Wojciech ZABIEROWSKI
TECHNICAL UNIVERSITY OF LODZ, DEPARTMENT OF MICROELECTRONIC AND COMPUTER SCIENCE,
Wólczanska 221/223, 90-924 Łódź

Accord sequences generation in agreement with harmony tonal rules using cascade artificial neural networks

Ph.D. Eng. Wojciech ZABIEROWSKI

He received the M.Sc. and Ph.D. degrees from the Technical University of Lodz in 1999 and 2008, respectively. He is an author or co-author of more than 80 publications: journals and most of them - papers in international conference proceedings. He was reviewer in four international conferences. He supervised more than 90 Msc theses. He is focused on internet technologies and automatic generation of music. He is working in linguistic analysis of musical structure.

e-mail: wojtekz@dmcs.pl

Abstract

The paper presents research related to domain of Computer Generated Music [1], namely it treats on accords sequence generation. The goal of the research is obtained with application of neural networks. Data preparation was conducted according to the tonal harmony theory. The paper also includes a discussion of the obtained results with respect to the actual process of music composition.

Keywords: tonal harmony, music, neural network, computer generated music.

Generacja sekwencji akordowych w zgodzie z zasadami harmonii tonalnej przy użyciu kaskadowego użycia sieci neuronowych

Streszczenie

Artykuł prezentuje badania dotyczące dziedziny pt. Computer Generated Music, a dokładnie generacji sekwencji akordowych. Przyjęto założenie, że akord to trójdźwięk. Pod uwagę brane były akordy w postaci zasadniczej wraz z pierwszym i drugim przewrotem w układzie skupionym. Do rozważań brano pod uwagę tonację C-dur jak podstawę odniesienia do przyszłej generalizacji wniosków i rozwiązań. Do rozwiązania zagadnienia zostały użyte sieci neuronowe w formie kaskady i sposób ich użycia jest nowym ich zastosowaniem. W rozwiązaniu zostały użyte także metody umożliwiające losowy wybór rozwiązania w celu uatrakcyjnienia powstałych sekwencji.. Przedstawiono algorytm, zgodnie z którym następuje finalne tworzenie i dobór trójdźwięków. Uzyskane w wyniku działania algorytmu produkcje, zostały przedstawione w postaci tradycyjnego zapisu nutowego w metrum 3/8. Do prezentowanej produkcji wykorzystano różne wartości rytmiczne w celu uatrakcyjnienia brzmieniowo-rytmicznego. Dane zostały przygotowane na podstawie zasad teorii harmonii tonalnej. Artykuł zawiera dyskusje uzyskanych wyników w odniesieniu do procesu kompozycji. Przedstawiono także dyskusję skuteczności sieci neuronowych do zastosowania w tworzeniu muzyki tonalnej.

Słowa kluczowe: harmonia tonalna, muzyka, muzykologia, sieci neuronowe, muzyka generowana komputerowo.

1. Introduction

There has been a continuous research throughout the world in the field of computer generated music. Many implementations as well as a lot of disputes on this topic have appeared and it is now perceived as a branch of science.

The computer participation in the process of music creation is not only restricted to being an aid for a composer. In fact, a machine is able to replace human in the creation process. There are some realizations that are purely machine-generated and which were able to reach the top of pop music charts, successfully competing with human-made works.

2. Problem specification

A music piece can be described with many music parameters. They define what we recognize as a music composition – a piece of art.

Notes, and thus sounds defined by them, do not constitute music. Moreover, sounds grouped in chords still remain more or less an ordered set of sounds/tunes.

One of the basic parameters which define music is a harmonic meaning of music constituents mentioned above.

Of course, harmonic meaning is here understood as a specific consonance of sounds. Basic harmonic constructs [2, 3, 4]**,** referred also to as functions, are called the harmonic/primary triad. Based upon scale degree, on which built, they are called; tonic (T), subdominant (S), and dominant (D).

It is the sequence of such harmonic functions that constitutes music. In addition, some subsidiary functions can be defined. Due to resemblance to the basic functions, they have names derived from their prototypes: e. g.: T^{II} – second order tonic. Subsidiary function usage may largely enrich harmonic sequence.

Rules of succession for playing specific harmonic functions chords are strictly defined (according to the applied harmonics). The most important relations are presented in Table 1.

Tab. 1. Basic harmony function successions
Tab. 1. Podstawowe następstwa funkcji harmonicznych

Harmonic function	Allowed consequent function
Tonic	S, D, T
Subdominant	T, (D)
Dominant	T

It must be remembered that every tune starts always with a tonic (T), and every stress built by a dominant (D) is solved to a tonic (T). In simple words, the process of music composition is based on a proper progression of harmonic functions.

3. Preparation of the input data

An attempt to implement tonal harmony rules (though simplified) in generation of harmonic sequences was made. Such an approach, following the trends of Computer Generated Music, could be an introduction to further works in this domain.

The list of chords was prepared in such a way, that their harmonic function was known in the given key fragment. This process went according to assumptions, described in the further part of the paper, for one chosen key. For this work only 21 different chords according to harmonic dependencies (C, d, e, F, G, a, hdim with I and II inversion of each one) were chosen. Majority of musical dependencies was taken into consideration, in accordance with tonal harmony [1, 2]. The secondary harmonic function was taken as a function with the same harmonic meaning as each one, the main harmonic function.

Neural networks were used. After numerous tests a three-layer perceptron-based model with 32 neurons in a hidden layer was chosen; 24 input- and 21 output-neurons were related to the specific values in the teaching sets for N1. An example of mapping is presented in Table 2. Then a three-layer perceptron-based model with 32 neurons in a hidden layer was chosen; 24 input- and 3 output-neurons were related to the specific values in the teaching sets for N2. An example of mapping is presented in Table 3.

Tab. 2. Exemplary values of the corresponding input- and output-neurons used in the network architecture N1
Tab. 2. Przykładowe wartości wejściowych i wyjściowych neuronów w zastosowanej architekturze sieci neuronowej N1

Neurons In								
In1	In2	In3	..	..	…	In22	In23	In24
1	0	0	0	0	0	1	0	0
Neurons Out								
Ou1	Ou2	...	..	…	…	Ou19	Ou20	Ou21
0	0	0	1	0	0	1	0	0

Tab. 3. Exemplary values of the corresponding input- and output-neurons used in the network architecture N2
Tab. 3. Przykładowe wartości wejściowych i wyjściowych neuronów w zastosowanej architekturze sieci neuronowej N2

Neurons In								
In1	In2	In3	..	..	…	In19	In20	In21
1	0	0	0	0	0	1	0	0
Neurons Out								
Ou1	Ou2	Ou3						
0	1	0						

If a chord exists, it corresponds to '1' at the input of the related neuron. Absence of a chord is entered as '0'.

The input data was prepared in accordance with the tonal harmony rules providing sufficiently numerous dependence groups, to serve as teaching sets for the neural network.

The sets comprised about 60 sequences (the simplest case) up to 300 when dependencies between triads (also in I and II inversion of auxiliary functions) are taken into consideration.

During the generation of accords, sequences with neural networks, one chord (including their harmonic meaning) was input into it. The neural network was expected to output a proper chord that fulfils harmonic progression rules. But as known from the earlier research [10] this could provide few acceptable values. This gave us the necessity of choosing a proper chord. We used P module for draws [8]. Then the second neural network module described before went to work. As a result, it gave us the next harmonic function meaning. The result from P module gave us the next chord in the sequence. It is described as CGM in Fig. 1. The range of accords was limited to one key in order to simplify the resulting dependencies.

The schema of the working system is given in Fig. 1. The symbols in the figure denote:

- Ai – input chord vector
- Fhi – input harmonic function vector binded with Ai
- Bi – vector of results from N1 neural network
- AC – chord result given by P module
- Fh – harmonic function for next iteration

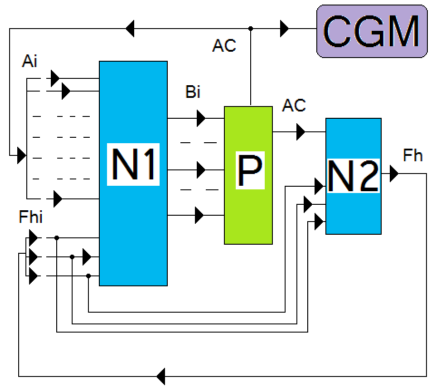

Fig. 1. Schema of proposed system
Rys. 1. Schemat działania systemu

4. Research

Learning process, depending on the method used, consisted of up to 40 000 iterations. In the process [6,7] the following methods were taken into consideration:

- Gradient descent back-propagation, (with and without momentum)
- Resilient back-propagation,

In addition, the following transfer functions were used there [5]:

- Log-sigmoid transfer function – Matlab - *logsig(n)*;

Hyperbolic tangent sigmoid transfer function – Matlab *tansig(n)*.

$$a = \mathrm{tansig}(n) \qquad a = \frac{2}{1+e^{-2n}} - 1 \tag{1}$$

$$a = \mathrm{logsig}(n) \qquad a = \frac{1}{1+e^{-n}} \tag{2}$$

As a result of the neural network teaching process, a good adjustment was achieved. Its quality was measured by the mean square error – MSE, and the accurate attribution error - Er, see Tables 4 and 5.

The magnitude of the classification error Er means a number of direct error matches with respect to the expected precise answers, and can be defined as follows:

$$Er = \frac{er}{L} * 100\% \tag{3}$$

where: er – number of erratic classifications, L – test set size.

The most interesting results for architecture with one hidden layer and 32 neurons in the layer as well as the obtained adjustment and mean squared error (MSE) are shown in Tables 4 and 5. The presented solution architecture was empirically chosen. All of the sets were prepared with the default network parameters taken into account. Various network configurations with different numbers of neurons per hidden layer were tried out. More setups were investigated but not all of them are included because of the unsatisfactory results they produced.

Concluding from the Table 4, the reached level of the mean squared error of learning process for this issue is, in fact, the same in every case. The same can be observed for the classification error *Er*, which shows similar values in different network configurations. Due to this, more attention has to be paid to the configurations that allow obtaining the similar adjustment level with less calculation overhead.

Tab. 4. The results of learning processes and their Er error values for N1
Tab. 4. Wyniki uczenia wraz z błędem Er dla N1

Network architecture	**Lerning time (iter.)**	**Transfer functions (t = tansig)**
24/32/21	40 000	t/t/t
24/32/21	40 000	t/t/t
24/32/21	600	t/t/t
Learning function	**MSE**	**Er**
traingd	0,032112	42
taingdm	0,019321	50
trainrp	0,019222	50

Tab. 5. The results of learning processes and their Er error values for N2
Tab. 5. Wyniki uczenia wraz z błędem Er dla N2

Network architecture	**Lerning time (iter.)**	**Transfer functions (t = tansig)**
24/24/3	40 000	t/t/t
24/24/3	40 000	t/t/t
24/24/3	400	t/t/t
Learning function	**MSE**	**Er**
traingd	0,133451	92
taingdm	0,193212	92
trainrp	0,192222	92

The network efficiency obtained in the proposed solution is of a limited quality. The results seem to be unsatisfactory, as only every second step of the generation process occurred to be unambiguous. Thorough analysis with respect to the tonal harmony theory shows that the alleged adjustment errors are very small or absent.

It occurred that for some combinations of harmonic progressions the networks did not propose unambiguous solutions. The network, for example, proposed two possible chords. Data analysis showed that under the specific conditions of data input into the network, the number of sounds proposed by the network was also higher. It is so because there is more than one combination that fulfils the definition of a chord in the proper harmonic meaning used in the research. Owing to this, it was possible for the network to choose different chords for the given harmonic functions and still fulfil rules of tonal harmony.

Do such results disqualify the presented method? No, if in addition the randomization and chord recognition [8] modules are added to the presented process – P module. Then the obtained results are quite different. The result ambiguity is removed and its quality is satisfactory.

The second architecture gives more interesting results without the necessity of adding a draw module. Classification of the error *Er* gives us almost the certainty that the network works correctly. It goes from the theory of tonal music. With the given chord and the previous harmonic function of the chord you have one solution of harmonic meaning in this specific case.

An important question is how the results relate to the actual process of music composition. The same dilemma of composition process is shared by a music composer [9]. It is his genius – his mind depending on emotions - that decides about sounds, chords, and harmonic dependencies.

This issue can be solved by adding another decisive stage implemented into the neural network. This stage will somehow reflect the knowledge about human emotions (as I wrote in [8]).

In the presented research, the simulation of a human-composer activity was implemented by means of the additional stages of the composition process, realized as a neural network [8]. Process randomization mechanisms for choice of the sound and harmonic meaning are also added.

The result of the system operation, as it was described, is presented as a note record for various rhythmic values/parameters. The excerpt of the note record can be treated as the excerpt of a music composition.

Fig. 2. Generated music score in 3/8 measure
Rys. 2. Uzyskany efekt muzyczny w metrum 3/8

5. Summary

Due to the features of one key notation, the solutions obtained with the assumed simplifications could be automatically transponed up and down the scale and need only small tuning. The future works on application of a neural network to the problems of Computer Generated Music are planned to be focused on including several new parameters, like rhythm or tempo, into the composition process.

The results taken from our system give us opportunity to use this accord music line for further composition process.

Another area of interest is an attempt to define, reflect and record the influence of human emotions in a form acceptable by a neural network – if only to a limited extent.

Such works will undoubtedly contribute to the worldwide progress in the Computer Generated Music domain. The future works will especially support application of the methods of computer science to musicology, including the commercial solutions used in musical instruments, composition tools, and maybe also, music-therapy and rehabilitation by sounds.

6. References

[1] Baggi L. D.: Computer-Generated Music, IEEE Computer, 1991.
[2] Lerdahl F., Jakedoff R.: A Generative Theory of Tonal Music, MIT Press, 1983.
[3] Sikorski K.: Harmonia cz. I, Polskie Wydawnictwo Muzyczne, 1991.
[4] Wesołowski Fr.: Zasady muzyki, Polskie Wydawnictwo Muzyczne, Kraków 2004.
[5] Demuth H., Beale M.: Neural Network Toolbox: User’s guide version 4, The Math Works, Inc., 2000.
[6] Tadeusiewicz R.: Sieci neuronowe, BG AGH, Kraków, 2000.
[7] Rutkowski L.: Metody i techniki sztucznej inteligencji, PWN, Warszawa, 2006.
[8] Zabierowski W., Napieralski A.: Chords classification in tonal music with artificial neural networks, Polish Journal of Environment Studies vol17/4C/2008, wyd. HARD Olsztyn, ISSN 1230-1485, 2008.
[9] Gang D., Lehamnn D.: Melody Harmonization with Neural Nets, International Computer Music Conference, Banff, 1995.
[10] Zabierowski W.: Accord generation in agreement with harmony tonal rules using artificial neural network., Pomiary Automatyka Kontrola, 2010, vol. 56, s. 1551-1553, ISSN 0032-4140.

otrzymano / received: 03.08.2011
przyjęto do druku / accepted: 03.10.2011 *artykuł recenzowany*