

Evolution in Musical Contexts: The Software DARWIN

MODALIDADE: COMUNICAÇÃO

Carlos de Lemos Almada

Universidade Federal do Rio de Janeiro – calmada@globocom

Abstract: This paper examines some correlations between the fundamentals of the Darwinian Theory of Evolution and musical variation. DARWIN, a computational program inspired on Richard Dawkins (2000) algorithmic creation of “biomorphs”, is employed to produce six “mutations” in a structure of a given melodic line, a process that is recursively replicated along 20 “generations”. Through the use of an artificial selection strategy, an evolutionary lineage is obtained from the basic melody. Some discussions and possible developments for the research are proposed in the last section of the paper.

Keywords: Biomorphs. Theory of Evolution. Musical variation. Computational music.

Evolução em Contextos Musicais: O Software DARWIN

Resumo: Este artigo examina algumas correlações entre os fundamentos da Teoria da Evolução de Darwin e a variação musical. DARWIN, um programa computacional inspirado pela criação algorítmica das “bioformas” por Richard Dawkins (2000), é empregado para a produção de “mutações” na estrutura de uma dada linha melódica, num processo que é recursivamente replicado por 20 “gerações”. Por intermédio do uso de uma estratégia de seleção artificial é obtida uma linhagem evolutiva advinda da melodia básica. Algumas discussões e possíveis desdobramentos para a pesquisa são propostas na seção final do artigo.

Palavras-chave: Bioformas. Teoria da Evolução. Variação musical. Música computacional.

1. Introduction

This article integrates a broad research project dedicated to systematical studies on musical variation. It is based on Arnold Schoenberg’s theory of developing variation and *Grundgestalt* and has already produced, as main results, published papers and some computational tools according to two perspectives, analysis and musical composition.¹ The essential motivation for the present approach is to examine the limits of similarity between a given referential form and the infinite pool of possible variants which may derive from it through gradual and progressive divergence. Such investigation is based on the general principles of evolutionary biology, from Charles Darwin’s Theory of Natural Selection (or, as it is more known, the Theory of Evolution) and, more specifically, inspired on the concept of *biomorphs*, created by Richard Dawkins. As an adaptation of the process for biomorphs production it was developed the software DARWIN, intended to establish “evolutionary pathways” (another of Dawkins’ concept) originated out of a musical theme, through the production of successive “mutations” along a given number of generations, replicating

(evidently, in a much shorter time scale) the process of artificial selection observed in evolution of certain biological species.

2. Evolution, artificial selection, and the biomorphs

Published in 1859, *The origins of species* (DARWIN, 1859/2014) is one of the most important scientific books ever written. Its author, the English naturalist Charles Robert Darwin (1809-1882), after decades of research, systematical collection of an innumerable quantity of specimens (resulting from a long voyage around the world), experiments, and deep reflections, elaborated his revolutionary thesis: the evolution of life through natural selection of the fittest individuals. As argued Darwin, *variation* is the principal evolutionary mechanism responsible for such an extremely slow divergences between a referential living form and the myriad of its descendants – which ultimately would explain the existence of the almost infinite quantity of species.² In spite of a violent reaction of part of the contemporary scientific community and of the large number of critics to Darwin's central arguments, the Theory of Evolution has survived and became the basis for modern biology and genetics. Currently, evolutionary studies (associated to the branch known as New Evolution) are deeply related to concepts not even foreseen by Darwin, like chromosomes, randomly mutation, DNA and RNA filaments, genes, phenotypes, genetic code, nucleotides, enzymes, and so forth). One of the most known, important and respected (and, sometimes, controversial) new evolutionists is the English ethologist and biologist Richard Dawkins, author of several books on the subject. In the third chapter of *The blind watchmaker* (DAWKINS, 2000), he describes a computational program created by himself for simulating (at a very high speed) the effects of cumulative evolution of a live being. With this purpose, Dawkins introduced into the code of a simple algorithm employed to draw a fractal tree-like figure a series of mutations, influencing, at each generation, almost imperceptible modifications of specific aspects of the basic figure, associated – in his own terminology – to its “genes”. Deliberately, in order to simplify the process and his own program, Dawkins constrained the number of genes to nine, as well as the magnitudes of their possible alterations, allowing only differences of one unity (-1 or +1). Figure 1 presents some “genetic” distortions caused by Dawkins' mutations.

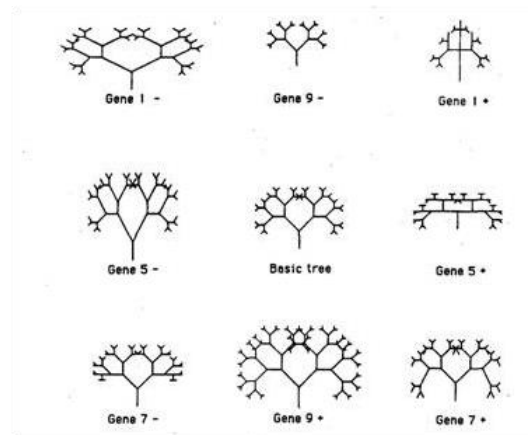


Figure 1: Basic figure (center) with mutations associated to some genes (in: DAWKINS, 2000: 66)

From each produced breed Dawkins selected one of the nine “children” (i.e., the mutations of the respective genes) to become “parent” of a new generation of variants. In other words, he adopted an artificial selection strategy (we will return to this point later). Dawkins comments his great amazement when he observed in his computer the results obtained after the 29th generation:

When I wrote the program, I never thought that it evolve anything more than a variety of tree-like shapes. (...) Nothing in my biologist’s intuition (...) prepared me for what actually emerged on the screen. I can’t remember exactly when in the sequence it first began to dawn on me that an evolved resemblance like an insect was possible. (DAWKINS, 2000: 73)

Reproducing many times the experiment, he found other types of evolved “creatures” (Figure 2), which he generically named as “biomorphs”.³



Figure 2: Some species of biomorphs (in: DAWKINS, 2000: 75)

3. Musical evolution with DARWIN

Inspired by these fascinating results and impressed with the close relationship that exist between the production of Dawkins’ biomorphs and the processes of musical variations,

I decided to elaborate DARWIN, a software destined to produce gradual and cumulative mutations on the structure of a referential musical form – the “patriarch” of an intended lineage.⁴ In the current version, DARWIN has six possible *mutational agents* that affect somehow the “genetic code” (Gc) of a given referential melodic line. In this context, Gc is a two-dimensional structure which displays the abstracted intervallic and rhythmic contours of the melody, translated as sequences of integers, according to conventions previously established in the research (Example 1).⁵



$$\text{Genetic code} \begin{cases} \langle +2+2-9 \rangle & [I] \\ \langle +4+2+2+8 \rangle & [R] \end{cases}$$

Example 1: Genetic code of a simple melodic fragment

The mutational agents operate basically in the same manner: the program randomly chooses a “gene” of the theme’s Gc (interval or duration) and applies to it a specific algorithm. The result is always a relatively close variation of the “parent”, since just one of its elements was modified (see Table 1 for a description of the six agents).⁶ Therefore, at least hypothetically, the longer the cumulative evolutionary process (i.e., the sum of the mutation steps), the larger it will be the divergence of resemblance between ancestor and descendants.

Genetic component	Mutational agent	Description	Examples (from Ex.1)
Intervallic structure	1. Inversion	Multiplies the selected gene by (-1).	$\langle +2+2-9 \rangle \Rightarrow \langle +2-2-9 \rangle$
	2. Change of register	If the selected gene is positive, subtracts it from 12. In the contrary case, adds 12 to it.	$\langle +2+2-9 \rangle \Rightarrow \langle -10+2-9 \rangle$
	3. Expansion	Adds to the selected gene (or subtracts from it) a randomly chosen prime number {1,2,3,5,7} if it is, respectively, positive or negative.	$\langle +2+2-9 \rangle \Rightarrow \langle +2+2-10 \rangle$
	4. Contraction	Adds to the selected gene (or subtracts from it) a randomly chosen prime number {1,2,3,5,7} if it is, respectively, negative or positive.	$\langle +2+2-9 \rangle \Rightarrow \langle +2+2-4 \rangle$
Rhythmic structure	5. Augmentation	Multiplies the selected gene by 2.	$\langle +4+2+2+8 \rangle \Rightarrow \langle +4+4+2+8 \rangle$
	6. Diminution	If the selected gene is even, divides it by 2. In the contrary case, subtracts one unity from it.	$\langle +4+2+2+8 \rangle \Rightarrow \langle +4+2+1+8 \rangle$

Table 1: DARWIN’S mutational agents

Apart from several contact points between DARWIN and the Dawkins’ program, the former has an exclusive feature destined precisely to determine the “parenthood” degree between two related forms: the *coefficient of similarity* (Cs),⁷ an index expressed as a real

number from 0 (total contrast) up to 1 (identity). DARWIN considers two types of Cs's: (a) relative, that measures the similarity between a given mutation and its immediate progenitor and (b) absolute, that measures the similarity between the mutation and the “patriarchal” form (i.e., the musical structure that initiates the evolutionary process). Even though, in most cases, the tendency is that both Cs's decrease along of the generations, sometimes the absolute index surpasses the relative one, which indicates the occurrence of a sort of “regression”, as eventually happens in biological circumstances (for example, a grandson that looks more like his grandfather than properly his father). In the present experiment, the use of Cs measurement aims to contribute with more precision to the evaluation of the evolutionary trajectory of the sequence of mutations.

4. Application

Once a theme is chosen as progenitor of the lineage to be built, DARWIN allows numerous evolutionary pathways, depending specially on the criteria adopted for selection of the fittest mutation at each generation.⁸ I adopt here a similar protocol to that Dawkins used to obtain the biomorphs: at each generation, after a careful examination of the entire breed, one “individual” is chosen as the most “adequate” mutation for musical purposes, becoming thus the reference for a new generation.⁹ In spite of how subjective it may sound, this procedure is not essentially different from the artificial selection employed in agriculture and animal creation (dogs, for example), in which the choice of the fittest result depends on its *phenotypic qualities* (i.e., the physical expression of genetic transformations). In this case, the phenotypes of the variations are specifically related to their musical attributes, as melodic contour, range, balance, rhythmic organization, implicit harmony, and so on.

Example 2 presents a referential melodic line (or *progenitor-theme*) for the subsequent derivative process, a simple and clearly tonal structure, input in the program as a monophonic midi file.



Example 2: The progenitor-theme

The application here proposed consists on the production of an evolutionary lineage of descendants from the progenitor-theme along 20 generations. I conducted the evolutionary line according to what seemed to me the most appropriate (in this case, taking into

account the above mentioned musical attributes), by selecting the fittest mutation at each generation and using it as parent for the next step. Considering the limits of scope and extension of the present study, only the main experiment's outcomes are listed below. Example 3 displays the extremities of the genealogical tree referring to the evolutionary pathway covered by the successive choices made, omitting the intermediary stages.



Example 3: Genealogical tree's extremities of the progenitor-theme's evolutionary lineage

Table 2 provides a report of the complete sequence of produced “biothemes” (in analogy to Dawkins' creatures) with their respective mutational agents and coefficients of similarity (relative and absolute). These are also presented in graphic format (Figure 3).

5. Analysis and discussion

The results obtained in this experiment leads to interesting questionings. Firstly, as we can observe in Example 3, a (purely) visual comparison between the progenitor-theme and its last “heir” (biotheme #20) suggests that they are deeply contrasting ideas. This divergence is mainly due to the slow and gradual “erosion” of the rhythmic structure of the successive variants, which consequently undermines the motivic parallelism organization, generally considered as a decisive factor for melodic coherence.¹⁰ Moreover, the cumulative evolutionary process affected the pitch structure, causing also less evident changes in the implicit harmonic original context. However, hearing both melodies we perceive that some degree of identity between them were maintained, which is probably due to their general

melodic contours (Figure 4). This is supported by results of empirical studies on musical cognition,¹¹ which suggest that the resemblance between melodic ideas is more conditioned by their respective contours than by absolute pitches.

Generation	Mutational agent	Relative Cs	Absolute Cs
0 (progenitor)	0	1	1
1	4	0.998	0.9982
2	1	0.947	0.946
3	5	0.758	0.758
4	2	0.752	0.712
5	1	0.744	0.698
6	6	0.741	0.596
7	1	0.614	0.569
8	3	0.612	0.469
9	5	0.583	0.491
10	4	0.392	0.338
11	5	0.378	0.299
12	6	0.359	0.289
13	1	0.353	0.209
14	6	0.335	0.270
15	3	0.335	0.195
16	4	0.325	0.201
17	5	0.309	0.249
18	2	0.301	0.178
19	5	0.286	0.231
20	3	0.281	0.162

Table 2: List of the 20 descendants, considering mutational agents applied, and their respective relative and absolute coefficients of similarity

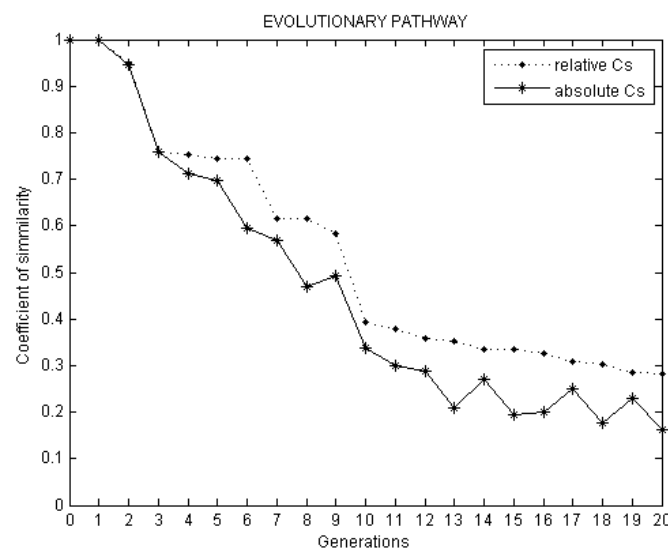


Figure 3: Graph of the lineage's evolutionary pathway presenting the behavior of absolute and relative Cs's along 20 generations

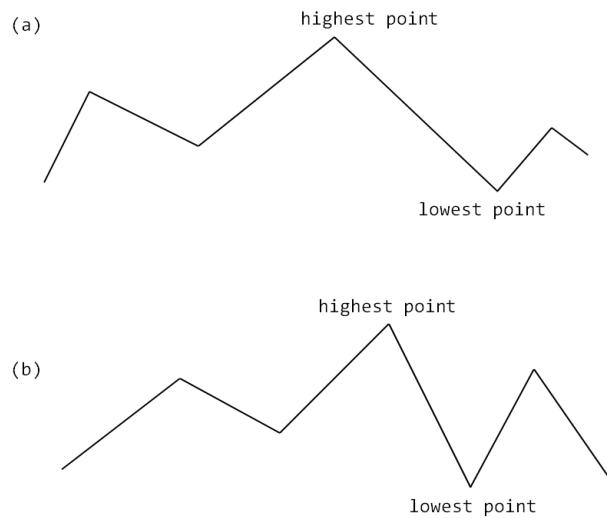


Figure 4: Comparison of the general melodic contours of the progenitor-theme (a) and biotheme#20 (b)

A second important and correlate question is: which is the limit of similarity between a given variant and its referential form? Or, alternatively, when a musical descendant loses any “blood ties” with the lineage patriarch? Would it be possible to establish a clear-cut border separating the minimally similar from the irreconcilable contrast?¹² If affirmative, could it be associated to a null value for the absolute Cs?¹³ In this case, how many generations (30, 50, 100?) it would take to achieve this presumed frontier? Anyway, a proper investigation in this direction needs firstly to be preceded by a detailed revision (and, probably) improvement of the algorithm currently used to calculate the Cs values, since, as was observed during the realization of the experiment, there are still some little distortions in the results.¹⁴

6. Concluding remarks

Taking this study as an initial step of a new branch of the research, it is possible to envisage some promising developments, not only related to cognitive aspects of variation and similarity (as suggested in the previous subsection), but also concerning composition. In this aspect, for example, one could employ DARWIN for creation of thematic ideas with different degrees of resemblance with a referential form. Disregarding how close or remote from the original source could be such ideas, is fascinating to think that all of them would be somewhat mutually related (even those with null Cs!) and have a common ancestor, in an analogous manner as happens with the living species.

Another possibility for further research, following again one of the Dawkins' ideas, is to try reconstructing evolutionary pathways from a given "formula" of modifications of a biotheme's Gc. Since DARWIN preserves all "genetic" information of a produced lineage, it means that each pair of "genes" (reference/mutation) can be compared and their differences quantified. Figure 5 illustrates this aspect, comparing the Gc's of the progenitor-theme with that one of biotheme#20. A matrix of absolute differences informs the "genetic distances" between each internal element, which can ultimately be associated to the degree of divergence between both thematic structures.

	intervallic Gc													rhythmic Gc																	
prog.	8	-1	1	-5	2	2	1	2	-8	-1	-2	-2	-1	8	-3	4	2	2	6	2	2	2	2	2	6	2	4	4	4	4	2
bt#20	1	6	-1	0	-1	7	-4	-10	5	-3	2	8	-3	-4	-5	8	2	2	12	2	4	2	2	2	6	4	2	4	4	2	8
diff.	7	7	2	5	3	5	5	12	13	2	4	10	2	12	2	4	0	0	6	0	2	0	0	0	0	2	2	0	0	2	6

Figure 5: Genetic codes of the progenitor-theme and biotheme#20 and matrix of internal differences

As can be observed, this bias permits to conclude that the intervallic mutations were considerably more far-reaching than the rhythmic ones, since the number of "zeros" indicates the absence of "genetic" transformation.

References:

- ALMADA, Carlos de L. O Sistema-Gr de composição musical baseada nos princípios de variação progressiva e *Grundgestalt*. *Música e Linguagem*, Vitória, v. 2, n. 1, p.1-16, 2013a.
- _____. Comparação de contornos intervalares como parâmetro de medição de similaridade. In: ENCONTRO INTERNACIONAL DE TEORIA E ANÁLISE MUSICAL, (3.), 2013, São Paulo. *Anais...* São Paulo: ECA-USP, 2013b, p.205-214.
- _____. Simbologia e hereditariedade na formação de uma *Grundgestalt*: a primeira das *Quatro Canções* Op.2 de Berg. *Per Musi – Revista Acadêmica de Música*, Belo Horizonte, n.27, p. 75-88, 2013c.
- _____. A variação progressiva aplicada na geração de ideias temáticas. In: SIMPÓSIO INTERNACIONAL DE MUSICOLOGIA, (2.), 2011, Rio de Janeiro. *Anais ...* Rio de Janeiro: UFRJ, 2011. p.79-90.
- DARWIN, C. *The origins of the species*. London: John Murray, 1872). Available in: <https://www.andrew.cmu.edu/user/jksadegh/A%20Good%20Atheist%20Secularist%20Skeptical%20Book%20Collection/Charles%20Darwin%20-%20The%20Origin%20of%20Species%20-%206th%20Edition.pdf>
- Access in: 18/feb/2015
- DAWKINS, Richard. *The blind watchmaker*. 3rd edition. London: Penguin Books, 2000.
- LERDAHL, Fred & JACKENDOFF, Ray. *A generative theory of tonal music*. Cambridge: The MIT Press, 1983.
- MARVIN, Elizabeth. *Tonal/atonal: Cognitive strategies for recognizing transposed melodies*. In: BAKER, James. *Music theory in concept and practice*. Rochester: University of Rochester Press, 2009. p.217-236.



MAYR, Desirée & ALMADA, Carlos. Uma aplicação da análise derivativa na música romântica brasileira: O primeiro movimento da *Sonata para Violino e Piano* Op.14, de Leopoldo Miguéz. In: CONGRESSO DA ASSOCIAÇÃO BRASILEIRA DE TEORIA E ANÁLISE MUSICAL, (1.), 2014. Salvador. *Anais ...* Salvador: UFBA, 2014, p.19-37.

McADAMS, Stephen & MATZKIN, Daniel. Similarity, invariance, and musical variation. In: *Annals of the New York Academy of Sciences: the biological foundations of music*, New York, v. 930, p. 62-76, 2001.

Notes

¹ For the published papers, see, for example, ALMADA (2011;2013a;2013b;2013c) and MAYR & ALMADA (2014).

² This becomes even more impressive when we think that “the actual animals that have ever lived on Earth are a tiny subset of the theoretical animals that *could* exist”. (DAWKINS, 2000: 90). Or, in Dawkins’ terms, the “genetic space” (the universe of total possible variants) would be virtually infinite. For more information, see (Ibid.: 81-91).

³ Several sites in the web provide interesting information about Dawkins and his biomorphs (including download of a program for their production). See, among others:

https://www.youtube.com/watch?v=Ok_tcAEbHHw

<https://richarddawkins.net/?gclid=CNDc16eTgsQCFYrm7AodinMA-Q>

<http://www.codeproject.com/KB/recipes/EvolvingCrittters/EvolvingCrittters.zip>

<http://www.emergentmind.com/biomorphs>

⁴ Written in the computational language MATLAB, DARWIN is an adaptation of other programs created during the above mentioned broad research: axG_thG and DV. Both are destined to production of concrete musical variants from basic forms through application of transformational operations and are inserted in a genetic algorithm complex, named *geneMus*. Their versions as applications are available for download in the site www.musmat.org

⁵ For sequences of intervals, the unity corresponds to one semitone, and the signs “+/-“ to intervallic directions (ascendant/descendant). In the case of rhythmic sequences, the unity is equal to a 16th note, and the signs “+/-“ represent, respectively, sound and rest.

⁶ Selected genes appear in bold.

⁷ For a detailed description of the algorithm that calculates Cs, see ALMADA (2013b).

⁸ For example: “choose the lowest Cs of each breed (looking for maximizing the divergences)”; “alternate the mutational agents along the generations”; “use recursive strategies” (i.e., applying always the same mutational agent to the selected results), etc.

⁹ At first glance, considering the current number of six mutational agents, there will be 6²⁰ (approximately 3x10¹⁵!) possible lineages. However, the total of possibilities is much greater, since the mutations, by definition, occur affects a random gene. In other words, two applications of a given mutational agent to a referential form will probably produce different results. This implies an existence of – also in musical evolutionary domain – of an endless “genetic space” for a given original form.

¹⁰ See, for example, LERDAHL & JACKENDOFF (1983: 51-53).

¹¹ See DOWLING (1978) and MARVIN (2009), among others.

¹² In this respect McCADAMS & MATZIN (2001) present a very interesting discussion about the limits of variation under a cognitive/empirical perspective. Their results point to the existence of a kind of fuzzy area separating similarity and contrast.

¹³ One can also argue which of the two types of Cs – relative or absolute – could be the most adequate index for measurement of similarity. In this respect, it is interesting to observe the behavior of both curves considering generations 10 to 20 (Figure 3): while the relative Cs decreases almost linearly, the absolute Cs outlines a “sawteeth” trajectory, alternating deviations and regressions of similarity.

¹⁴ The expansion of the number of mutational agents is another possibility for improvement of the program.