Łukasz MAZUROWSKI

West Pomeranian University of Technology-Szczecin

An algorithmic composition - the chosen aspects

***Abstract**. The paper focuses on the issue of making the music composition process automatic that can be placed within the frames of machine learning. In the work the author presents methods for algorithmic composition based on human-defined compose rules and rules extracted from sample music. Moreover, author's algorithmic composition system is presented in which the composition process is modelled as a states sequence, where states describe selected parameters of a music pattern.*

***Streszczenie.** Praca skoncentrowana jest na problematyce automatyzacji procesu komponowania muzyki, która może być umiejscowiona w ramach machine learning. W pracy zostały ukazane metody komponowania algorytmicznego oparte na zasadach kompozycji zdefiniowanych przez człowieka oraz komponowaniu opartym na regułach uzyskanych z przykładowych utworów. Zaprezentowano również autorski system komponowania, w którym proces komponowania zamodelowano jako sekwencję stanów, będących opisem wybranych parametrów muzycznego wzorca. (**Komponowanie algorytmiczne-wybrane aspekty**).*

Keywords: algorithmic composition, MIDI, generative music.
Słowa kluczowe: komponowanie algorytmiczne, MIDI, muzyka generatywna.

Introduction

The idea of composing a song by the machine as an act of artificial music may be regarded by many critics as an unwanted intrusion into the sphere reserved and dominated by a man. Despite this skeptical point of view there are many systems of computer music composition. In particular there are a few approaches that differ in the stage of composition process autonomy: fully automatic composition systems, algorithmic composition systems, in which the composition process is automatic with a little human interference, and Computer-Aided Algorithmic Composition systems, in which the composer's work is supported by the machine. This paper is devoted to the algorithmic composition systems which are described in details in the next part.

Methods for algorithmic composition

The solutions to the problem of algorithmic composition can be grouped in two classes: the methods in which a composer defines the composition rules and the methods in which the composition rules are generated from exemplary music compositions.

In fact, there are obviously the composition systems that use the hybrid solutions in which the music composition process is based both on the rules defined using the music theory and the ones obtained from the ready music examples.

The composition methods oriented on the rules defined by the user treat the composition rules as a function of probability distribution of the occurrence of the music data sequences described by means of the selected music parameters. The best known approach is the use of Hidden Markov Model (HMM) in which the future events (such as the appearance of a particular note) only depend on the previous events. The sample solution – the CAMUS system – is presented in [1].

Moreover, solutions using cellular automates and genetic algorithms can be distinguished in the domain of composition based on the defined rules. Cellular automates defined as a cells matrix, where cells change in time according to the fixed rules and generate patterns, constitute the model in the form of patterns propagation. A composition can be treated as an optimization problem where the music creation mechanism constitutes genetic algorithm. This algorithm is a process initiated by a random individuals population (e.g. music patterns), which as a result of reproduction and mutation of selected individuals evolves in time to the next generation solutions. A critical element of this process is a fitness function whose defining is a complex problem and because of this an approach to generation interactive assessment by a human is used more frequently. The examples of the use of the genetic algorithms are described in [2] and [3].

An interesting approach is the use of Petri's nets in which music composition is based on defining music objects (single sounds, phrases, etc.) which connections are described by means of properly coded expressions. A sample solution is presented in [4].

Methods for algorithmic composition aimed at extracting music patterns from exemplary compositions constitute a subset of solutions from a machine learning domain. S. Dubnov proposed using incremental parsing and prediction suffix tree (PST) as a mechanism for separating music style from input compositions (the process is described in detail in papers [5] and [6]). Another solution is using artificial neural networks as a simplified model of the human brain (because of this using them in the artificial music composition domain appears to be intuitive) which separates patterns from the given data structures. In this area the most recent studies are described in [7] and [8].

Finally, the works of prof. David Cope [9] – [12] should be mentioned whose studies have been based on the idea of Mozart's music dice game. Cope's solution is to find and select simple music signatures from the works of music classicists (Bach, Mozart) using template matching methods. Then the templates are replicated and recombined to make a new melody. The evolving solution is a result of an augment transition network used in processing natural language.

Staying in the topic connected with natural language processing the work of Jakedoff and Lardahl should be mentioned [13]. They made an attempt to analyze the listener's perception of the music structure and indicated the similarity between the music composition analysis and syntax of a sentence in natural language analysis. The authors, raising the sphere of tonal harmony analysis, examine the mutual relations between the music objects in the structure similar to the Chomsky's tree for the natural language.

In the next part of this work author's research on the algorithmic composition system using framework developed by Eerola and Toiviainen (the MIDI Toolbox [14] implemented in Matlab) are described.

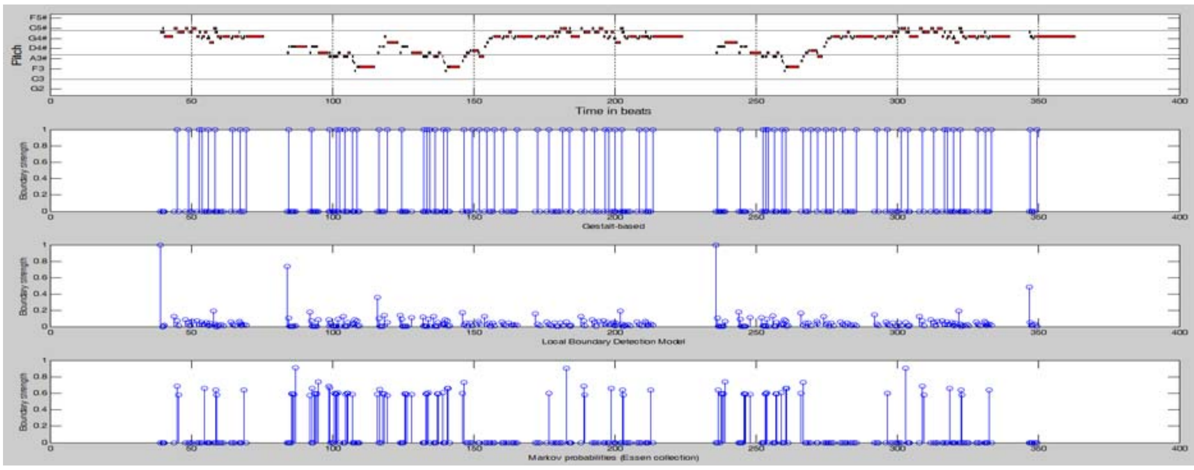

Fig. 1. Differences in the music segments detection for Abba's song Dancing Queen. On the top the pianoroll plot of the melody, below successively the segmentation results for the Gestalt-based algorithm, Local Boundary Detection Model and Markov's model.

Algorithmic composition system prototype

The presented algorithmic composition system operates in a music set described in the symbolic form (MIDI, Musical Instrument Digital Interface). The MIDI transfers information between electronic musical devices by means of a set of commands described in [15]. In the presented system only a selection of messages is used and their interpretation is a data structure that is described in the next part of the work.

Before processing music information the MIDI file is converted to the form of a matrix presented in Table 1 where columns represent following note parameters: onset time in beats (1), duration in beats (2), MIDI channel (3), pitch (in MIDI number) (4), velocity (5), onset time in seconds (6) and duration in seconds (7).

Table 1. An example of a note matrix - C4 - C5 octave

Onset (beat)	Duration (beat)	MIDI channel	Pitch	Velocity	Onset (sec.)	Duration (sec.)
0	0,4167	1	60	100	0	0,2500
0,4167	0,4167	1	62	100	0,2500	0,2500
0,8333	0,4167	1	64	100	0,5000	0,2500
1,2500	0,4167	1	65	100	0,7500	0,2500
1,6667	0,4167	1	67	100	1	0,2500
2,0833	0,4167	1	69	100	1,2500	0,2500
2,5000	0,4167	1	71	100	1,5000	0,2500
2,9167	0,4167	1	72	100	1,7500	0,2500

Then all the channels used in the composition that contain separate instrumental parts are separated from the note matrix. Every channel is assigned to one of four defined classes which describe the sort of instrumental part: melody, bass melody, accompaniment and drums.

To determine the potential channels belonging to the class of melody, it is proposed to calculate the cover coefficient of the sounds that overlap in time in the set of all the sounds that occur in the channel. The number of sounds in the channel *j* is *n*, and the number of all the channels present in the compositions (except for the channel 10) is *N*. For every *i* sound from channel *j* its position in relation to the beginning of the time row series of all the sounds in the *j* channel is calculated and it constitutes the sum of the onset and the duration value of the *i* sound. The condition of shading the sound *k* by the sound *i* fulfills the inequality (1):

$$(1)\quad \forall_{i\in 1:n-1} \forall_{j\in N} \forall_{k\in 2:n} onset_{ij} + duration_{ij} \rangle onset_{kj}$$

The total number of all the sounds that fulfill the inequality (1) was marked as *C*. The ratio of *C* to the total number of sounds in the *j* channel is the cover coefficient that has a value from the interval [0,1]. A 0 value means that in the given channel there is not the phenomenon of polyphony. A 1 value indicates that all the sounds in the channel are in consonance. The condition of belonging to the melody class is described by the inequality (2):

$$(2)\quad \forall_{j\in N} \frac{C_j}{n_j} \le threshold$$

As a result of the conducted experiments it was assumed that the *j* channel potentially connects the melody (it can belong to the melody class) if the ratio of C_j to n_j is lesser than the *threshold* 0.2. Otherwise the given channel is assigned to the accompaniment class.

After determining the melody class the channel, which stemplot dominant sounds pitch value (D_{pitch}) is minimal, is designated from among the class members. This is described by the formula (3):

$$(2)\quad \forall_{j\in melody} \min(D_{pitch\,j})$$

Channel, for which the inequality (3) is fulfilled, is marked as a member of the bassmelody class. If in the processed composition the channel 10 is used it is always marked as drums (this follows from the MIDI standard, where the channel 10 is reserved for the percussion instruments).

Separated and properly classified channels are then segmented in order to identify patterns that occur in them. In the used framework there are three methods of segmentation: one using the Gestalt-based algorithm [16], the Local Boundary Detection Model proposed by Cambouropoulos [17] and the last one - the Markov's model described in [18]. The exemplary segmentation output is shown in Figure 1 and it refers to designating segment in the Abba's song Dancing Queen.

The mentioned methods present an acceptably good division into segments, however, subsequently the model proposed by Cambouropoulos is used as the most accurately reflecting the division into segments and the one recommended by Eerola and Toiviainen in [14].

Subsequently, the matrix of possible connections is made for the patterns extracted from every channel. This matrix is used in the process of creating an output composition. An exemplary transition matrix for the accompaniment class channel and the pianoroll plot of the most common pattern are shown in Figure 2.

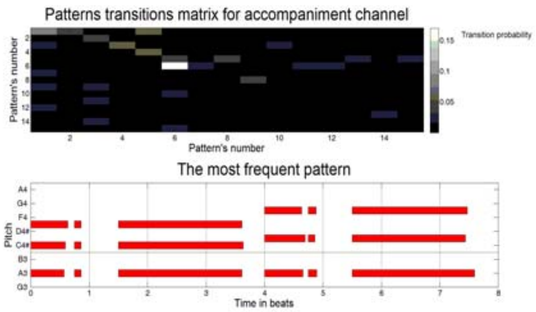

Fig. 2. An exemplary transition matrix for the patterns (top) and the pianoroll plot for the most common pattern (bottom)

After designating possible connections between the patterns the process of creating the output composition begins. The set that is going to appear in the new output composition is chosen from the patterns extracted from all the defined classes. The patterns appearing in a given class are chosen taking into account the transition matrix for the given class segments.

The final result is a music composition that is a recombined version of the input song. The recombination is based on using patterns from the input song in the new generated song. The transition matrices, which are described in the model and refer to the composition process described as a Markov's chain, constitute a basis of composition rules obtained from the exemplary input songs

In the exemplary composition process the MIDI song Kool & The Gang - Celebration was used. The output composition contains three MIDI channels: melody, drums and accompaniment classes. In Figure 3 there is a pianoroll plot for the output composition created by means of the previously described algorithmic composition system. Channels are labelled as ch0, ch1 and ch2. In the example, three patterns for drum class were used, two for accompaniment class and three for melody class.

Conclusion

Algorithmic composition systems described in the work present current problems in the music machine learning domain. In Poland this problem still stays in the scientific niche, although, what deserves an attention is the work of Zabierowski [19, 20], in which he describes the issue of using linguistic methods in reference to music creation and cascade artificial neural networks which are based on the information derived from the harmony rules - the music theory.

At this stage the author's researches focused only on finding recurrences and recombination among patterns constituting output composition. The possible directions of further research include the use of defined composition rules (consistent with the theory of music) and the operators locally modifying the patterns found in the song.

An important issue in the music composition systems is evaluation of the composition results. The problem lies in the subjective individual opinion of the evaluator. The majority of composition systems are evaluated by presenting exemplary generated songs. Some scientists suggest quantitative analysis of the composition by asking evaluators about their music preferences. The qualitative results analysis is similar to the Turing's test: the evaluators indicate which composition has been generated by a computer system and which by a man. The author suggests becoming acquainted with the results of his research in the sources presented below.

The results of the song composition methods described by the classicists (Cope, Biles, Kirnberger), and the method proposed by the author, can be heard during the broadcast "Trącić myszką" of the Polish Radio Szczecin on the webpage: http://www.myszka.org/view/Podcasty/ entitled "Komputerowy kompozytor w domowym studiu nagraniowym" and "Synteza i przetwarzanie dźwięku".

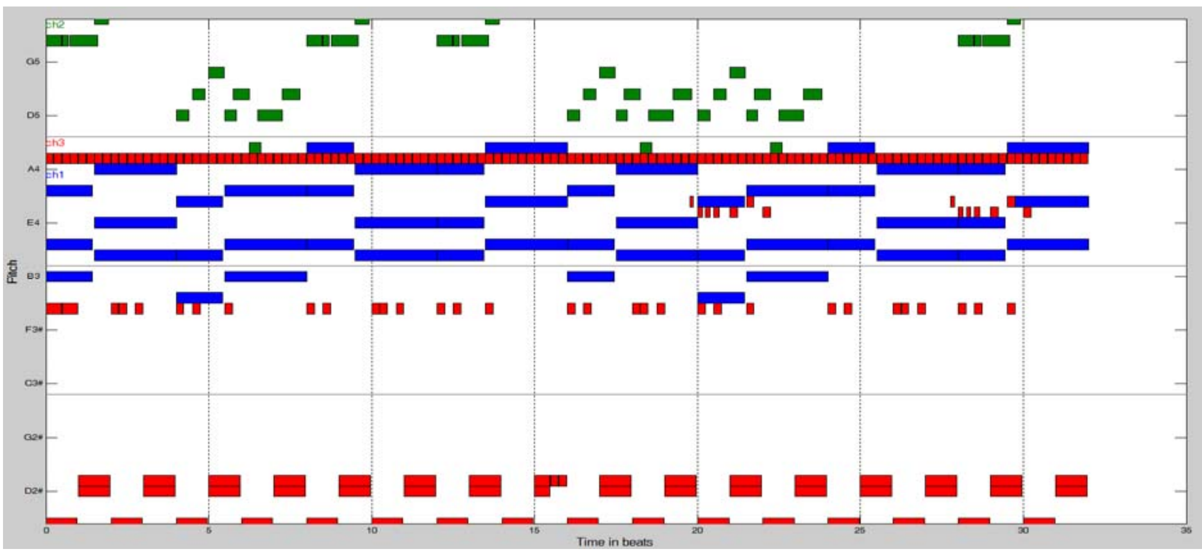

Fig. 3. The pianoroll plot for the output composition based on Kool & The Gang - Celebration. In the composition the patterns of the drum, accompaniment and melody classes were used.

Wydanie publikacji zrealizowano przy udziale środków finansowych otrzymanych z budżetu Województwa Zachodniopomorskiego.

REFERENCES

[1] Miranda E., Composing Music with Computers (Music Technology), *Focal Press*, 2001
[2] Biles J. A., GenJam: a genetic algorithm for generating jazz solos, In Proceedings of *ICMC'94*, University of Michigan Library, *Ann Harbor, MI: MPublishing*, 1994, 131-137

[3] Papadopoulos G., Wiggins G., A genetic algorithm for generation of jazz melodie solos, In Proceedings of *STEP'98*, 1998, 7-9.
[4] Goffredo, H., Sametti, A., Scoresynth: A System for the Synthesis of Music Scores Based on Petri Nets and Music Algebra, *IEEE Computer* 24*(*1991), No. 7, 56-60
[5] Dubnov S., Assayag G., Lartillot O., Bejerano G., Using machine-learning methods for musical style modeling, *IEEE Comput*. 36(2003), No. 10, 73-80
[6] Lartillot O., Dubnov S., Assayag G., Bejerano G., Automatic modeling of musical style, In Proceedings of *ICMC'02*, 2002
[7] Chen C. C. J., Miikkulainen R., Creating melodies with evolving recurrent neural network, In Proceedings of *IJCNN'01*, 2001
[8] Correa D. C., Levada A. L. M., Saito J. H., Mari J. F., Neural network based systems for computer-aided musical composition: supervised x unsupervised learning, In Proceedings of *ACM SAC'08*, 2008
[9] Cope D., Recombinant music using the computer to explore musical style, *IEEE Comp. 24 (1991)*, No. 7, *IEEE Computer Society Press Los Alamitos*, 22-28
[10] Cope D., Computer modeling of musical intelligence in EMI, *Computer Music Journal 16(1992)*, *The MIT Press*, 69-83
[11] Cope D., Experiments in musical intelligence, *A-R Editions*, 1996
[12] Cope D., The algorithmic composer, *A-R Editions, 2000*
[13] Lerdahl, F., Jakedoff, R., A Generative Theory of Tonal Music, *MIT Press*, 1983
[14] Eerola T., Toiviainen P., MIR in Matlab: The Midi Toolbox. Proceedings of *ICMIR'04*, 2004, 22-27
[15] MIDI Manufacturers Association: MIDI Message Table 1, [online] http://www.midi.org/techspecs/midimessages.php
[16] Tenney J., Polansky L., Temporal gestalt perception in music, Journal of Music Theory 24(1980), No. 2, *Duke University Press*, 205-241
[17] Cambouropoulos E., Musical rhythm: A formal model for determining local boundaries, accents and metre in a melodic surface, In Proceedings of Music, Gestalt, and Computing - Studies in Cognitive and Systematic Musicology, 1997, *Springer-Verlag*, Berlin, 277-293
[18] Bod R., Memory-based models of melodic analysis: challenging the gestalt principles. *Journal of New Music Research 31(2002)*, *Swets & Zeitlinger*, 27-37
[19] Zabierowski W., Napieralski A., The problem of grammar choice for verification of harmonisation process in tonal harmony, *TCSET'2008*, Proceedings of the International Conference "*Modern Problems of Radio Engineering, Telecommunications and Computer Science*", Lviv-Slavsko, Ukraina, 19-23 February 2008, *Publishing House of Lviv Polytechnic National University,* 2008, 86-88
[20] Zabierowski W., Accord sequences generation in agreement with harmony tonal rules using cascade artificial neural networks, *PAK 2011*, No. 11, *Wydawnictwo PAK*, 1394-1396

***Author**: mgr inż. Łukasz Mazurowski, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie, ul. Żołnierska 49, 71-210 Szczecin, E-mail: lmazurowskii@wi.zut.edu.pl*