Composing Music with Neural Networks and Probabilistic Finite-State Machines

Tomasz Oliwa, Markus Wagner

Artificial Intelligence Center, The University of Georgia, United States of America
Artificial Intelligence Workgroup, Universität Koblenz-Landau, Germany

Introduction

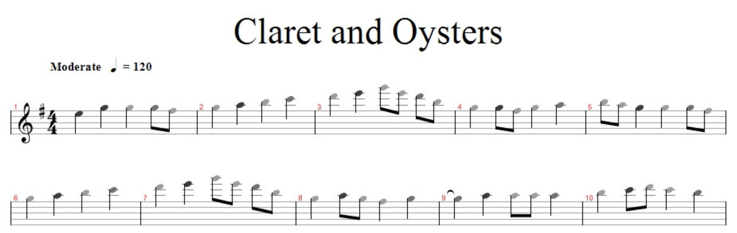
Two novel music composition systems are presented, based on

- Time-Delay Neural Networks (TDNN)
- Probabilistic Output Labelled Transition System (POLTS), updated

The systems acquire musical knowledge by inductive learning and learn **key features** of a musical database. They are able to produce complete musical scores, inheriting these **key features**, for multiple instruments.

Music Representation

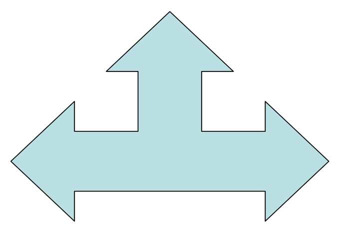
Songs in the MIDI format or a musical score have a corresponding encoding in the **abc** language. This encoding can further be transformed into an internal integer representation, allowing the same **rich expressiveness** and coevally intuitive **numerical computability**.



Three encodings of the same song are shown here as an example of the transformation possibilities.

The abc language can capture musical elements like accidentals, broken rhythms, chords, gracings, rests, slurs, ties etc.

Musical score / MIDI format



abc notation

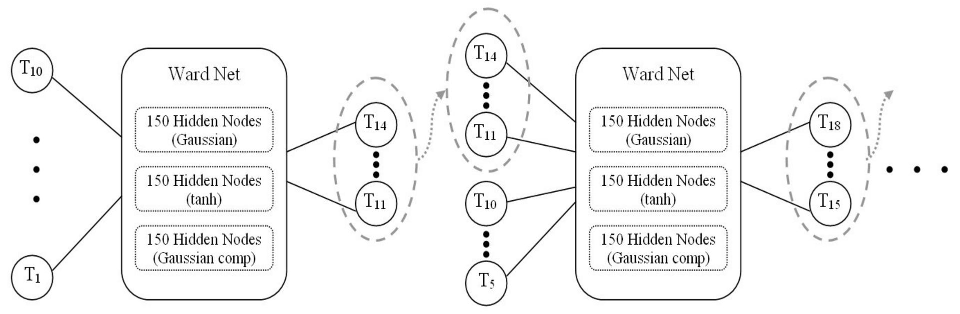
```
E2 G2 G2 G F
G2 A2 B2 c2
d2 e2 g e d B
G2 G F G2 A2
```

Internal representation

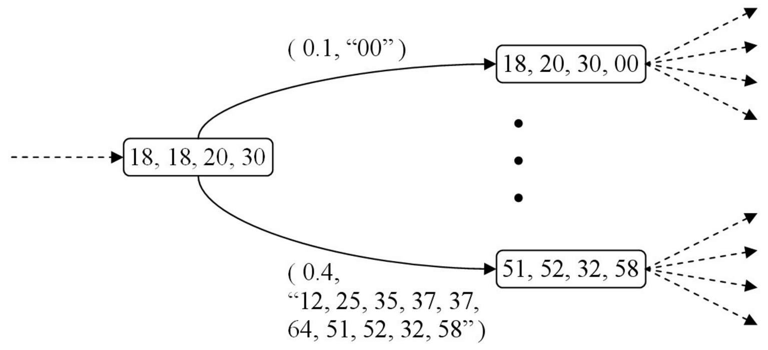
```
10, 18, 18, 20, 16,
18, 22, 26, 30,
34, 38, 48, 40, 36, 28,
18, 20, 16, 18, 22
```

Methods

- Time-Delay Neural Network (TDNN)
  input patterns are successively delayed in time



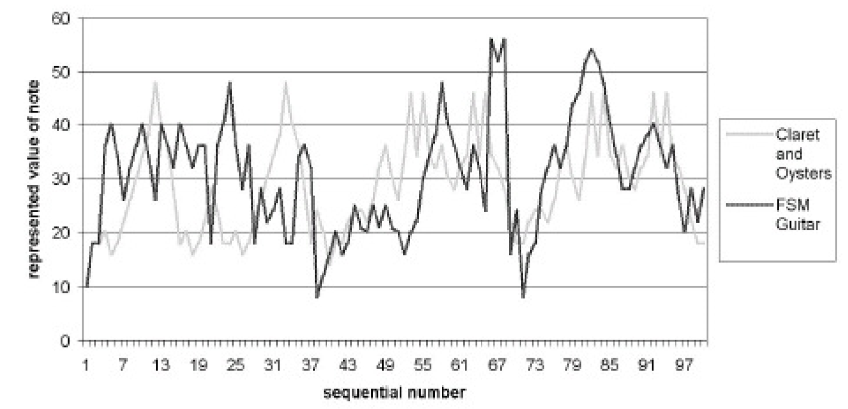
- Probabilistic Output Labelled Transition System (POLTS)
  learned conditional probabilities



Results

The songs composed by POLTS inherited key ideas of the musical database throughout **whole songs**.

Sample song generated by the POLTS system in its musical score.



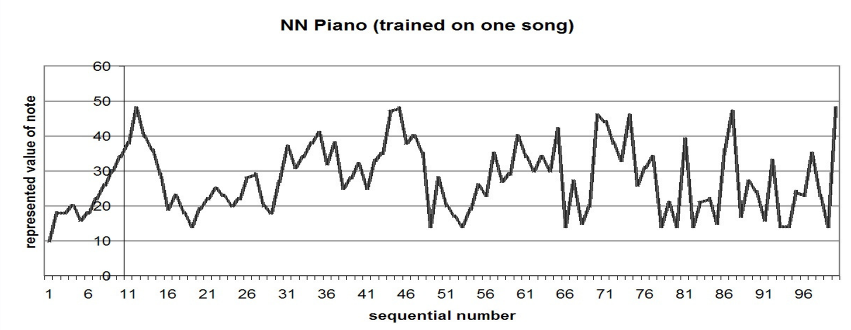
The comparison of the song structures shows:

Repeating patterns exist in the graph. The "landscape" of the song shares similar "hills", for example the notes 50-57 and 80-87.

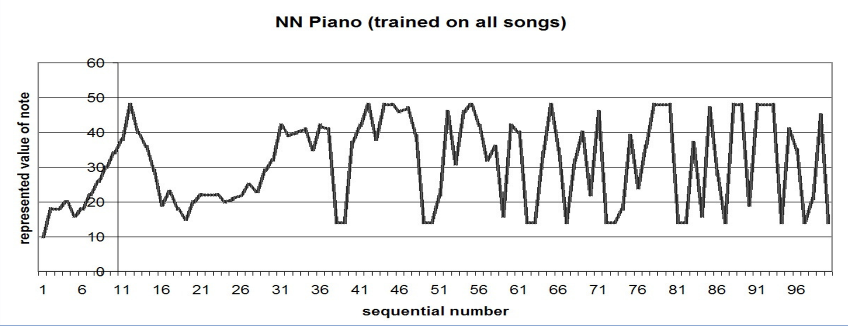
The TDNN composed songs that inherited key ideas of the musical database. They start to focus/oscillate more often between extreme notes in an advanced progress of the song.

Sample song, the TDNN was trained on the whole database.

Sample song, the TDNN was trained on only one song.



In the second half of both the songs, the structure becomes non-coherent as opposed to the song created by POLTS.



If trained on all songs, the structure starts to become non-coherent faster.

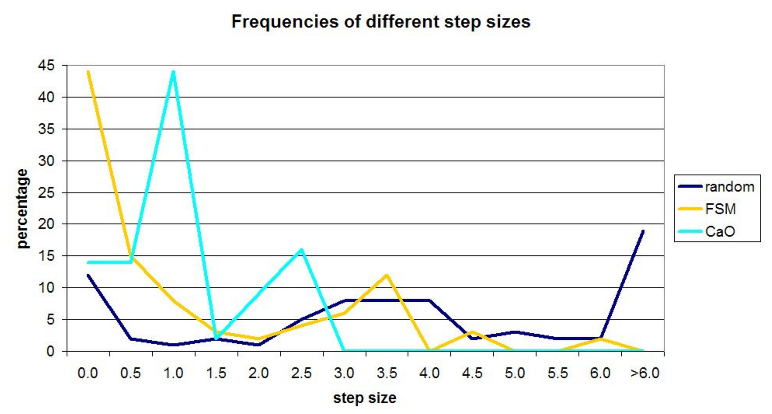
Comparison

Musical Intervals - Consonances and Dissonances

Definition: "Musical interval" = difference in pitch between two notes

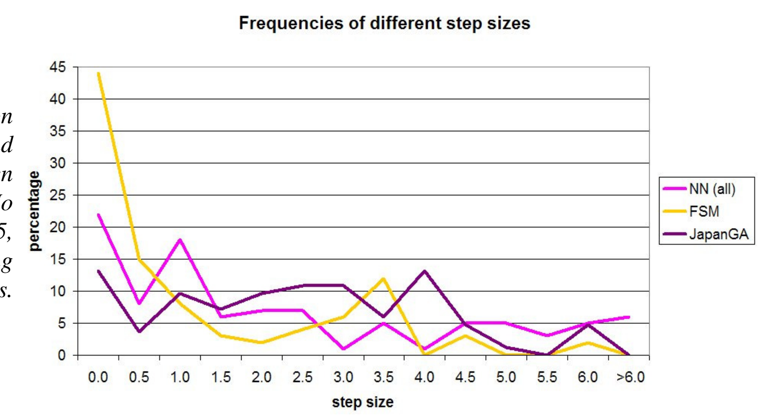
Characterizations into consonances and dissonances of intervals have constantly changed over centuries and therefore are not a very useful criterion.

Nevertheless, we can give detailed statistics of the musical intervals in songs.



The randomly generated song did not have any preference for any interval. Although higher intervals were used by the FSM occasionally, preferences for medium and low size intervals were observed. "Claret and Oysters" has a preference for lower intervals as well.

Although a qualitative comparison is hardly possible, at least a trend towards not arbitrary chosen intervals can be noticed. No approach favors intervals beyond 5, but below that mark each song reflects its own bias.



Conclusion

In general, critics tend to describe their creations as "…compositions only their mother could love…"

In our case

- the local contours make sense, but
- the songs lack thematic structure, having minimal phrase structure and rhythmic organization.

Our advantages

- learned **key elements**, e.g. distinct scale, musical style
- induced bias to produce short coherent sequences at once
- no explicit human modeling needed

References

- Baratè, A., Haus, G., Ludovico, L.A.: Music analysis and modeling through petri nets. In: CMMR 2005. LNCS, vol. 3902, pp. 201–218. Springer, Heidelberg (2006)
- Biles, J.A.: Genjam: A genetic algorithm for generating jazz solos (June 15, 1994)
- Eck, D., Schmidhuber, J.: Finding temporal structure in music: Blues improvisation with lstm recurrent networks (2002)
- Jeon, Y.-W., Lee, I.-K., Yoon, J.-C.: Generating and modifying melody using editable noise function. In: CMMR 2005. LNCS, vol. 3902, pp. 164–168. Springer, Heidelberg (2006)
- Kohonen, T.: A self-learning musical grammar, or associative memory of the second kind. In: IJCNN, Washington DC, vol. I, pp. I–1–I–6, IEEE, Los Alamitos (1989)
- Miranda, E.R., Biles, J.A. (eds.): Evolutionary Computer Music. Springer, Heidelberg (2007)
- Mozer, M.: Neural network music composition by prediction: Exploring the benefits of psychoacoustic constraints and multi-scale processing (1994)
- Schoenberger, J.: Genetic algorithms for musical composition with coherency through genotype. (2002)
- Unehara, M., Onisawa, T.: Construction of music composition system with interactive genetic algorithm (October 2003)
- Walshaw, C.: The abc notation system (1993), http://abcnotation.org.uk