

# Structural Harmony Method in the Context of Deep Learning on Example of Music by Valentyn Sylvestrov and Philipp Glass

Anna Shvets  
IBM  
Armii Krajowej str. 16, 30-50 Kraków,  
Poland  
[annashvets11@gmail.com](mailto:annashvets11@gmail.com)

The article shows the potential of structural harmony method application in deep learning models training. Harmonic sequences, being represented as the paths in the system of graphs and then schematised, serve as the source of the deep learning models input. The visual character of the dataset generated using schemes allows the application of supervised machine learning technics and is suitable for time series analysis. Therefore, two neural network architectures – Convolutional and Convolutional Long-Short Term Memory were tested on example of music by postmodern composers Valentyn Sylvestrov (represented by vocal cycle “Silent songs”) and Philipp Glass (represented by piano cycle “Metamorphosis”).

*Harmonic sequences clustering. Supervised learning. Deep learning. CNN. Convolutional LSTM.*

## 1. INTRODUCTION

The development of training methods opened great possibilities in big data analysis, which influences the domains of knowledge, traditionally based on qualitative analysis, introducing quantitative analytical base. However, the selection of features being analysed, should be performed in a careful manner, as the selection of very few features (such as one musical interval) may be misleading in considering such large amount of data as the history of music styles (Nakamura & Kaneko 2018).

In fact, the core of the western music development resides in the tonality formation, its extension and dissemination, basing on functional relations and alterations within a given tonality, and the modulation strategies between tonalities: the establishment of the functional relationship began with Palestrina compositions and the decomposition of tonality manifested in the elliptic chains in the late Wagnerian, and more broadly – late romantic music. Therefore, the analysis of the functional relations within tonality should be considered as the most relevant for the research in tonal music.

## 2. STRUCTURAL HARMONY METHOD

A structural harmony method has been presented and showed its efficiency in functional relations representation (Shvets & Desainte-Catherine 2015). It comprises two levels of schematization – the system of graphs, representing a matrix for harmonic paths selection and a folded version of chosen (by the user) harmonic sequences, and the graph schemes, shadowing the graph pattern. The colour distribution on the second level of representation allows reflecting functional affiliation of each chord in the harmonic sequence. A colour distribution for functional differentiation has received a positive users’ feedback in a recent project (Malandrino, Pirozzi & Zaccagnino 2018) and showed its efficiency in the pedagogical context (Pistone & Shvets 2014). Figure 1 shows the transformation of the score representation (Figure 1a) into graph (Figure 1b) and scheme (Figure 1c) representations consequently.

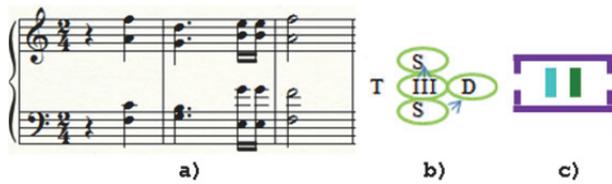


Figure 1: Harmonic sequence S-D-III-S represented using structural harmony method.

The condense chord cluster representation of the second level can now be treated as the observation unit and allows to determine the probability (allows to train a classifier) of a chord cluster appearance in the composition of a target composer.

### 3. DATA DESCRIPTION

For the purposes of this article several compositions by postmodern composers Valentyn Sylvestrov and Philipp Glass were chosen. Five cameral compositions from the vocal cycle *Silent songs* (1987) by Sylvestrov (*Ailing spirit, There were storms, wind and weather, La belle dame sans merci, A melancholy time, Goodbye world*) as well as five pieces of the *Metamorphosis* (1988) piano cycle by Glass have been selected for the analysis. Figure 2 shows a set of schemes for each selected piece.

The dataset was divided into training set, containing 11 schemes (for each composer), and a test set, consisting of 5 schemes. The training data was used to train a visual recognition model (with underlying deep learning model architecture) available at Watson Studio-7r under IBM cloud services platform *bluemix*, then the test set for each

composer has been shown to the model, which resulted in a very good learning rate: 92% for both test sets. Such an excellent result can be explained by a good data selection, which excluded the instances from one class in another class.

### 4. CNN-LSTM NEURAL NETWORK MODEL CASE

The capability of the neural network to distinguish harmonic structures appearing in music of target composers is not the only option it offers. Another use case of the structural harmony application as a dataset input to the neural network is dictated by a time series character of the data: the appearance of one scheme after another is not random and represents the dramaturgy of harmonic development of a specific composition. The exploitation of this parameter is possible with the application of Long-Short Time Memory (LSTM) neural architecture, since this type of Recurrent Neural Network (RNN) is capable of retaining the data of a time series domain (up to 400 time series). The combination of LSTM with Convolutional Neural Network (CNN) architecture gave good results in the activity recognition demonstrated in a sequence of images (Donahue et al. 2015, Sainath et al. 2015). As a derivation of this structure, a Convolutional LSTM was proposed, characterizing by convolutional structures in both the input-to-state and state-to-state transitions, and showing better performance in capturing spatiotemporal correlations (Xingjian et al. 2015). Keras library supported this architecture with a ConvLSTM layer, which was used for building a custom model.

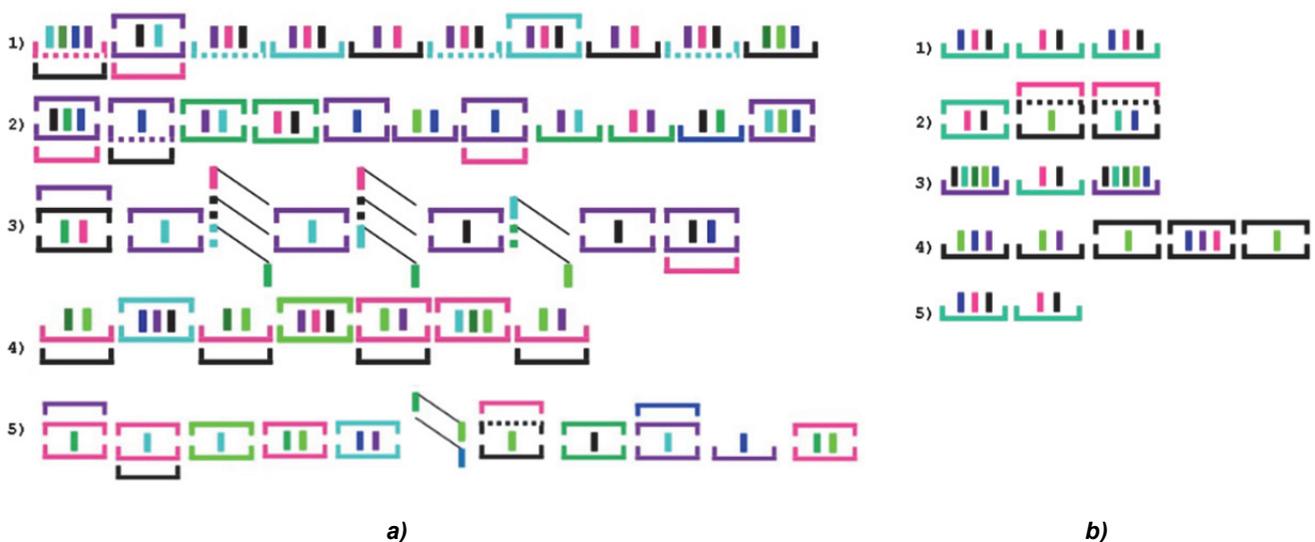


Figure 2: Schemes used for data feeding to deep learning models: a) schemes of harmonic structures for Sylvestrov's example; b) schemes of harmonic structures for Glass' music example.

As the goal of this model consisted in a sequence prediction, full schemes lists of first three pieces from *Metamorphosis* cycle were chosen for the model training (instead of separate schemes from each piece selection, which was the case for training dataset used in Visual Recognition model). The input data was transformed to *numpy* arrays and resized to 32x32 pixels with preservation of RGB channel using *CV2* library. The features were normalised and labels were vectorised using *LabelBinarizer* from *sklearn* library. The *ConvLSTM2D*, *Dropout*, *Flatten* layers along with fully-connected *Dense* layers were added to Keras sequential model. The categorical cross entropy was used as a loss function and ADAM algorithm – as optimiser. The build model gave 88% of accuracy with 18% of loss on the 100<sup>th</sup> epoch of training. The model was tested using *Jupiter notebooks* runtime environment hosted by *Google Colab*.

## 5. SUMMARY

The article showed the possibilities of structural harmony application in different neural network architectures, such as serialised Visual Recognition model (with underlying Convolutional Neural Network architecture) and custom Convolutional LSTM neural network model. The extracted features from a dataset represented the logic of harmonic functional development of a music pieces by postmodern composers Valentyn Sylvestrov and Philipp Glass, and further dataset generation with the use of structural harmony may include other homophonic tonal music styles. Predictions made by serialised versions of Convolutional LSTM may be used in algorithmic composition for harmonic sequences generation.

## 6. REFERENCES

- Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K. and Darrell, T. (2015) Long-term recurrent convolutional networks for visual recognition and description. *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, Boston, 7-12 June 2015, pp.2625–2634.
- Malandrino D., Pirozzi D. and Zaccagnino R. (2018) Visualization and Music Harmony: Design, Implementation, and Evaluation. *22nd International Conference Information Visualisation*. Fisciano, 10–13 July 2018, pp.498–503. IEEE.
- Nakamura E. and Kaneko K. (2018) Statistical Evolutionary Laws in Music Styles. arXiv:1809.05832 [physics.soc-ph] 21 February 2019.
- Pistone, P. and Shvets, A. (2014) Investigation of the activity-based teaching method in e-learning musical harmony course. *Proceeding of Electronic Visualisation and the Arts (EVA Florence 2014)*, Florence, 7–8 May 2014, pp.107–112. Firenze University Press, Florence.
- Sainath, T. N., Vinyals, O., Senior, A. and Sak, H. (2015) Convolutional, long short-term memory, fully connected deep neural networks. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, 19–24 April 2015, pp.4580–4584. IEEE.
- Shvets A., Desainte-Catherine M., (2015) Schemographe: Application for a New Representation Technique and Methodology of Analysis in Tonal Harmony. *Lecture Notes in Computer Science*, Volume 9027, pp.212–223.
- Vinyals, O., Toshev, A., Bengio, S. and Erhan, D. (2015) Show and tell: A neural image caption generator. *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, Boston, 7-12 June 2015, pp.3156–3164.
- Xingjian, S. H. I., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K. and Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, pp.802–810.