

# Harmonizing Melodies in Real-Time: the Connectionist Approach

Dan Gang \*  
Institute of Computer Science  
Hebrew university  
Jerusalem 91904, Israel  
dang@cs.huji.ac.il

Daniel Lehmann  
Institute of Computer Science  
Hebrew university  
Jerusalem 91904, Israel  
lehmann@cs.huji.ac.il

Naftali Wagner  
Department of Musicology  
Hebrew university  
Jerusalem, Israel  
nwagner@hum.huji.ac.il

## Abstract

We describe a sequential neural network for harmonizing melodies in real time. It models aspects of human cognition. This neural network succeeds reasonably well, if we take into consideration the constraints imposed by real time processing. The model exploits efficiently the available sequential information. The net contains a sub-net for meter that produces a periodic index of meter, providing the needed metric awareness. The net learns the relations between important notes of the melody and their harmonies and is able to produce harmonies for new melodies in real time, i.e., without knowledge of the future development of the melody.

## 1 Introduction

One promising and challenging task is to build a computational model of the cognitive processes that originate in the listener's mind while listening to music. Listening, among all other musical activities at the cognitive level, is a universal musical activity that all human cultures share. Composers, consciously or unconsciously, send hidden messages to an "ideal listener", according to a "stylistic ideal" [CG95]. By means of tension and relaxation the composer creates, in the listener's mind, expectations that can be realized or violated. It seems that an "artificial listener" is not only important by itself but should be used as an essential part of future experiments in building computational models for performance and composition.

Ray Jackendoff, in [Jac91], suggested a model for a musical parsing of listening, by analogy with evidence from the processing of language. Jackendoff proposed a parallel multiple analysis model that

contains a parser able to develop some concurrent analyses. When the parser encounters a choice point among competing interpretations, processing splits into simultaneous branches, one for each possibility. Yet, no known approach has lead to building a computational model that succeeds in establishing an "artificial listener". This task seems to be very complex and far from being achieved. However, some basic attempts were made to model schematic harmonic expectations of the listener [BT91]. A computational model of measurement of the Degree of the Realized Expectations (DRE) was suggested in [BG96].

## 2 Background

In this paper, we examine the problem of real time harmonization of simple melodies. For this, we build a Human Neural Network model (HNN) that reflects some cognitive processes that a musician may perform to solve this task. We found that exploiting cognitive insights into the way humans perform the task is worthwhile for building an artificial system. We used our personal retrospective cognitive experi-

---

\*Dan Gang is supported by an Eshkol Fellowship of the Israel Ministry of Science

ence to direct the choice of the model.

This work is based on a previous system that was designed for the task of harmonizing melodies (not in real time) [GL95]. There, we describe a net capable of learning simple harmonized melodies and to generalize what it has learned by harmonizing melodies it has never seen. The net is a hierarchical Jordan’s sequential net that includes a sub-net trained by back propagation. The sub-net is trained separately to extract harmonic hints from the melody for each measure. The full harmonization is produced by the sequential net that is able to learn the regularity of the musical style from examples. The harmonic context is used directly in this sequential net while the melodic context is used only indirectly through the sub-net that influences the harmonization. The reduced set of important notes extracted from the melody were propagated from the sub-net to the sequential net only one measure ahead. As a consequence the network sees a very limited scope of future information. We found that, for the task of melody harmonization of simple Western European melodies, the system was able to produce quite impressive results. Building a system capable of harmonizing melodies in real time can be a basis for other real time interactive applications, exploiting this approach. An example of a system that may benefit from the harmonic capabilities is NetNeg [GGRL96]. NetNeg is a hybrid interactive architecture for composing polyphonic music in real time. New rules and heuristics may be formulated to exploit the new harmonic framework that is provided by the system for real time harmonization, in producing the counterpoint parts of the melody.

### 3 The Problem

In this paper we suggest an approach to solve a task that is quite complicated even for a human musician. The task is harmonizing in real time unfamiliar simple Western European melodies. The musician has to follow after a melody line performed in real time, trying to produce simultaneously an appropriate harmonization. For the musician real time harmonization is a “risky” musical task and it is relatively rarely performed.

Musicians are often find themselves playing in a jam session, enjoying the freedom to invent new music in a real time improvisation process. In these situations, the interacting performers share a priori information and agreements, before starting the session. Such information may contain the form and structure of the piece and agreed chord progression as a framework (e.g., Blues progression). In case of Jazz



Figure 1: Two harmonizations are presented: the upper one is the output of our system, the lower one is found in the book. The harmonization resulting from the non real time system was the same as the book’s harmonization.



Figure 2: Three harmonizations are presented: the upper one is the output of our system, the lower one is found in the book. The middle one was obtained from the non real time system.

improvisation a melody line exists and may be used as a reference for the improvisation. The performers interact by sending and receiving verbal and non-verbal messages and by sensitively listening to each other. Each of the performers may change dynamically his/her performance by choosing new notes, chords and rhythms. Another related task which is neither improvisation, nor the harmonization of an unfamiliar melody, arises when the harmonizing is done for a familiar melody in real time. In this situation, the musician can recall from his memory the melody but not the harmony outline. The question here is how does the musician exploit this information when trying to harmonize this known melody.

We chose to experiment, as a first case study, with harmonization of unfamiliar melodies in real time. A human model for real time harmonization of unfamiliar melodies should deal with: sequential and temporal information, memorizing the past and using it for the prediction of the next sequential elements, awareness of the location in the metric hierarchy and structure, harmonic and melodic expectations and reduction processes. All of these are preliminary terms in the listening act. The advantage in dealing with the real time harmonization of an unfamiliar melody, compared to listening, is that we can evaluate the performance of the system. For example, we can compare the quality of the results of our system with the harmonization appearing in the examples of the source books, or judge it according to our taste and experience.

## 4 Description of the Model

The previous model for harmonizing melodies is extended to tackle real time harmonization constraints. The net is capable of learning harmonized melodies, using a version of back propagation algorithm, and generalizing what it has learned by harmonizing in real time new melodies. The net used for real time harmonization differs from the previous one (for batch, i.e., non real time harmonization) in the following: the net described in [GL95] used information (the important notes) gathered from the whole measure's melody to harmonize the measure. In particular, it used information concerning the latest part of the measure to provide harmony for the first part of the measure. This does not fit the real time paradigm. Our new real time net is fed the melody itself, a note at a time, in real time. In this way the first note of a measure will influence its harmony but this harmony will not be influenced by the rest of the measure's melody.

The hierarchical net contains a sub-net for meter, that produces a periodic index of meter, providing the needed metric awareness. This sub-net has proved itself essential in obtaining a better harmonization: by differentiating between identical situations the meter index enables more chord changes. As a case study we chose to use just the melody's notes and periodic beat stream modulo the meter of the first and third beats for each measure. We assume for this phase of experiments that the examples contains "ideal" performance - no pitch and rhythm errors, exact timing and duration.

The 4-layer sequential net learns the sequence of chords, as a function of the melody's notes and the index of meter. The output units represent 14 chords - 7 harmonic degrees of the Major scale, and 7 more for their Dominant 7th chords. The output layer is fed back in the state units of the input layer. The output layer represents the predictions or expectations of the net for the next chord. The state units with the same 14 chords represent the context of the current chord sequence. The net also includes two internal hidden layers. The second hidden layer represents the chromatic scale by twelve units. The layer is partially connected with the output layer, establishing the appropriate pitch to chords relations. These connections are fixed, i.e., do not learn. The input layer contains four pools of units which are connected to the first hidden layer, except for the melody pool which is connected to the second hidden layer. The first pool contains 14 units for the state units. The second pool is the output layer of the sub-net for meter and it contains two units to represent the first and the third beats. The third pool contains twelve units to represent the pitch class of the melody's notes. This pool of units is fully connected to the second hidden layer, but some of the connections are fixed and some are learnable. In this way, we were able to impose external representation on the second hidden layer. The fourth pool is plan units which are used to label different sets of notes' sequences.

## 5 Running the Net

### 5.1 The Learning Phase

For the learning phase the net was given eighteen simple harmonized melodies, and learned to reproduce those examples. For this experiment, we fed the net only with the first and third beats of the melody for each measure. An index of meter for these beats was provided accordingly. We assume that, when predicting the appropriate chord the net has access to the note that appears in the first beat or the third beat

of the melody. No memory of the first beat melody is used when predicting the chord for the third beat. However, the state units in the input layer of the sequential net represent the context of the chord sequence by memorizing something of the whole past of the sequence. The value of the state units at time  $t$  is the sum of its value at time  $t - 1$  multiplied by some decay parameter and the value of the corresponding output unit at time  $t - 1$ .

## 5.2 The generalization phase

The network's generalization capability has been tested by giving it new melodies to harmonize. The resulting harmonization was found to be functionally quite appropriate by trained musicians. However, some problems occur in real time harmonization compared to the book harmonization or to the output of the net for non real time harmonization. The chords resulting from the harmonization of the song *Lane in the Forest* (see Figure 1) are similar to the source and the resulting harmonization from the non real time system, except for measure 6 and 10. In these measures the net has difficulty to choose from G7, C or Dm chords to harmonize the B note.

The chords resulting from the harmonization for the song *Supercalifragilisticexpialidocious* (see Figure 2) reveal the problem in harmonizing melodies in real time. The limited scope of the information about the future which is inherent in the real time harmonization task damages the performance of the system. The fact that the net is able to use only the melody's notes on the first and third beats for each measure, may lead to the wrong choice of a G7 chord for the second half of measure 2 and the lack of G7 on measure 15. In these two cases the information of the notes in the fourth beat, that is not available in real time, might help in choosing the right chords for beat three. This information was available to the non real time system, which produced the book harmonization in measure two and hesitated between G7 and C in measure 15. However, this does not explain why the net chose a C chord for the second half of measure 4. The problem might be the lack of accumulation of information from the beginning of the measure. This problem could be, perhaps, cured by memorizing some of the melody context.

## 6 Conclusion

We suggest a sequential neural net as a model for human harmonization of melodies in real time. This neural network uses a simple representation and returns appropriate results, if we take into considera-

tion the real time constraints. The model is exploiting efficiently the available sequential and temporal information for such a task. In the learning phase, the net learns the relations between a reduced set of notes from the melody (the first and third beat) and their harmonies. The net establishes the harmonic context and learns schematic expectations that are influenced by the melody. Using its long-term memory that reflects the representative examples and its short-term memory that establishes the current context, the net is able to predict the next sequential element. The periodic index of meter provides the required local metric awareness to the net. In our model no global structure information, e.g. about the AA'BA' structure, is available and this is obviously a severe limitation. It seems that, the listening act and the harmonization of melodies in real time share a lot of common characteristics. A successful computational human model may shed some light on some aspects of the listening process. The model enables us to hypothesize how all of these perceptual and cognitive components function together to achieve a cognitive musical task (i.e., harmonizing melodies in real time).

## References

- [BG96] J. Berger and D. Gang. Modeling musical expectations: A neural network model of dynamic changes of expectation in the audition of functional tonal music. In *Proceedings of the Fourth International Conference on Music Perception and Cognition*, Montreal, Canada, 1996.
- [BT91] J. J. Bharucha and P. M. Todd. Modeling the perception of tonal structure with neural nets. In P. M. Todd and D. G. Loy, editors, *Music and Connectionism*. M.I.T, 1991.
- [CG95] D. Cohen and R. Granot. Constant and variable influences on stages of musical activities. *Music Perception*, 24:197–229, 1995.
- [GGRL96] C. Goldman, D. Gang, J. Rosenschein, and D. Lehmann. A hybrid interactive architecture for composing polyphonic music in real time. In *Proceedings of the International Computer Music Association*, 1996.
- [GL95] D. Gang and D. Lehmann. An artificial neural net for harmonizing melodies. In *Proceedings of the International Computer Music Association*, 1995.
- [Jac91] R. Jackendoff. Musical parsing and musical affect. *Music Perception*, 9(2):199–229, 1991.