms. In *Proceedings of the 1994 International Computer Music Conference*, ICMA, San Francisco.

Huron D. (1991). Tonal consonance versus tonal fusion in polyphonic sonorities, *Music Perception*, **9**, 135-154.

Jacob B. (1995).Composing with Genetic Algorithms. In *Proceedings of the 1995 International Computer Music Conference*, ICMA, San Francisco.

Langton C.G. (Ed.) (1997). *Artificial Life. An overview*, Cambridge, Mass., MIT Press.

Miranda E. R. (1999). Modelling the Evolution of Complex Sounds: A Cellular Automata Approach, *Proceeding of AISB'98 Symposium on Creative Evolutionary Systems*, Edinburgh College of Art and Division of Informatics, University of Edinburgh.

Mitchell M. (1996). *An Introduction to Genetic Algorithm*, Cambridge, Mass., MIT Press.

Schoenberg A. (1984). *Style and idea*, Berkeley, University of California Press.

Shellemberg E. G. and Trehub S. E. (1994). Frequency ratios and the perception of tone patterns, *Psychonomic Bullettin & Review*, **1**, 191-201.

Sims K. (1993). Interactive Evolution of Equations for Procedural Models. *The Visual Computer*, **9** (8), 466-476.

Takala T., Hahn J., Gritz L., Geigel J., and Lee J. W. (1993). Using Physically Based Models and Genetic Algorithms for Functional Composition of Sound Signals, Synchronized to Animated Motion. In *Proceedings of the 1993 International Computer Music Conference*, ICMA, San Francisco.

Zentner M. R. and Kagan J. (1996). Perception of music by infants, *Nature*, **383**, 29.

Zentner M. R. and Kagan J. (1998). Infants, perception of consonance and dissonance in music, *Infant Behavior and Development*, **21**, 483-492.