A Neural Probabilistic Model for Predicting Melodic Sequences

Srikanth Cherla
Artur Garcez
Tillman Weyde

CITY UNIVERSITY LONDON

International Workshop on Machine Learning and Music
European Conference on Machine Learning
September 16, 2013

Outline

Introduction: Modelling Melodic Sequences

Background: Music, Information Theory & Neural Networks

Approach: Predicting melodic sequences with RBMs

Initial Results: Encouraging with scope for improvement

Next

Introduction: Modelling Melodic Sequences

Background: Music, Information Theory & Neural Networks

Approach: Predicting melodic sequences with RBMs

Initial Results: Encouraging with scope for improvement

Sequential Information in Music

- A wealth of information in notated music.
- Increasingly availabile
  - In different formats (MIDI, Kern, GP4, etc).
  - For different kinds of music (classical, rock, pop, etc.)
- Analysis of sequences key to extracting information.
- Melody — Good starting point for a broader analysis.

Relevance

Scientific:

- Analysing
  - Compositional practices
  - Musical style & structure
- Music education
- Organizing music data
- Musical expectation

Creative:

- Music generation
- Compositional assistance

Next

Introduction: Modelling Melodic Sequences

Background: Music, Information Theory & Neural Networks

Approach: Predicting melodic sequences with RBMs

Initial Results: Encouraging with scope for improvement

Context

Music & Information Theory

- Multiple-viewpoint Systems (Conklin & Witten, 1995)
- Statistical modelling of melodies (Pearce & Wiggins, 2004)
- Folk Melody classification (Conklin, 2013)

Neural Networks

- Neural Language Models (Bengio et al., 2003)
- RBM-provisor (Bickerman et al., 2010)
- TC-RBM (Spiliopoulou & Storkey, 2011)

Motivation

- Multiple-viewpoint systems — comprehensive & thorough framework for music analysis.
- Recent success of deep neural networks in natural language processing & computer vision.
- Neural networks may be a viable alternative to n-gram models for music analysis within this framework.

Overview of Approach

In this research, the following are explored

- An event-based representation of musical sequences.
- An alternative to Markov models to learn these sequences.
- Two-fold evaluation
    1. Cross-entropy comparison
    2. Folk melody classification

Next

Introduction: Modelling Melodic Sequences

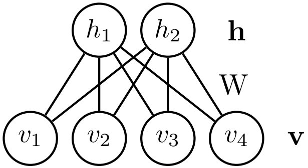
Background: Music, Information Theory & Neural Networks

Approach: Predicting melodic sequences with RBMs

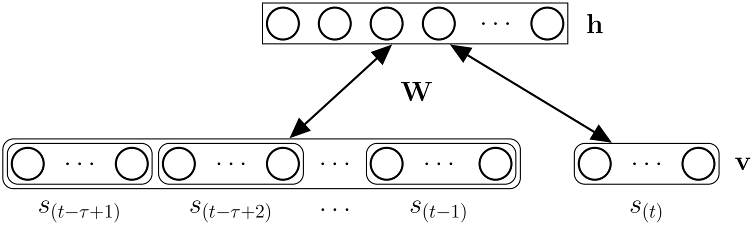
Initial Results: Encouraging with scope for improvement

Restricted Boltzmann Machine

- A bipartite graphical model with binary stochastic units.
- Can be trained to model $p(\mathbf{v})$ using Contrastive Divergence learning algorithm.
- Data in visible layer, features in hidden layer.
- Is readily scalable to deeper network architectures.



Neural Probabilistic Music Prediction



- Consists of softmax visible units.
- Pitch subsequence $s_{(t-\tau+1)\dots t}$ in visible layer.
- RBM trained generatively, tested discriminatively (Larochelle & Bengio, 2008).
- Models the conditional distribution $p(s_t | s_{(t-\tau+1)\dots(t-1)})$
- Absence of event represented with an additional node.

Next

Cross Entropy Comparison

- Dataset: 185 chorale melodies from the Essen Folk Song Collection (EFSC) (Schaffrath & Huron, 1995).
- Same data folds as (Pearce & Wiggins, 2004)
- Training hyperparameters
  - $n_{hid} \in \{\mathbf{100}, 200, 400\}$
  - $\eta \in \{0.01, 0.05\}$
  - $w_{cost} \in \{0.0001, 0.0005\}$
  - $\mu_{ini} = 0.5, \mu_{fin} = 0.9$
- Slightly better cross entropy estimates over a range of subsequence lengths.

n	2	3	4	5	6	7	8	9	∞
n-gram	2.737	2.565	2.505	2.473	2.460	2.457	2.455	2.451	2.446
RBM	2.698	2.530	2.490	2.470	2.454	2.433	2.536	2.486	N/A
	(0.100)	(0.112)	(0.134)	(0.125)	(0.129)	(0.127)	(0.134)	(0.135)	N/A

Folk Melody Classification

- Dataset: A set of folk melody collections of 7 different origins from the EFSC.
- Overall accuracy of 61.74%.

	Nova-Scotia	Alsace	Yugoslavia	Switzerland	Austria	Germany	China	*Recall*	*Total*
Nova-Scotia	117	6	2	2	2	13	10	0.770	152
Alsace	8	33	11	7	15	15	2	0.363	91
Yugoslavia	15	14	54	9	17	7	3	0.454	119
Switzerland	6	9	10	33	22	11	2	0.355	93
Austria	5	16	10	14	41	14	4	0.394	104
Germany	14	23	10	15	14	132	5	0.620	213
China	11	3	2	2	5	1	213	0.899	237
Precision	0.665	0.317	0.545	0.402	0.402	0.684	0.891		

Conclusions & Future Work

We demonstrated the following

- A distributed model for melodic prediction.
- Application of the model to folk melody classification.

Some interesting directions for future work

- Extensions to harmonic sequences.
- Predicting other musical dimensions.
- Learning higher-level structure.
- Improving predictions.
- Interesting applications.

More details of the model available in (Cherla et al., 2013).

Acknowledgements

We would like to thank

Dr. Darrell Conklin (Universidad del País Vasco)
Dr. Marcus Pearce (Queen Mary University London)
Son Tran (City University London)

References I

Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Jauvin.
A Neural Probabilistic Language Model.
Journal of Machine Learning Research, 3:1137–1155, 2003.

Greg Bickerman, Sam Bosley, Peter Swire, and Robert M. Keller.
Learning to Create Jazz Melodies Using Deep Belief Nets.
In *International Conference on Computational Creativity (ICCC)*, 2010.

References II

Srikanth Cherla, Tillman Weyde, Artur Garcez, and Marcus Pearce.
A Distributed Model for Multiple-viewpoint Music Prediction.
In *International Society for Music Information Retrieval Conference*, 2013.

Darrell Conklin.
Multiple Viewpoint Systems for Music Classification.
Journal of New Music Research, 42(1):19–26, 2013.

Darrell Conklin and Ian H Witten.
Multiple viewpoint systems for music prediction.
Journal of New Music Research, 24(1):51–73, 1995.

References III

Hugo Larochelle and Yoshua Bengio.
Classification using discriminative restricted Boltzmann machines.
In *International Conference on Machine Learning (ICML)*, pages 536–543. ACM Press, 2008.

Marcus Pearce and Geraint Wiggins.
Improved methods for statistical modelling of monophonic music.
Journal of New Music Research, 33(4):367–385, 2004.

Helmut Schaffrath and D Huron.
The essen folksong collection in the humdrum kern format.
Menlo Park, CA: Center for Computer Assisted Research in the Humanities, 1995.

References IV

Athina Spiliopoulou and Amos Storkey.
Comparing probabilistic models for melodic sequences.
In *Machine Learning and Knowledge Discovery in Databases*, pages 289–304, 2011.

Thank you!

Questions?