Proceedings of the Federated Conference on
Computer Science and Information Systems pp. 733–737 ISBN 978-83-60810-51-4

Computer models for algorithmic music composition

Łukasz Mazurowski
West Pomeranian University of
Technology-Szczecin
49 Żołnierska Street, 71-210
Szczecin, Poland
Email:
lmazurowski@wi.zut.edu.pl

Abstract—**The algorithmic composition models used in the domain of systems generating music compositions are presented in the paper. Moreover, the model based on the transition matrix of music events (music patterns described by notes, measures and durations) and classification of the instrumental parts appearing in the input music work is presented. An exemplary implementation of the model is described using MIDI Toolbox implemented in Matlab. In the summary possible extensions of the presented model are described as well as the place of system functioning results in the form of output music compositions are indicated.**

I. 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 human. However, it appears that in this area the machine begins to play an increasingly important role and it becomes at least a helpful tool used in the music work composition. In the history of music machine learning development there have been many music composition computer systems and among them there are two the most important approaches that differ in the stage of composition process autonomy: 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 automatic composition systems which are described in details in the next part**.**

II. Computer Models for Algorithmic Composition

There are two main approaches to solving the problem of algorithmic composition. The first method consist of models that use the composition rules defined in an arbitrary way by a person. The second one consist of models in which the composition rules are generated from the exemplary music works.

A. The models with composition rules defined by a user

The composition models orientated to the rules defined by a user treat the composition rules as a probability distribution function of the music sequences occurrence that are described by means of 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.

Moreover, solutions using cellular automates and genetic algorithms can be distinguished in the domain of composition based on the defined rules. Cellular automates are 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. The exemplary solution – the CAMUS system – of the Markov's Model and cellular automates usage is presented in [1].

Composition can be treated as an optimization problem where the music creation mechanism constitutes a genetic algorithm. This algorithm is a process initiated by a random individuals population (e.g. music patterns), which as a result of the reproduction and mutation of the selected individuals evolves in time to the next generation solutions. A critical element of this process is a fitness function whose definition is a complex problem and because of this an approach to the generation using interactive assessment by a human is used more frequently. The examples of the use of the genetic algorithms are described in [2] and [3].

B. The models with composition rules generated from exemplary music works

The algorithmic composition models orientated to extracting music patterns from exemplary compositions constitute a large subset of the solutions from a machine learning domain. S. Dubnov in the works [4] and [5] proposed using incremental parsing and the prediction suffix tree (PST) as a mechanism for separating music style from input compositions. Another solution is applying the widely used artificial neural networks as a simplified model of the human brain which separates patterns from the given data structures. In this area the most recent studies are described in [6] and [7].

Prof. David Cope's researches constitute an important element of developing algorithmic composition [8]-[11]. He uses the idea of Mozart's music dice game and grammars describing syntax of music compositions. 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.

In the next part of this work author's research on the algorithmic composition system using a framework developed by Eerola and Toiviainen (the MIDI Toolbox [12] implemented in Matlab) are described.

III. The Algorithmic Composition Model with the Music Patterns Classification

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 [13]. 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 a music information, the MIDI file is converted to the form of a matrix presented in the 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).

In the Table 1 the octave from the sound C4 to C5 (pitches from 60 to 72) is presented. The sound pitch is represented in so called MIDI numbers. The note matrix proposed by Eerola and Toiviainen can be presented in the form of a pianoroll plot (Figure 1), which depicts the composition duration in time beats. The vertical axis contains pitch ranges. On the horizontal axis there are time ranges in beats. The rectangles represent single sounds from the composition. Figure 1 depicts pianoroll plot for the octave from C4 to C5 sounds.

In the next part of this work a pianoroll plot is used to show a piece of music generated in the described model, as a well readable graphical representation of the music data.

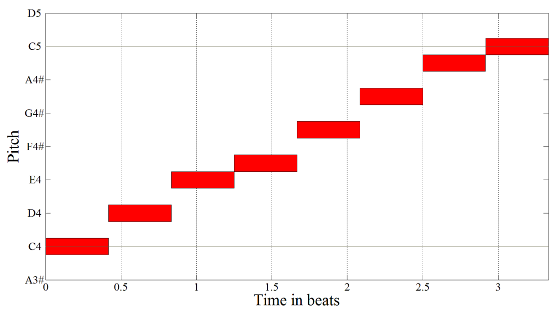

Fig. 1 Pianoroll plot of the C4-C5 octave.

The first stage in the music information processing is the separation of all the channel used in the song that contain separate instrumental parts. Every channel is assigned to one from four defined classes which describe a kind of instrumental part: *melody*, *bass*, *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 which is reserved for percussion instruments) 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 covering the sound k by the sound i fulfills the inequality (1):

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

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 the phenomenon of polyphony is absent. A 1 value indicates that all the sounds in the channel cover each other. The condition of belonging to the melody class is described by the inequality (2):

$$\forall_{j\in N} \frac{C_j}{n_j} \leq treshold \tag{2}$$

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 smaller 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):

TABLE I.
AN EXAMPLE NOTE MATRIX

Onset (beat)	Duration (beat)	MIDI channel	Pitch (MIDI number)	Velocity	Onset (seconds)	Duration (seconds)
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

$$\forall_{j \in melody} \min \left(D_{pitch_j} \right) \tag{3}$$

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

Then the processed and classified channels undergo segmentation in order to select patterns that they contain. There are three methods of segmentation in the employed framework: the method using the Gestalt-based algorithm and described in [14], the Local Boundary Detection Model proposed by Cambouropoulos in [15] and Markov's model presented in [16].

The exemplary segmentation result for the three mentioned methods is presented in Figure 2 and it applies for determining music segments in the Norway folk song Läksin Minä Kesäyönä. At the top of Figure 2 there is a pianoroll plot depicting a main melody segmentation (sounds pitch) of the mentioned song in the time function in beats. Under the plot there are stemplots depicting the probability of the boundary occurrence in the music sequence for a given note (strength of a boundary following each note) for a method based on the Gestalt-based algorithm (2a), the Local Boundary Detection Model (2b) and Markov's probabilities (2c) from the Essen collection (6,252 folk songs collected by Schaffrath in 1995, mainly from the Germanic regions, a handful of folksongs from other regions of the world).

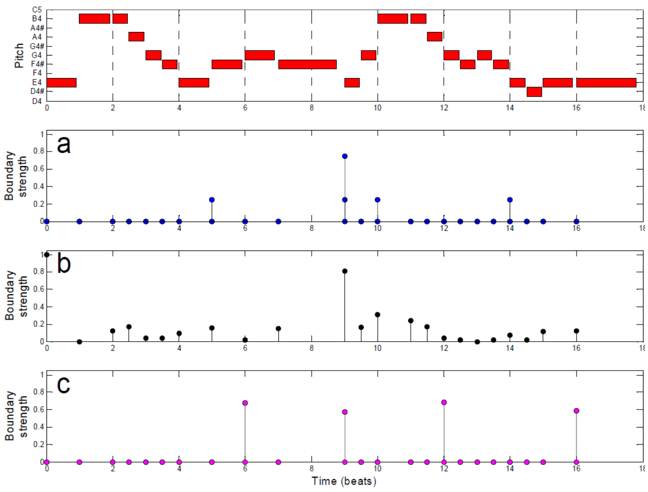

Fig. 2 The pianoroll plot of the main melody (top) of the Läksin Minä Kesäyönä and the segmentation results for the Gestalt-based algorithm (a), Local Boundary Detection Model (b), Markov probabilities from Essen collection (c). Source: [12].

The mentioned methods generate different results of the segmentation, however, next the model proposed by Cambouropoulos (Local Boundary Detection Model) and recommended by Eerola and Toiviainen in [12] was used as a one that most accurately reflects division boundaries of the sequence in the music composition.

Subsequently, the matrix of possible connections is made for the patterns extracted from all the channels. This matrix is used in the process of creating an output composition. An exemplary transition matrix for the patterns of the *drums* class channel and the pianoroll plot of the most common pattern is shown in Figure 3. The transition matrix (on the top of Figure 3) depicts probable connections between the discovered patters in the drums track in Sasha's song Ecuador. At the bottom of Figure 3 there is a pianoroll plot of the most common patters in the analyzed track.

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 consideration a transition matrix for a particular class.

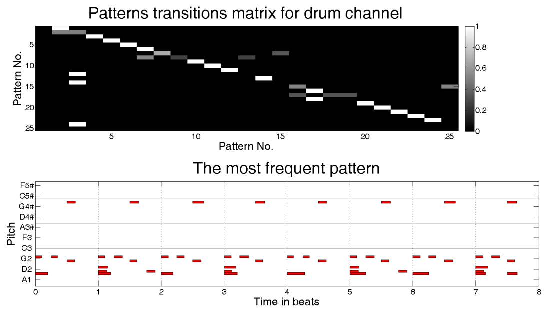

Fig. 3 Exemplary transition matrix for drum patterns (top) and the most common drum pattern (bottom) for Sasha song Ecuador.

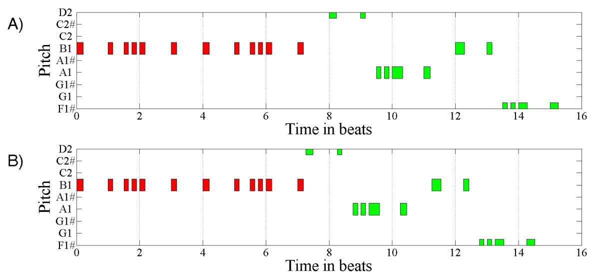

Fig. 4 Connection of two exemplary patterns: allowing of (A) and not allowing of (B) time signature (meter).

At the beginning, a theme motif is drawn from a class containing main melody patterns. The theme motif is a combination of the patterns distinguished from an input melody track and connected taking into account the probabilities appearing in the transition matrix.

After creating a new main melody the next step is creating a chord progression matching the generated melody. The rhythmical chord progression patterns connection matching is implemented similarly to the example described basing on the composition theme motif creation process. As for the main melody, the new instrumental tracks (recombinations of the input instrumental parts taking into account the transition matrices between patterns) are created for the instrumental class tracks appearing in the input composition.

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 a 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.

An important and critical point in the patterns connection process is taking under consideration rhythm. In the current approach the mechanism of patterns connection is used and it takes into account measures distribution – the smallest musical pieces implementing time signature – meter. Meter describes schema that defines the value of notes duration and the configuration of accentuations in the measure.

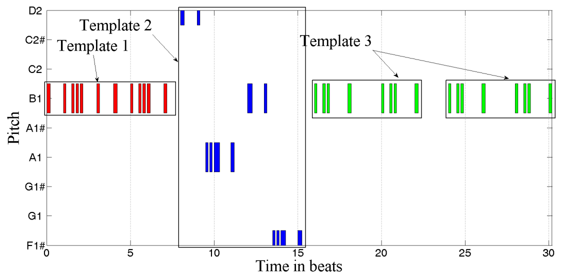

Fig. 5 A piece of an output composition belonging to the bass class.

Meter is written in the form of two numbers one above another. The bottom number describes a basic metric unit of a song (note duration). The top number describes the number of metric units (notes with a given duration) in the measure. For example, 2/4 means that there are two crotchets in the measure (number 1 stands for a semibreve, 2 for a halfnote, 4 for a crotchet, 8 for a quaver etc.). It also means that within a measure there can be metric units which duration sum is 2 crotchets, that is 1 crotchet and 2 quavers (1/4 + 1/8 + 1/8), 4 quavers (1/8 + 1/8 + 1/8 + 1/8), etc. In the described model the mechanism of searching meter using the onset distributions across a measure and autocorrelation function of onset times of notes described in [17] pages 29-31 was used.

In Figure 4 there is a connection of two exemplary patterns (in different shades) allowing of (4a) and not allowing of (4b) a song meter. Not allowing of a meter causes rhythm disturbance.

An exemplary result (a short composition) in the form of a pianoroll plot is presented in Figure 5.The patterns used in the composition are marked with different colours. The connections between the patterns allow of the song meter marked as 4/4. Every pattern used in the composition has the length of two measures and the total length of a piece is 8 measures..

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 person. The author suggests becoming acquainted with the results of his research in the sources presented in the summary.

IV. 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 [18, 19], 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 the patterns constituting the 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.

For the further research development the author suggests using genetic algorithms which can constitute a mechanism that goes beyond the limitations of the transition matrices connection. The author proposes using uncommon mutation operator for genetic algorithms. Within the frames of mutation operator the following can be used: transposition of a piece of the pattern or a single note (sound pitch shift by a particular number of semitones up or down), a note or notes replication or removal from the pattern (increase or decrease of music ornamentation) or inversion of the sound pitches or their rhythmic values.

Another possible extension of the research is generating compositions on the basis of many input songs, described by means of Markov's models and constituting common transition matrices for each attribute (domain) of a sound or a music piece. An important scientific problem is then studying the possible connections between the previously recognized patterns from the different input songs.

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". In the works [20] and [21] the author also presents his research related to the domain of the Computer Generated Music. These papers contain a discussion of the obtained results with respect to the actual process of music composition, using the genetic algorithm and Markov's chains.

References

[1] E. Miranda, "*Composing Music with Computers (Music Technology)*", Focal Press, 2001, pp. 124-129.

[2] J. A. Biles, "GenJam: a genetic algorithm for generating jazz solos" in *Proc. of ICMC*, University of Michigan Library, Ann Harbor, MI: MPublishing, 1994, pp. 131-137.

[3] G. Papadopoulos, G. Wiggins, "A genetic algorithm for generation of jazz melodie solos" in *Proc. of STEP*, Jyväskylä, Finland, 1998, pp. 7-9.

[4] S. Dubnov, G. Assayag, O. Lartillot, G. Gejerano,

[5] "Using machine-learning methods for musical style modeling". IEEE *Computer*, vol. 36, pp. 73-80, October 2003.
[6] O. Lartillot, S. Dubnov, G. Assayag, G. Bejerano, "Automatic modeling of musical style" in *Proc. of ICMC*, 2002.
[7] C. C. J. Chen, R. Miikkulainen, "Creating melodies with evolving recurrent neural network" in *Proc. of INNS-IEEE*, Piscataway, NJ, 2001,pp. 2241-2246.
[8] D. C. Correa, A. L. M. Levada, J. H. Saito, J. F. Mari, "Neural network based systems for computer-aided musical composition: supervised x unsupervised learning" in *Proc. of ACM SAC*, New York, 2008, pp. 1738-1742.
[9] D. Cope, "Recombinant music using the computer to explore musical style", *IEEE Computer*, vol. 24, pp. 22-28, July 1991.
[10] D. Cope, "Computer modeling of musical intelligence in EMI", *Computer Music Journal*, vol. 16, no. 2, pp. 69-83, Summer 1992.
[11] D. Cope, "*Experiments in musical intelligence*", A-R Editions, 1996.
[12] D. Cope, "*The algorithmic composer*", A-R Editions, 2000.
[13] T. Eerola, P. Toiviainen, "MIR in Matlab: The Midi Toolbox" in *Proc. of ICMIR*, 2004, pp. 22-27.
[14] *MIDI Manufacturers Association: MIDI Message Table 1*, [online] http://www.midi.org/techspecs/midimessages.php
[15] J. Tenney, L. Polansky, "Temporal gestalt perception in music", *Journal of Music Theory*, vol. 24, no. 2, pp. 205-41, 1980.
[16] E. Cambouropoulos, "Musical rhythm: A formal model for determining local boundaries, accents and metre in a melodic surface" in *Proc. of Music, Gestalt, and Computing - Studies in Cognitive and Systematic Musicology*, Springer-Verlag, Berlin, 1997, pp. 277-293.
[17] R. Bod, "Memory-based models of melodic analysis: challenging the gestalt principles", *Journal of New Music Research,* vol. 31, pp. 27-37, 2002.
[18] T. Eerola, P. Toiviainen, "MIDI Toolbox. MATLAB Tools for Music Research", *University of Jyväskylä: Kopijyvä*, Finland, 2004.
[19] W. Zabierowski, A. Napieralski, "The problem of grammar choice for verification of harmonisation process in tonal harmony", in *Proc. of the International Conference "Modern Problems of Radio Engineering, Telecommunications and Computer Science"*, Lviv-Slavsko, Ukraina, 2008, pp. 86-88
[20] W. Zabierowski, "Accord sequences generation in agreement with harmony tonal rules using cascade artificial neural networks", *PAK*, no. 11, pp. 1394-1396, 2011.
[21] Ł. Mazurowski, "Application of the artificial intelligence methods to the process of generating musical compositions", B.S thesis, Dept. Comp. Sci., West Pomeranian University of Technology, Szczecin, Poland, 2007.
[22] Ł. Mazurowski, "The method of generating melodic contour based on the conception of Markov's process", M.S. thesis, Dept. Comp. Sci., West Pomeranian University of Technology, Szczecin, Poland, 2009.