



A Deep Learning Algorithm for the Development of Meaningful Learning in the Harmonization of a Musical Melody

Michele Della Ventura^(✉)

Music Academy “Studio Musica”, Department of Music Technology, Treviso, Italy
micheledellaventura.mdv@gmail.com

Abstract. The interest of musicians and computer scientists in AI-based automatic melody harmonization has increased significantly in the last few years. This research area has attracted the attention of both teachers and students of Theory, Analysis and Composition, looking for support tools for the learning process. The main problem is that the systems designed and developed up to now harmonize a melody written by a user without considering the didactic and therefore cognitive aspects at the basis of a “significant learning”: given a melody, the system returns a harmonization finished without any user input. This paper describes a self-learning algorithm capable of harmonizing a musical melody, with the aim of supporting the student during the study of Theory, Analysis and Composition. The algorithm, on the basis of the ascending and descending movement of the sounds of the melody (soprano), proposes the sounds for the bass line: the Viterbi algorithm was applied to evaluate the probability of the best match between the melody sounds and the provided Markov chains, to reach the “optimal” state sequences. Subsequently, the algorithm allows the user to complete the chords for each sound of the bass line (tenor and alto), or to create the complete chords. Examples of musical fragments harmonized in this way demonstrate that the algorithm is able to respect the concatenation rules of the tonal functions which characterize classical tonal music.

Keywords: artificial intelligence · Hidden Markov Model · melodic harmonization · significant learning · symbolic music generation

1 Introduction

Even if the use of artificial intelligence in the music composition process is nothing new, in recent years it has become a solid reality, destined to have more and more space. One of the main fields of research interest concerns the possibility of creating the accompaniment (or harmonization) of a melody. There have been several efforts made towards this task in the past, using different approaches: from hidden Markov models (HMMs) [1–3] to deep learning models [4, 5]. These systems had in common the goal of educating the computer (so that it could learn autonomously from various situations) rather than programming it, as happened in Rule-Based Algorithms [6–8]: algorithms based on specific rules of musical grammar mathematically formalized.

Nowadays, there are dozens of platforms that use AI for the harmonization of a melody or for automatic musical composition. The most important platforms carried out to date include the following: Flow Machines, IBM Watson Beat, Google Magenta's NSynth Super, Jukedeck, Melodrive, Spotify's Creator Technology Research Lab e Amper Music. Most of these systems work using deep learning networks, a type of artificial intelligence that depends on analyzing large amounts of data.

In many cases, assisted music composition systems are used as a support tool for the teaching and learning activities of the Theory, Analysis and Composition discipline. The aim would be to stimulate the creativity of the student, i.e. his ability to produce ideas and objects that are new, original, appropriate, and to which a value is attributed, which can be of a social, spiritual, aesthetic, scientific and technological nature [9]. However, a passive use of these tools could affect the student's active learning [10]: what is called "meaningful learning" [11, 12] would thus be lacking.

Starting from the assumption that the learning process must allow the student to develop the skills useful for the specific discipline, this paper presents a new algorithm (that taking inspiration from the previous algorithms) able to support the student in harmonizing a musical melody: given the sounds of the melodic line, it is able to define the sounds of the bass line, leaving the student with the possibility of completing the harmony independently. For each melody sound, the Viterbi algorithm was applied to evaluate the probability (defined through the Markov chains) of the best matching of the bass line sound.

The structure of this paper has been organized as follows. In the Introduction, the context of this study is presented, followed by a review of related studies on automatic melody harmonization and an analysis of the characteristics that these systems present in order to define the research goals. Section 2 explores the concept of "significant learning", which is the basis of the proposed algorithm. This is followed (Sect. 3) by a description of the (mathematical) method used to achieve the goal. Section 4 shows some experimental tests that illustrate the effectiveness of the proposed method. Finally, in Sect. 5 the paper ends with concluding remarks on the current issues and future research possibilities with respect to the efficient enhancement of educational practices and technologies.

2 Harmonization and Significant Learning

The first step in the study of musical composition is to know the musical grammar rules through the 4-voice harmonization (bass, tenor, alto and soprano) of a bass line (see Fig. 1b). Above each sound of the bass line, the sounds that make up the respective musical chord must be arranged so as to obtain a melody (soprano), that is, a succession of sounds that have different heights and which together with the sounds of the other voices (bass, tenor and alto) form a harmonic texture, a music that is pleasant and pleasing to the ear [13].

The musical chord (built on a specific degree of the musical scale) is a set of three notes [14]: the root note, and intervals of a third and a fifth above the root note (see Fig. 1a). As can be seen in Fig. 1b, in 4-voice harmonization there is always a sound that is doubled (doubled sound).

a)

Scale degrees: I II III IV V VI VII

Fifth
Third
Root

b)

Scale degrees: I I VII I V VI

Doubled sounds

● Soprano (line)
● Alto (line)
● Tenor (line)
● Bass (line)

Fig. 1. Example of harmonization of a bass line.

The next step consists in harmonizing (always for 4 voices) a melodic line (soprano). This is a more complex operation than the previous one. The note of the bass line directly represents a degree of the scale with respect to the tonality of the piece of music and therefore it is sufficient to arrange the sounds of the other voices according to the previous chord (concatenation of chords). Instead, for each note of the melody it is possible to find different chords with which to harmonize it, according to the needs and the sound one intends to obtain [15]. This means arranging a melody in a harmonic context, submitting a sound texture to the music [13], choosing and evaluating the alternatives among the possible chords [16]. For example, the note “Do” can belong to the following chords (see Fig. 2):

- all chords where C is the root;
- all chords that contain C as a third;
- all chords that contain C as a fifth;
- all chords that contain C as the seventh.

Fig. 2. Examples of possible chords for the “C” note of the melody.

It is evident that the harmonization of the melody presents greater difficulties and requires more time to carry out than the harmonization of the bass line. Furthermore, the student must possess specific skills and competences which allow him to observe a note from several points of view. All this could become a pretext for the student to justify the use of platforms capable of automatically harmonizing a melody, without being aware of the result obtained.

The type of study that should prevail in the theory, analysis and composition learning process is what is called “significant learning”, that is to say the process through

which new information, entering into relation with pre-existing concepts, acquire a deep meaning, linked to a variety of information and contexts. This allows us to remember the acquired knowledge for a long time and to really understand the meaning of what we are learning [17]. It is a mechanism made possible by the active attitude towards what the student has to learn and by the connections he is able to make with the information he/she already possesses (given that understanding the connections between the various elements requires an effort and a more complex operation than learning a simple definition). This type of learning not only allows for cognitive development, but also increases the student's sense of self-efficacy, defined by Albert Bandura as the awareness of being able to dominate specific knowledge and situations [18]. It is a consequence that derives from the greater mastery of information, their links and the contexts in which they apply [19].

In order to achieve meaningful learning, therefore, the motivation to learn actively is fundamental, and therefore the tools that the student can use in the learning process are also important.

The algorithm presented in this article is inspired by these considerations and is proposed as a support tool for the teaching activity of the Theory, Analysis and Composition discipline, because, given a melody, it is able to suggest to the student possible solutions for the bass line, leaving the student to complete the tenor and contralto lines, taking care not to make any mistakes in the musical grammar. In this way the student has the possibility to observe (and memorize) how the movements of the melody and the bass line vary without the interference of other sounds. Each bass line proposed by the algorithm tries to respect the harmonic functions of the chords, functions taken from Schenkerian analysis [16] which give a piece of music different intentions depending on their resolution/concatenation with the preceding chords and they follow.

3 Methodology for Melody Harmonization

This paragraph illustrates the method used by the algorithm to analyze the melody and propose the sounds of the bass line consistent with the theories described in the previous paragraphs.

The algorithm takes its cue from an important assumption of the harmonization rules: a bass sound can be harmonized in different ways in order to obtain a better melodic line or to avoid errors in musical grammar [14]. This paper does not illustrate the rules of musical grammar and not even the errors that it asks to avoid in musical harmonization: in the first place because a manual should be illustrated and there would be the risk of creating confusion for the reader who is inexperienced in the field of music; secondly, because the algorithm was designed without presetting any rules of musical grammar, as its goal is to analyze the melody trend and define the sounds of the bass line (and subsequently those of the other voices). Therefore, harmonizing a sound means not only deciding which sounds must compose the chord, but also which sound must be doubled (see example in Fig. 1). The disposition of the sounds in the 4 voices (harmonization) and the trend of the melody are two things connected to each other since the first determines the second and vice versa: in Fig. 3a with the same bass line there is a melody different while in Fig. 3b with the same melody there is a different bass line, according to the sounds of the chord, their disposition and their doublings.

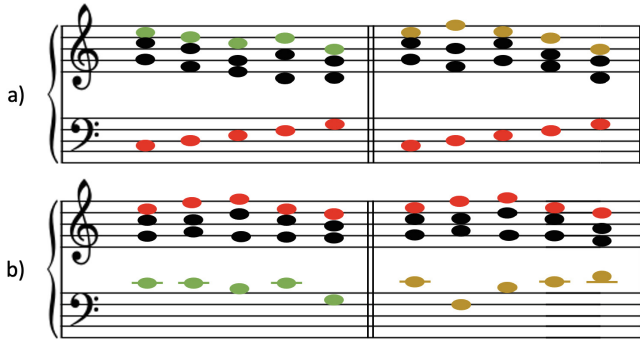


Fig. 3. Example of the relationship between harmonization and melody progression.

Analysis of the Melody

In the previous paragraph it was highlighted that each sound of the melody can belong to different chords and this can determine a different harmonization: a different degree of the musical scale and therefore different sounds and doublings.

The model developed for the automatic harmonization of a melody includes a self-learning phase in which the algorithm, through the reading and analysis of musical scores written in the form of a 4-voice choral, defines:

- (1) the degrees of the scale for each chord underlying a sound of the melody (see Fig. 4);
- (2) the ascending (a) or descending (d) trend between two consecutive sounds of the melody (see Fig. 4);
- (3) the distance (in semitones) between two consecutive sounds of the melody (called musical interval) (see Fig. 4).

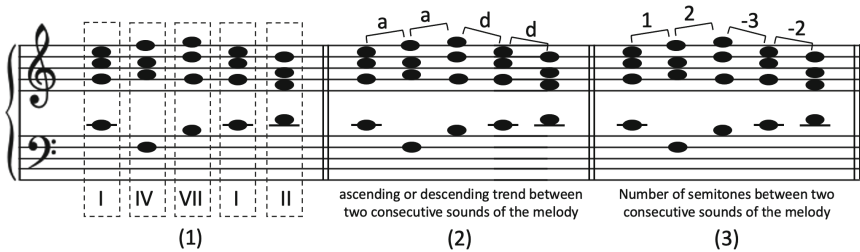


Fig. 4. Analysis of the musical melody.

This type of analysis can be done using the Markov process [20]. From reading musical scores it is possible to construct a transition matrix

$$P = (p_{ij}) \tag{1}$$

in which are represented: the probabilities that one degree of the musical scale X_d resolves on another degree of the musical scale based on the ascending or descending

movement of the melody and the number of semitones between the various sounds of the melody:

$$p_{ij} = P(X_{d+1} = j \mid X_d = i) \tag{2}$$

Figure 5 shows an excerpt of the transition matrix: column 2 and row 2 refer to the degrees of the musical scale; column 3 and row 3 refer to the number of semitones which separate the sound of the melody of the first chord (indicated in column 2) and the sound of the melody of the second chord (indicated in line 2); column 4 and row 4 refer to the ascending or descending movement of the melody as it moves from one chord to the next.

		FUTURE STATE																						
		I										II												
		1	2	3	4	5	1	2	3	4	5									
		a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	
I	1	a		0,014		0,014																		
	d																							
	2	a																						
	d																							
	3	a																						
	d																							
	4	a																						
	d																							
	5	a																						
	d																							
...																								
...																								
II	1	a		0,003																				
	d																							
	2	a		0,092		0,153	0,015																	
	d																							
	3	a																						
	d																							
	4	a																						
	d																							
	5	a																						
	d																							
...																								
...																								

Fig. 5. Example of transition matrix derived from the reading of more than 500 chords concatenation.

Definition of the Bass Line

To define the bass line (which will determine the definition of the sounds of the chord and their arrangement) it is possible to use the Viterbi algorithm associated with the transition matrix represented in Fig. 5 [21]. The probability of the most probable path ending in state k with observation “i” is

$$p_{ij}(i, x) = e_1(i) \max_k (p_k(j, x - 1) \bullet p_{ki}) \tag{3}$$

where “i” represents the probability to observe element “i” in state “I”, “j” represents the probability of the most probable path ending at position x-1 in state “k” with element “j”, and p_{ki} represents the probability of the transition from state “I” to state “k”.

The Viterbi algorithm is used to compute the most probable path (as well as its probability) [22]. It requires knowledge of the parameters of the transition matrix and

a particular output sequence and it finds the state sequence that is most likely to have generated that output sequence [23]. It works by finding a maximum over all possible state sequences [24].

Figure 6 shows an example of trellis diagram. The number of possible states depends on the musical scores used for the training phase and therefore the possibility of obtaining a bass line as coherent as possible with the tradition of musical grammar rules is directly proportional to the number of chord concatenations analyzed.

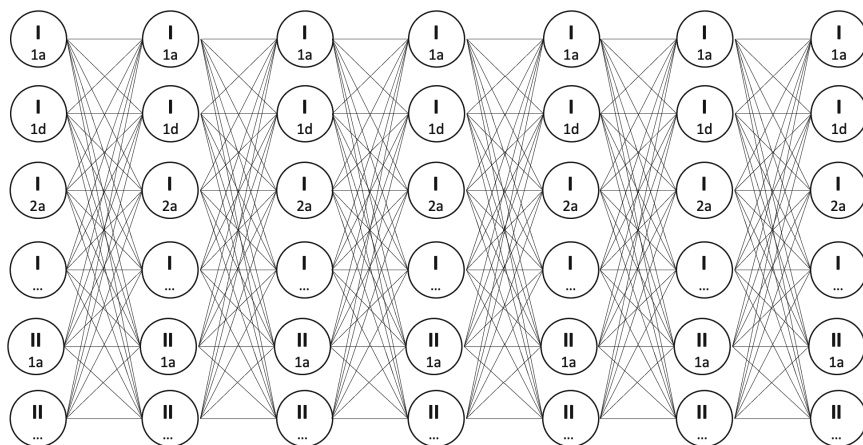


Fig. 6. Excerpt of the trellis diagram.

4 Results and Evaluation

The model presented in this paper is part of a pilot project which aims to investigate the effectiveness of the use of technologies in the teaching/learning process. In the specific case considered in this paper, the research made it possible to develop an algorithm capable of supporting the student in the study of theory, analysis and composition (or autonomously harmonizing a melodic line in the form of a 4-voice choral).

The algorithm does not provide any limitation with respect to the dimensions of the transition matrix, which is automatically dimensioned based on the characteristics of the music scores used during the algorithm training phase: non-modulating musical scores have been used (written in the form of chorale for 4 voices) in order to speed up the procedures for reading and collecting the necessary data (as described in the previous paragraph).

The algorithm was tested in 3 different steps during the training phase:

- 1) after reading about 500 concatenations of different chords,
- 2) after reading about 1000 concatenations of different chords,
- 3) after reading about 3000 concatenations of different chords.

The result supplied important information so as to be able to continue with the test. In particular, in each of the 3 verification steps, some musical melodies were proposed to the algorithm and it was observed that the bass line musically improved as the cases analyzed increased (see Fig. 7). In the first step the algorithm failed (for some proposed melodies) to conclude the bass line: this was determined by the fact that the algorithm found possibilities of movement of the melody, all with the same probability (derived from the transition matrix). In the second step (after reading about 1000 concatenations of different chords), the algorithm was able to finish the bass line even if in some cases in the last sounds (of the bass line) it proposed chords that did not give a final meaning to the musical piece. In the last step, the algorithm concluded the bass line satisfying also the musical cadence aspect.

1) I IV III I ?

2) I VI V I VII I V VII II I II III II III I

3) I IV IV III II I IV V III IV V VI IV V I

Fig. 7. Example of the results for the 3 steps.

A second type of test was performed to evaluate the algorithm. In this case, 10 students (including 2 dyslexic students) from the third year of a Music High School were involved, and they were asked to use the software while carrying out 2 exercises related to the melody harmonization. To simplify the testing procedures, some fixed test melodies have been chosen. These melodies were not randomly chosen but needed to be relatively simple and in major key rather than minor as our training data had substantially more major key musical pieces.

It was possible to notice that all the students consulted the algorithm to see the bassline it proposed. Three students then partially modified this bass line while the other students limited themselves to completing the other voices: they still had to figure out which type of chord could be inserted, which sounds were missing and which ones could be doubled. Also in this case, 2 students harmonized some sounds in two different ways, respecting (without errors) the melody and the bass.

It is therefore clear that machine learning systems are not born “ready”, with a predefined and “embedded” knowledge, but this is acquired over time, as new situations, new documents, data and information are encountered. In this case, if the algorithm encounters new cases in reading new musical scores the solutions can be greater and different.

In order to implement the machine learning mechanisms, and therefore the algorithm begins to “come to life” and to learn automatically according to the preset settings, both the training phase and the subsequent phase are fundamental (bound to the continuous availability of new data on which to retrain the model).

5 Discussion and Conclusions

Research in the field of artificial intelligence has begun to investigate the concept of metacognition [25]: the ability to “learn to learn”, knowing how to abstract from a specific domain of knowledge strategies for solving certain problems, even in a new and different context.

In this paper, a method for melody harmonization (mainly for bass line creation) has been presented: it is a part of a pilot project which aims to investigate the effectiveness of the use of technologies in the teaching/learning process. It is a method (algorithm) still under development but the first results have highlighted its potential.

The proposed method allows to obtain a bass line (as a guide for the resolution of a given exercise) to encourage the student to develop new ideas, apply previous knowledge and develop new skills: increasing his involvement and enriching the learning experience. The different harmonization proposed by the algorithm or created by individual students can be analyzed by other students and teachers, thus providing further information to increase the training dataset and refining the transition matrix, useful to guide the Viterbi algorithm to provide more successful harmonization.

New tests must be carried out both as regards the autonomy of the algorithm in harmonizing a melody (through a careful analysis of the proposed results), and as regards the number of students called to use the software as a support during the harmonization of a melody. In the latter case it is possible to evaluate the effectiveness of the algorithm as a teaching tool.

Future work can be directed in two directions. First of all, it is important to extend the training dataset, increasing the number of chord concatenations read and analyzed by the algorithm: including more complicated chords such as those of the seventh and ninth (which offer different possibilities for harmonizing the chords and therefore for their resolutions and doublings). Secondly, in addition to the task of harmonizing the melody (or generating the bass line), it would also be interesting to study the generation of a bass line conditioned by some chords inserted by the user within the melody.

The challenge will be to evolve in two parallel directions: to help students become stylistically unique and competent, and to understand how to use these new tools to enhance their creativity and explore new frontiers.

References

1. Paiement, J.F., Eck, D., Bengio, S.: Probabilistic melodic harmonization. In Conference of the Canadian Society for Computational Studies of Intelligence, pp. 218–229 (2006)
2. Tsushima, H., Nakamura, E., Itoyama, K., Yoshii, K.: Function- and rhythm-aware melody harmonization based on tree-structured parsing and split-merge sampling of chord sequences. In: Proceedings of the International Society for Music Information Retrieval Conference (2017)
3. Ventura, M.D.: A self-adaptive learning music composition algorithm as virtual tutor. In: Maglogiannis, I., Iliadis, L., Macintyre, J., Cortez, P. (eds.) Artificial Intelligence Applications and Innovations. AIAI 2022. IFIP AICT, vol. 646, pp. 16–26. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-08333-4_2
4. Briot, J.P., Hadjeres, G., Pachet, F.D.: Deep learning techniques for music generation: a survey (2017). <http://arxiv.org/abs/1709.01620>
5. Lim, H., Rhyu, S., Lee, K.: Chord generation from symbolic melody using BLSTM networks. In: Proceedings of the ISMIR 2017 (2017)
6. Boenn, G., Brain, M., De Vos, M., Ffitch, J.: ANTON: a rule based composition system. In: Proceedings of the International Computer Music Conference 2011, University of Huddersfield (2011)
7. Everardo, F., Aguilera, A.: Armin: automatic trance music composition using answer set programming. *Fundam. Inform.* **113**(1), 79–96 (2011)
8. Kippen, J., Bel. B.: Computers, composition, and the challenge of “new music” in modern India. *Leonardo Music Journal*, Massachusetts Institute of Technology Press (MIT Press), 1994, vol. 4, pp.79–84 (1994)
9. Bartoli, G.: *Psicologia della creatività. Le condotte artistiche e scientifiche*. Roma, Monolite, 2005, p. 19 (2005)
10. Ahmed, A.A., Ganapathy, A.: Creation of automated content with embedded artificial intelligence: a study on learning management system for educational entrepreneurship. *Acad. Entrep. J.* **27**(3), 1–10 (2021)
11. Ausbel D.P.: *Educazione e processi cognitivi*. Ed. Franco Angeli, Milano 2004 (2004)
12. Novak, J.: *L'apprendimento significativo*. Ed. Erickson, Trento 2001 (2001)
13. Adorno, T.W.: *Kierkegaard. La costruzione dell'estetico*, Longanesi, Milano 1983 (1983)
14. Coltro, B.: *Lezioni di armonia complementare*. Ed. Zanibon (1997)
15. Schoenberg, A.: *Theory and Harmony*. Univ of California Pr; Reprint edition (1992)
16. de la Motte, D.: *Manuale di armonia*, Bärenreiter (1976)
17. Novak, J.: *L'apprendimento significativo*. Erickson, Portland (2001)
18. Bandura, A.: *Social Foundations of Thought and Action*. Prentice-Hall, Englewood Cliffs, NJ (1986)
19. Bandura, A.: *Self-Efficacy: The Exercise of Control*. W.H. Freeman, New York (1997)
20. Ames, C.: The Markov process as a compositional model: a survey and tutorial. *Leonardo* **1989**(22), 175–187 (1989)
21. Viterbi, A.: Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Inf. Theory* **1967**(13), 260–269 (1967)
22. Forney, G.: The viterbi algorithm. *Proc. IEEE* **1973**(61), 268–278 (1973)
23. Bockmayr and Reinert (2011). *Markov chains and Hidden Markov Models*. Discrete Math for Bioinformatics WS 10/11
24. Della Ventura, M.: Human-centred artificial intelligence in sound perception and music composition. In: Abraham, A., Pllana, S., Casalino, G., Ma, K., Bajaj, A. (eds.) *Intelligent Systems Design and Applications*. ISDA 2022. LNNS, vol. 646, pp. 217–229. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-27440-4_21
25. Flavell, J.H.: First discussant's comments. What is memory development the development of? *Hum. Dev.* **14**, 272–278 (1971)