Comparing Probabilistic Models for Melodic Sequences

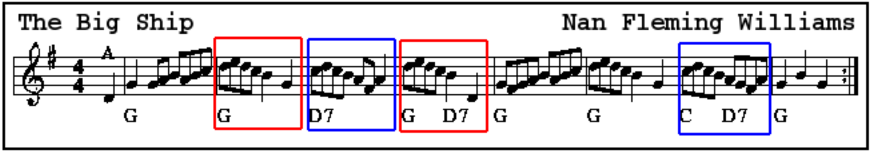
Athina Spiliopoulou, Amos Storkey
a.spiliopoulou@ed.ac.uk, a.storkey@ed.ac.uk

Institute for Adaptive and Neural Computation
School of Informatics
University of Edinburgh

ECML-PKDD 2011
September 7

The Problem

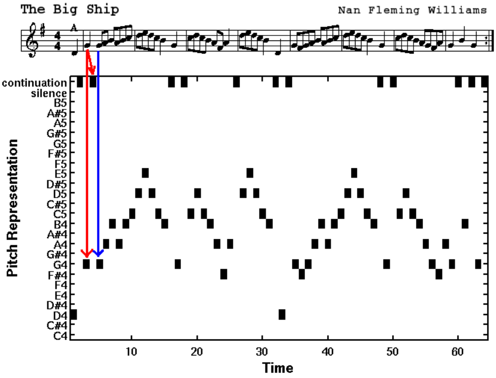
- Learn a generative model for melody directly from a set of musical sequences.
  - Capture the musical structure automatically
  - Avoid utilizing prior musical knowledge



- Challenging structural aspects
  - Repetition
  - Componential influences

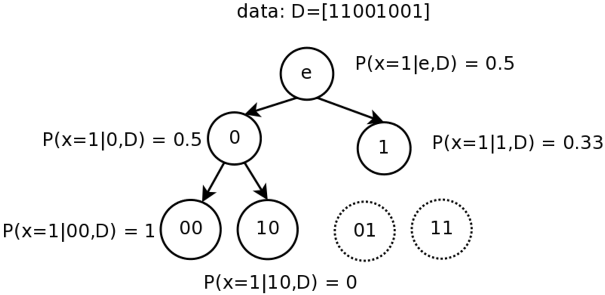
The Data

- 117 Reels from the Nottingham Folk Music Database
  - 80 for training, 37 for testing
  - All pieces in the key of G and in 4/4 meter
  - Information: pitch & duration of the melody notes
    - Discretize to 8th notes
    - Truncate to 2 octaves: C4-B5
    - Use *continuation* value for longer events



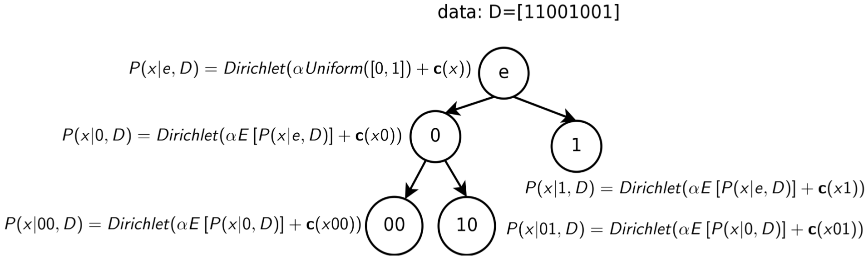
Variable Length Markov Model (VMM)

- State-of-the-art results in automated melody generation (Dubnov et al., 2003; Paiement, 2008)
  - Parses melody into a lexicon of musical motifs
- Length of history to consider is not fixed but depends on the context
- Probabilistic Suffix Tree
  - nodes: labeled by contexts, identified by distributions $P(x_t|context)$
  - depth of node: length of the context
  - learning: add nodes that occur frequently enough in the data



Dirichlet Variable Length Markov Model (Dirichlet-VMM)

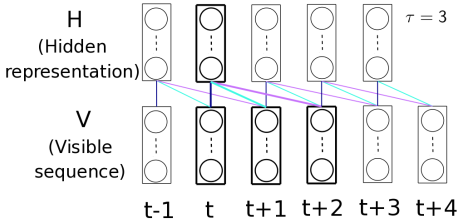
- Smoothing: Bayesian approach
  - Introduce an appropriate prior distribution at each node of the tree
- Hierarchical model: each conditional distribution in the tree is sampled by a Dirichlet distribution centered at the sample multinomial of the parent node.



Time-Convolutional Restricted Boltzmann Machine (TC-RBM)

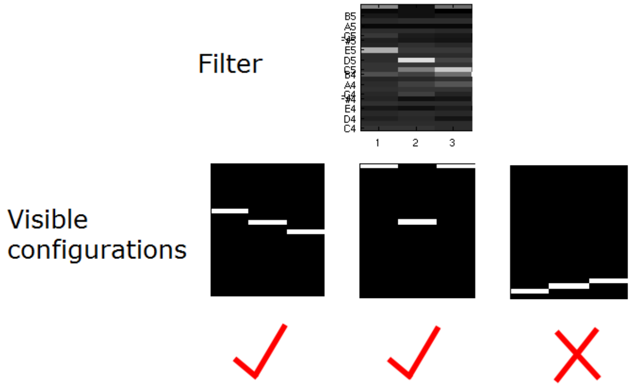
- Distributed representation of the input space
- Latent units receive input from subsequence, not individual time step
- Energy function in bilinear form - efficient learning algorithm

$$E(\mathbf{V}, \mathbf{H}|\theta) = -\sum_{t} \left(\mathbf{c}^{\mathsf{T}} \mathbf{v}_t + \mathbf{b}^{\mathsf{T}} \mathbf{h}_t + \sum_{k=0}^{\tau-1} \mathbf{v}_{t+k}^{\mathsf{T}} \mathbf{W}_k \mathbf{h}_t \right), \quad P(\mathbf{V}, \mathbf{H}|\theta) = \frac{\exp(-E(\mathbf{V}, \mathbf{H}; \boldsymbol{\theta}))}{Z(\boldsymbol{\theta})}$$



Time-Convolutional Restricted Boltzmann Machine (TC-RBM)

- Distributed representation of the input space
- Latent units receive input from subsequence, not individual time step
- Energy function in bilinear form - efficient learning algorithm



Model Evaluation

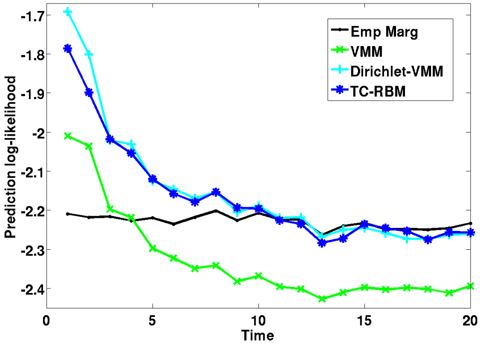
- What is a good generative model for music?
  - How do we evaluate the "musicality" of model samples?
  - Interesting direction: how do music teachers assess students' compositions?
- Proxies for quantitative evaluation of the models
  - Prediction log-likelihood
  - Kullback-Leibler divergence between the statistics of model samples and test sequences
- Qualitative evaluation
  - Examine the latent features of the TC-RBM
  - Listen to model generations

Experiments: Prediction Task

- How well can the models predict the future in test pieces?

Prediction log-likelihood for the τ-th future time step

$$\log L_\tau(\boldsymbol{\theta}, M; \mathbf{D}_n) = \frac{1}{T_n}\sum_{t=1}^{T_n} \log P(d_{t+\tau}|d_1, \ldots, d_t, \boldsymbol{\theta}, M)$$



Experiments: Kullback-Leibler divergence between model and data statistics

- How well do model statistics match the data statistics?

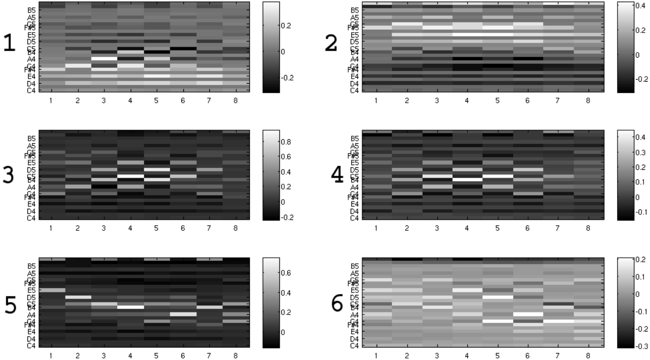
P: empirical frequencies in the data $\quad$ Q: empirical frequencies in model samples
d_t: observation at time t

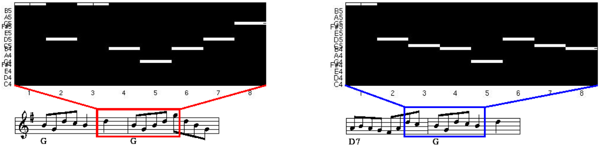
Order 1: $D_{\mathrm{KL}}(P(d_t) \| Q(d_t)) = \frac{1}{N} \sum_{n=1}^{N} P(d_t^n) \log \frac{P(d_t^n)}{Q(d_t^n)}$

Order 2: $D_{\mathrm{KL}}(P(d_t, d_{t+1}) \| Q(d_t, d_{t+1}))$

	order 1	order 2	order 3	order 4
TC-RBM	0.064	**0.273**	**0.872**	**2.420**
Dirichlet-VMM	**0.045**	0.302	1.158	2.594
VMM	0.187	0.481	1.331	3.242

Experiments: Learning Musical Features with the TC-RBM





Conclusions and Current Research

- Problem
  - Melody generation
- Models
  - Dirichlet Variable Length Markov Model
  - Time Convolutional Restricted Boltzmann Machine
- Results
  - Models outperform current state-of-the-art in automated melody generation
  - Hidden units in TC-RBM learn interesting music features

- Models that deal with inter- and intra-piece variability
- Introduce hidden structure in the Dirichlet-VMM

Thank You!

Comparing Probabilistic Models for Melodic Sequences
Athina Spiliopoulou and Amos Storkey, ECML-PKDD 2011