

3A: mAchine learning Algorithm Applied to emotions in melodies

Cláudio Gomes¹, Josue Barros¹, Marco Leal¹, Thiago Nascimento¹

¹Departamento de Ciências Exatas e Tecnologias - Universidade Federal do Amapá (UNIFAP)
Caixa Postal 261 – 68.906-970 – Macapá – AP – Brazil

josue.barros1986@gmail.com, {claudiorogerio, marcoleal, thiago.nascimento}@unifap.br

Abstract. *At every moment, innumerable emotions can indicate and provide questions about daily attitudes. These emotions can interfere or stimulate different goals. Whether in school, home or social life, the environment increases the itinerant part of the process of attitudes. The musician is also passive of these emotions and incorporates them into his compositions for various reasons. Thus, the musical composition has innumerable sources, for example, academic formation, experiences, influences and perceptions of the musical scene. In this way, this work develops the mAchine learning Algorithm Applied to emotions in melodies (3A). The 3A recognizes the musician's melodies in real time to generate accompaniment melody. As input, The 3A used MIDI data from a synthesizer to generate accompanying MIDI output or sound file by the programming language Chuck. Initially in this work, it is using the Gregorian modes for each intention of composition. In case, the musician changes the mode or tone, the 3A has an adaptation to continuing the musical sequence. Currently, The 3A uses artificial neural networks to predict and adapt melodies. It started from mathematical series for the formation of melodies that present interesting results for both mathematicians and musicians.*

1 Introduction

In all cultures music is present. Regardless of musical intentions, there was a process of musical creation. In this way, the musical composition is present in the part of a rich and particular process in the musical production of each environment. The composition can be the combination of sounds and silences within a space and time. This way, there are two important agents, the generator, or musician, and the listener [1].

The musician, during the written moment of composition, defines different ways the construction of his work, and style with a sonorous interaction. This construction is defined by their experiences and musical knowledge. Depending on the musical discourse, the musician can use various tools, instruments: guitar, saxophone, trumpet, flute; as well as digital instruments: synthesizers, percussion, and digital drums, etc.

With the emergence of the MIDI protocol, it allowed the development and appearance of countless digital synthesizers such as Casio, Korg, Roland, etc. There was also interest in musical research centers for the alternative development of modules or systems of sound synthesis or by MIDI protocol. Thus, with the development of instruments compatible with MIDI technology, the possibilities of output and the facilities with the control and verification in the construction and composition of melodies [2].

The technological advance allowed several solutions that help the musician in his compositions. The process of computer-assisted composition is an example in evolution that allows greater clarity and affinity of the musician to the computer [1, 3]. The LZ77 algorithm was used from parts of a musical base for the generation of melodies [4]. The LZ77 algorithm is a generic algorithm that uses the Euclidean distance concept to select the best fitness among the chunks or parts of each song in genotype generation. In addition, interesting melodies can be generated from inheritances from different parts of the music base. Thus, this model pre-processes to generate new melodies.

2 3A algorithm

The self-organizing map (SOM) is a learning algorithm used in some applications: image recognition, text, financial market, etc [5]. Basically, the SOM has input parameters and a single output. This learning consists of adapting numerous input possibilities, and without increasing the cost of processing, obtaining satisfactory results for each application. Each input parameter has specific importance to the evolution of algorithm learning, that it could vary temporarily. Thus, at each moment the inputs are evaluated to the algorithm learns and presents the best alternative.

The 3A utilizes a midi synthesizer device connected to a desktop or notebook computer by processing this data from the Chuck language to output audio or midi, as shown in figure 1. For this initial works, the 3A consists in a SOM with MIDI inputs and output. The inputs are 3 MIDI notes played temporarily, the output is MIDI note played as a melody to suggestions for accompaniment to musician in real time. Based on inputs, the 3A decides the harmonic field played by musician. The harmonic field is the map of learning and its neurons are the notes played temporarily, as seen in figure 2.



Figure 1: General view from 3A

In figure 2, the first note in SOM is the specific

musical tone, in other words, in all lines the first note is the music field tonal and the other notes are part of musical tones. Each MIDI note played activates respective neurons, for example, if the G note was played, 3A activates 5th neuron from the first line, 1st neuron from the second line, 4th neuron from the third line. After three sequential notes played, 3A identifies among the neurons which showed the highest synapse activated, deciding a musical tone and generating MIDI notes. It is considered the highest synapse activated the horizontal line with more notes marked in SOM. In case of there is a duality of synapse identification, in other words, lines in same quantity of neuron activated, the algorithm selects the leftmost synapse neuron. At each moment, 3A decides the harmonic field and generates notes, that it is the melody for accompaniment.

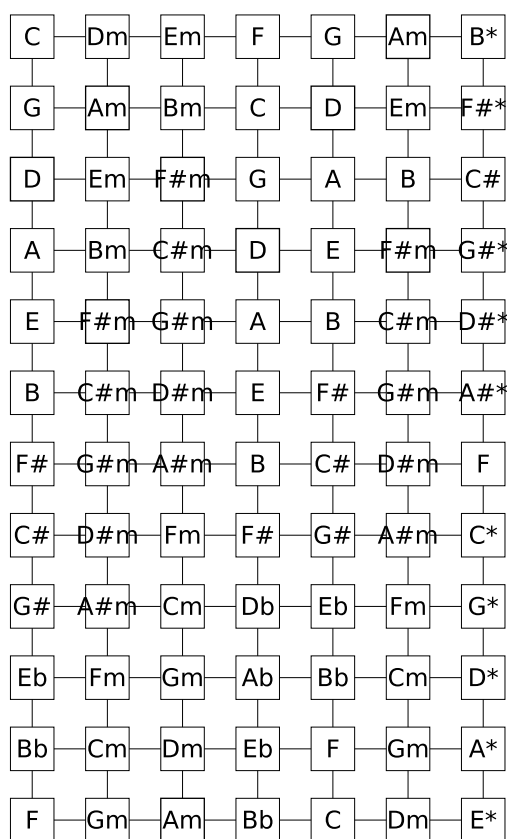


Figure 2: Self-organizing map of 3A

The 3A finds the musical tone among various to each line, and after that, suggest a note to accompany the input note in a same musical tone. This result can be transcribed in audio or MIDI. In figure 3 illustrates the machine learning model (3A) and still one of the activation neuron synapses.

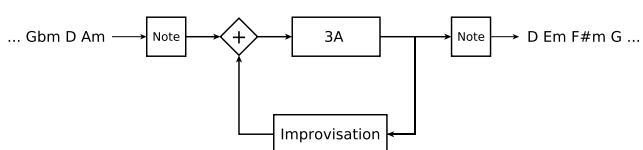


Figure 3: Structure of 3A algorithm

In first set of experiments, the sequence of notes

or silences generated by 3A follows a normalized random model that vary in the temporal dimension between whole, half, quarter, and quaver. For tests, a score made by melody based on the series prime numbers was used, because this sequence does not have mathematically completely identified a pattern. The algorithm was implemented from the Chuck language [6]. The results obtained show that the sequence acquired by 3A varies according to the initiation of weights for each input parameter. Regarding the architecture of tests, it is being used models of embedded systems like the model Raspberry Pi[7].

References

- [1] Jose D Fernández and Francisco Vico. Ai methods in algorithmic composition: A comprehensive survey. *Journal of Artificial Intelligence Research*, 48:513–582, 2013.
- [2] International MIDI Association et al. Midi musical instrument digital interface specification 1.0. *Los Angeles*, 1983.
- [3] John Strawn and Curtis Roads. *Foundations of computer music*. MIT Press, 1985.
- [4] Manuel Alfonseca, Manuel Cebrián Ramos, and Alfonso Ortega. A fitness function for computer-generated music using genetic algorithms. *WSEAS Transactions on Information Science and Applications*, 2006.
- [5] Teuvo Kohonen. *Self-organization and associative memory*, volume 8. Springer Science & Business Media, 2012.
- [6] Ge Wang, Perry R Cook, and Spencer Salazar. Chuck: A strongly timed computer music language. *Computer Music Journal*, 39(4):10–29, 2015.
- [7] Raspberry Pi Foundation. Raspberry pi. <http://www.raspberrypi.org>, may 2019. Accessed: 2019-05-05.