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

Accord generation in agreement with harmony tonal rules using 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 70 publications: journals and most of them - papers in international conference proceedings. He was reviewer in four international conferences. He supervised more than 80 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

Abstarct

The paper presents research related to the domain of Computer Generated Music [1], namely it deals with the harmonic sequence generation. The goal of the research is obtained with an application of neural networks. Data preparation was conducted using MIDI mechanisms and a music synthesizer. The paper includes also 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 akordów w zgodzie z zasadami harmonii tonalnej z wykorzystaniem sieci neuronowych

Streszczenie

Artykuł prezentuje badania dotyczące dziedziny Computer Generated Music, a w szczególności dotyczy generacji sekwencji harmonicznych. W tym konkretnym przypadku, problem następstw harmonicznych został ograniczony do następstwa akordowego przy założeniu stosowania akordów trójdźwiękowych. Podstawą rozważań jest harmonia tonalna. Do rozwiązania zagadnienia zostały użyte sieci neuronowe i sposób ich użycia jest nowym ich zastosowaniem. Zostały przebadane różne konfiguracje i metody uczenia. Dane zostały przygotowane przy wykorzystaniu mechanizmów MIDI. Artykuł zawiera dyskusje uzyskanych wyników w odniesieniu do procesu kompozycji. W przeprowadzonej dyskusji pojawiają się także propozycje rozwiązania niektórych problemów, co w zestawieniu z pewnymi wcześniejszymi pracami autora pokazuje skuteczność przyjętych rozwiązań. W opracowaniu został zamieszczony wygenerowany efekt muzyczny, w postaci zapisu nutowego, jako dowód na poprawność działania w zgodzie z zasadami harmonii 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 have appeared**,** as well as a lot of disputes on this topic, 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 a human being 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 on a scale degree they are built upon, 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 the harmonic sequence.

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

Tab. 1. Basic harmonic 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 always starts 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. Neural networks were used. After numerous tests a three-layer perceptron-based model with 32 neurons in a hidden layer was chosen. 16 input- and 16 output-neurons are related to the specific values in the teaching sets. An example of mapping is presented in Table 2.

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

Neurons In								
In1	In2	In3	In4	In5	…	In14	In15	In16
1	0	0	0	1	0	1	0	0
Neurons Out								
Ou1	Ou2	...	Ou8	Ou9	…	Ou14	Ou15	Ou16
0	0	0	1	0	0	1	0	0

To represent chord sounds with a specific harmonic meaning, a MIDI notation was used. This applies for a part of information stored in voice messages. Appearance of the specific value in a message corresponds to '1' at the input of the related neuron. Absence of a sound is entered as '0'.

Tab. 3. The values corresponding to the sounds in a MIDI notation
Tab. 3. Wartości liczbowe odpowiadające wartościom dźwiękowym w MIDI

Note	MIDI value (hex)	Decimal value (dec)
C	3c	60
D	3e	62
E	40	64
F	41	65
G	43	67
A	45	69
H	47	71

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

The sets consisted of 50 sequences (the simplest case) up to 300 when dependencies between triads (also in I and II inversion of auxiliary functions) were taken into consideration.

The input data complied with a rule that three sounds which define a chord possess a specific harmonic meaning. During the generation of harmonic sequences with neural networks, three sounds (including their harmonic meaning) were the input into it. The neural network was expected to output a proper sound chord that fulfills harmonic progression rules. The range of sounds was limited to one octave in order to simplify the resulting dependencies.

4. Research

A 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)*.

$$a = \text{tansig}(n)$$
$$a = \frac{2}{1+e^{-2n}} - 1 \quad (1)$$

$$a = \text{logsig(n)}$$
$$a = \frac{1}{1+e^{-n}} \quad (2)$$

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

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

Network architecture	Lerning time (iter.)	Transfer functions (t = tansig)
16/24/16	40 000	t/t/t
16/24/16	40 000	t/t/t
16/24/16	500	t/t/t
Learning function	**MSE**	***Er***
traingd	0,0216511	42
taingdm	0,018145	50
trainrp	0,018222	50

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

$$Er = \frac{er}{L} * 100\% \quad (3)$$

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

The most interesting results for an architecture with one hidden layer and 24 neurons in the layer, as well as the obtained adjustment and mean squared error (MSE) are shown in Table 4. 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 a 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 the 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 for 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.

The network efficiency obtained in the proposed solution is of limited quality. The results seem to be unsatisfactory, as only every second step of the generation process occurred to be unambiguous. The 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 harmonic functions. The data analysis showed that in 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 are more than one combinations that fulfill the definition of a chord used in the research. Owing to this, it was possible for the network to choose different sounds and harmonic functions and still fulfill the 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. Then the obtained results are quite different. The result ambiguity is removed and its quality is satisfactory.

An important question is, how the results relate to the actual process of music composition. 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.

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 the 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.

As it can be seen, the result is not ideal, because of noticeable parallelisms. However, this specific aspect of the harmony rules was not implemented in the presented research. Such a decision does not influence the conclusions about the whole system operations.

Fig. 1. Generated music score in 16/8 measure
Rys. 1. Uzyskany efekt muzyczny w metrum 16/8

5. Summary

Due to the features of octave notation, the solutions obtained with the assumed simplifications are automatically transposed up and down the scale and need only small tuning in the case of the sound domain extension to more than one octave. 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,. Another area of interest is an attempt to define, reflect and record the influence of human emotions in a form acceptable by the neural network – if only to a limited extent.

Such works undoubtedly will contribute to the worldwide progress in the Computer Generated Music domain. The future works will especially support application of the computer science methods to musicology, including the commercial solutions used in musical instruments, composition tools, and maybe also, to 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, vol. 17/4C/2008, wyd. HARD Olsztyn, ISSN 1230-1485, 2008.
[9] Gang D., Lehamnn D.: Melody Harmonization with Neural Nets, International Computer Music Conference, Banff, 1995.

otrzymano / received: 10.09.2010
przyjęto do druku / accepted: 01.11.2010 *artykuł recenzowany*

INFORMACJE

Informacje dla Autorów

Redakcja przyjmuje do publikacji tylko prace oryginalne, nie publikowane wcześniej w innych czasopismach. Redakcja nie zwraca materiałów nie zamówionych oraz zastrzega sobie prawo redagowania i skracania tekstów oraz streszczeń.

Artykuły naukowe publikowane w czasopiśmie PAK są formatowane jednolicie zgodnie z ustaloną formatką zamieszczoną na stronie redakcyjnej www.pak.info.pl. Dlatego artykuły przekazywane redakcji należy przygotowywać w edytorze Microsoft Word 2003 (w formacie DOC) z zachowaniem:
- wielkości czcionek,
- odstępów między wierszami tekstu,
- odstępów przed i po rysunkach, wzorach i tabelach,
- oznaczeń we wzorach, tabelach i na rysunkach zgodnych z oznaczeniami w tekście,
- układu poszczególnych elementów na stronie.

Osobno należy przygotować w pliku w formacie DOC notki biograficzne autorów o objętości nie przekraczającej 450 znaków, zawierające podstawowe dane charakteryzujące działalność naukową, tytuły naukowe i zawodowe, miejsce pracy i zajmowane stanowiska, informacje o uprawianej dziedzinie, adres e-mail oraz aktualne zdjęcie autora o rozmiarze 3,8 x 2,7 cm zapisane w skali odcieni szarości lub dołączone w osobnym pliku (w formacie TIF).

Wszystkie materiały:
- artykuł (w formacie DOC),
- notki biograficzne autorów (w formacie DOC),
- zdjęcia i rysunki (w formacie TIF lub CDR),

prosimy przesyłać w formie plików oraz dodatkowo jako wydruki na białym papierze (lub w formacie PDF) na adres e-mail: wydawnictwo@pak.info.pl lub pocztą zwykłą, na adres: Redakcja Czasopisma Pomiary Automatyka Kontrola, Asystent Redaktora Naczelnego mgr Agnieszka Skórkowska, ul. Akademicka 10, p.21A, 44-100 Gliwice.

Wszystkie artykuły naukowe są dopuszczane do publikacji w czasopiśmie PAK po otrzymaniu pozytywnej recenzji. Autorzy materiałów nadesłanych do publikacji są odpowiedzialni za przestrzeganie prawa autorskiego. Zarówno treść pracy, jak i wykorzystane w niej ilustracje oraz tabele powinny stanowić dorobek własny Autora lub muszą być opisane zgodnie z zasadami cytowania, z powołaniem się na źródło cytatu.

Przedrukowywanie materiałów lub ich fragmentów wymaga pisemnej zgody redakcji. Redakcja ma prawo do korzystania z utworu, rozporządzania nim i udostępniania dowolną techniką, w tym też elektroniczną oraz ma prawo do rozpowszechniania go dowolnymi kanałami dystrybucyjnymi.