First species counterpoint generation with VNS and vertical viewpoints

D. Herremans\textsuperscript{1}, K. Sörensen\textsuperscript{1}, D. Conklin\textsuperscript{2}

\textsuperscript{1}University of Antwerp
\textsuperscript{2}University of the Basque Country & IKERBASQUE

DMRN+8, London
Overview

Generating counterpoint music

Counterpoint

Variable Neighborhood Search
  How it works
  Experiments & Results

Machine Learning
  Markov Model (Vertical Viewpoints)
  VNS vs Random Walk & Gibbs Sampling
  Results

Conclusion
Acknowledgements

- *Lrn2Cre8*: project funded by the European Commission Framework Programme 7

- Consortium of 6 partners from 5 countries

- Understand the relation between machine learning and creativity, applied to music
  - Generate new musical structures based on learned models
  - New sampling method VNS applied to a controlled music generation problem
Computer aided composition (CAC)

Composing music = combinatorial optimization problem

- Music → combination of notes
- “Good” music → fits a style as well as possible
- Formalized and quantified “rules” of a style → objective function
5th species counterpoint

- Counterpoint & Cantus firmus

- Polyphonic baroque music
- Inspired Bach, Haydn,…
- One of the most formally defined musical styles
  → Rules written by Fux in 1725
Quantifying musical quality using rules

Examples of rules:

▶ Each large leap should be followed by stepwise motion in the opposite direction
▶ Half notes should always be consonant on the first beat, unless they are suspended and continued stepwise and downward
▶ All perfect intervals should be approached by contrary or oblique motion

→ 19 vertical and 19 horizontal subscores between 0 and 1
Quantifying musical quality using rules

\[ f_{cf}(s) = \sum_{i=0}^{19} a_i \cdot \text{subscore}_{cf_i}^H(s) \]  \hspace{1cm} (1)

Horizontal aspect

\[ f_{cp}(s) = \sum_{i=0}^{19} a_i \cdot \text{subscore}_{cp_i}^H(s) + \sum_{j=0}^{19} b_j \cdot \text{subscore}_j^V(s) \]  \hspace{1cm} (2)

Horizontal aspect \hspace{1cm} Vertical aspect

\[ f(s) = f_{cf}(s) + f_{cp}(s) \]  \hspace{1cm} (3)
Variable Neighborhood Search

- Local search with 3 neighborhoods
- Selection
  - Steepest descent
  - Based on adaptive score $f^\alpha(s)$

<table>
<thead>
<tr>
<th>$N_i$</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{sw}$</td>
<td>Swap</td>
<td>Swap two notes</td>
</tr>
<tr>
<td>$N_{c1}$</td>
<td>Change1</td>
<td>Change one note</td>
</tr>
<tr>
<td>$N_{c2}$</td>
<td>Change2</td>
<td>Change two notes</td>
</tr>
</tbody>
</table>
Variable Neighboorhood Search

- Excluded fragments
  - Tabu list
  - Infeasible
- Perturbation
  - Change r% of the notes randomly
- Adaptive weights mechanism
- Update best solution \( s_{\text{best}} \), based on original score \( f(s_{\text{best}}) \)
Results

Example of a generated fragment with score 0.556776.
Machine learning

- Specifying complex objective function by hand
- Automatically generate objective function
  → Learn from a corpus
- To evaluate this:
  → First species: optimal solution known

How does VNS perform compared to Random Walk and Gibbs Sampling?
1st species counterpoint

→ Represented as a sequence of dyads

\[
\begin{bmatrix}
60 \\ 48 \\
65 \\ 50 \\
64 \\ 52 \\
62 \\ 55 \\
60 \\ 57 \\
64 \\ 55 \\
\end{bmatrix} \ldots
\]
Vertical viewpoints method

- Horizontal & vertical aspects → linked

- 3 linked features per dyad:
  - Two pitch class intervals between the two melodic lines
  - Pitch class interval within the dyad
  - $\tau(b|a) = [5, 2, 3]$

- Dyad sequence transformed in abstract feature sequences (Sufficiently abstract to gather statistics in a corpus)

  → First order Markov model of abstract features
Deriving dyad TM from a viewpoint model

Let \( v = \tau(b|a) \) be the feature assigned by a viewpoint \( \tau \) to dyad \( b \), in the context of preceding dyad \( a \)

\[
P(b|a) = P(b, v|a) \quad \text{since } v \text{ is determined by } b \text{ and } a
\]

\[
= \frac{P(b, v, a)}{P(a)}
\]

\[
= \frac{P(v) \times P(a|v) \times P(b|a, v)}{P(a)} \quad \text{chain rule}
\]

\[
= \frac{P(v) \times P(a, v)}{P(v)} \times \frac{P(b|a, v)}{P(a)}
\]

\[
= P(b|a, v) \times \frac{P(a, v)}{P(a)}
\]

\[
= P(b|a, v) \times P(a) \times P(v) \times C_{ab}/P(a) \quad \text{ass. indep. of } a \text{ and } v
\]

\[
= P(b|a, v) \times P(v) \times C_{ab}
\]
Probability of a sequence with respect to the model:

\[ P(s) = \prod_{i=2}^{l} P(e_i|e_{i-1}) \]

Cross-entropy (to be minimised):

\[ f(s) = -\frac{1}{l} \sum_{i=2}^{l} \log(P(e_i|e_{i-1})) \]

For all dyads \( e_1, \ldots e_l \).
Experimental Setup

- 1000 pieces generated by VNS with rules → training
- Fragment with 64 dyads
- Fixed cantus firmus → $11^{61}$ total combinations
- First dyad fixed to $\begin{bmatrix} 60 \\ 48 \end{bmatrix}$
- Last two dyads fixed to $\begin{bmatrix} 59 \\ 50 \end{bmatrix}$ and $\begin{bmatrix} 60 \\ 48 \end{bmatrix}$
Experimental setup

- 3 Methods
  - VNS
  - Random Walk
  - Gibbs Sampling
- Complexity: number of TM lookups
- 10 runs for each method
- Stop criteria: optimum found or $30 \times 10^6$ TM lookups
- VNS TM lookups = $4 \times$ number of moves (overestimated)
Random Walk

- Start with initial fixed dyad.
- Repeat for 1 to $l$:
  - Select next dyad $e_i$ with probability $p(e_i|e_{i-1})$
  - If no next dyad with non-zero probability: dead end
- Several iterations
- On each iteration: solution stored if it is the best so far
Gibbs Sampling

- Repeat:
  - Select a non-fixed dyad
  - Consider all possible permitted dyads at that position
  - Compute the score of each modified piece
  - Construct probability distribution over these scores
  - Select a new piece based on this distribution

- Several iterations

- On each iteration: solution stored if it is the best so far
VNS vs Random Walk & Gibbs Sampling

- VNS: $f(s)^{opt}$ found after an average of $15.8 \times 10^6$ TM lookups
- GS & RW: optimum not found in any of the iterations
VNS vs Random Walk & Gibbs Sampling

Number of TM lookups

Objective function

RW
GS
VNS
Optimum

Number of TM lookups
The proposed VNS is a valid and flexible sampling method that outperforms both Random Walk and Gibbs Sampling using an objective function created by machine learning.

Future research:

- Multiple viewpoints
- More complex music, e.g. fifth species counterpoint using contrapuntal patterns approach of Conklin & Bergeron (2010).
- Learning on “real” data
First species counterpoint generation with VNS and vertical viewpoints

D. Herremans¹, K. Sörensen¹, D. Conklin²

¹University of Antwerp
²University of the Basque Country & IKERBASQUE

DMRN+8, London