University of Antwerp
Operations Research Group
ANT/OR

First species counterpoint generation with VNS and vertical viewpoints

D. Herremans[1], K. Sörensen[1], D. Conklin[2]
[1]University of Antwerp
[2]University of the Basque Country & IKERBASQUE

ORBEL 28, Mons

Overview

Acknowledgements

- *COMEX project*: Interuniversity Attraction Poles (IAP) Programme initiated by the Belgian Science Policy Office
- *Lrn2Cre8 project*: European Commission Framework Programme 7

- Consortium of 6 partners from 5 countries
- Understand the relation between machine learning and creativity, applied to music
  $\rightarrow$ Generate new musical structures based on learned models

$\Rightarrow$ New sampling method VNS applied to a controlled music generation problem

Computer aided composition (CAC)

Composing music = combinatorial optimization problem

- Music $\rightarrow$ combination of notes
- "Good" music $\rightarrow$ fits a style as well as possible
- Formalized and quantified "rules" of a style $\rightarrow$ 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

$\rightarrow$ 19 vertical and 19 horizontal subscores between 0 and 1

Quantifying musical quality using rules

$$f_{cf}(s) = \underbrace{\sum_{i=0}^{19} a_i.\mathsf{subscore\_cf}_i^H(s)}_{\text{horizontal aspect}} \quad (1)$$

$$f_{cp}(s) = \underbrace{\sum_{i=0}^{19} a_i.\mathsf{subscore\_cp}_i^H(s)}_{\text{horizontal aspect}} + \underbrace{\sum_{j=0}^{19} b_j.\mathsf{subscore}_j^V(s)}_{\text{vertical aspect}} \quad (2)$$

$$f(s) = f_{cf}(s) + f_{cp}(s) \quad (3)$$

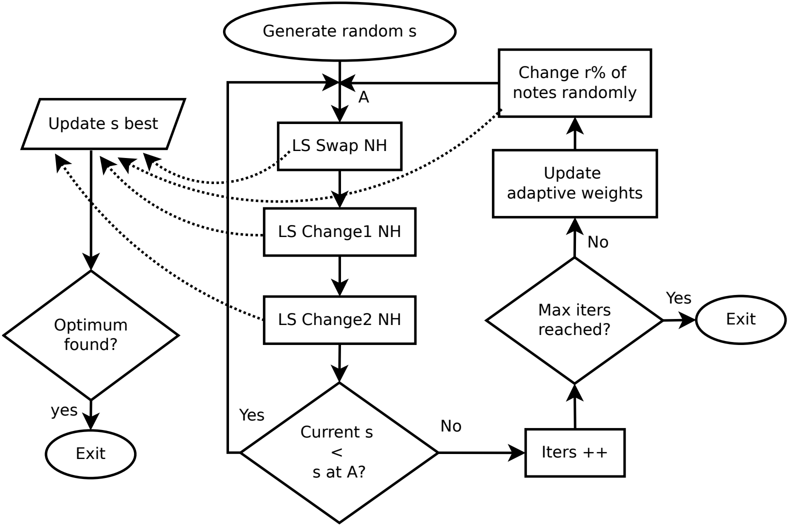
Variable Neigborhood Search

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

N_i	Name	Description
N_{sw}	Swap	Swap two notes
N_{c1}	Change1	Change one note
N_{c2}	Change2	Change two notes

Variable Neigborhood Search

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


Generate random s
A
Change r% of notes randomly
Update s best
LS Swap NH
Update adaptive weights
LS Change1 NH
No
Yes
LS Change2 NH
Max iters reached?
Exit
Optimum found?
yes
Yes
Current s < s at A?
No
Iters ++
Exit

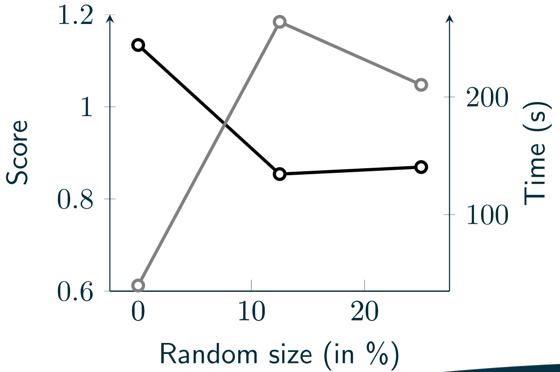
Experiments & Results

- Multi-Way ANOVA model with interaction effects, using R
- $R^2 = 0.98$

Parameter	Df	F value	Prob $(> F)$
N_{c1}	1	9886.2323	$< 2.2e^{-16}$
N_{c2}	1	15690.7234	$< 2.2e^{-16}$
N_{sw}	1	3909.2959	$< 2.2e^{-16}$
randsize	2	1110.1724	$< 2.2e^{-16}$
maxiters	2	322.6488	$< 2.2e^{-16}$
length	1	165.6053	$< 2.2e^{-16}$
adj. weights	1	4.0298	0.0448367
tt_{c1}	2	2.2575	0.1048791
tt_{c2}	2	8.271	0.0002646
tt_{sw}	2	3.2447	0.0391833

Experiments & Results

- Mean plot for the size of the random jump



Optimal parameter settings

Parameter	Value
N_{sw}	on with $tt_{sw} = \frac{1}{16}$
N_{c1}	on with $tt_{c1} = \frac{1}{16}$
N_{c2}	on with $tt_{c2} = \frac{1}{16}$
Random move	$\frac{1}{8}$ changed
Adaptive weights	on
Max. number of iterations	50

Results

- Example of a generated fragment with score 0.556776.

Machine learning

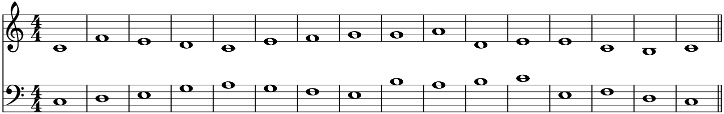
- Specifying complex objective function by hand

  $\downarrow$

- Automatically generate objective function
  $\rightarrow$ Learn from a corpus
- To evaluate this:
  $\rightarrow$ First species: optimal solution known

How does VNS perform compared to Random Walk and Gibbs Sampling?

1st species counterpoint

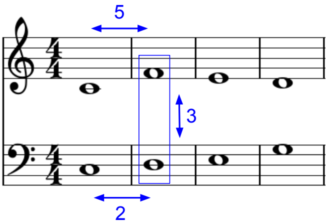


$\rightarrow$ Represented as a sequence of dyads

$$\begin{bmatrix}60\\48\end{bmatrix}\ \begin{bmatrix}65\\50\end{bmatrix}\ \begin{bmatrix}64\\52\end{bmatrix}\ \begin{bmatrix}62\\55\end{bmatrix}\ \begin{bmatrix}60\\57\end{bmatrix}\ \begin{bmatrix}64\\55\end{bmatrix}\dots$$

Vertical viewpoints method

- Horizontal & vertical aspects
  $\rightarrow$ 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)

  $\rightarrow$ Zero-order Markov model over abstract features

Deriving dyad TM from a viewpoint model

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

$$
\begin{aligned}
P(b|a) &= P(b, v|a) && \text{since } v \text{ is determined by } b \text{ and } a \\
&= P(b, v, a)/P(a) \\
&= P(v) \times P(a|v) \times P(b|a, v)/P(a) && \text{chain rule} \\
&= P(v) \times P(a, v)/P(v) \times P(b|a, v)/P(a) \\
&= P(b|a, v) \times P(a, v)/P(a) \\
&= P(b|a, v) \times P(a) \times P(v) \times C_{ab}/P(a) && \text{ass. indep. of } a \text{ and } v \\
&= P(b|a, v) \times P(v) \times C_{ab}
\end{aligned}
$$

Quality of a solution

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 $\rightarrow$ training
- Fragment with 64 dyads
- Fixed cantus firmus
  $\rightarrow 11^{61}$ 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 30×10^6 TM lookups
- VNS TM lookups = 4 * number of moves (overestimated)

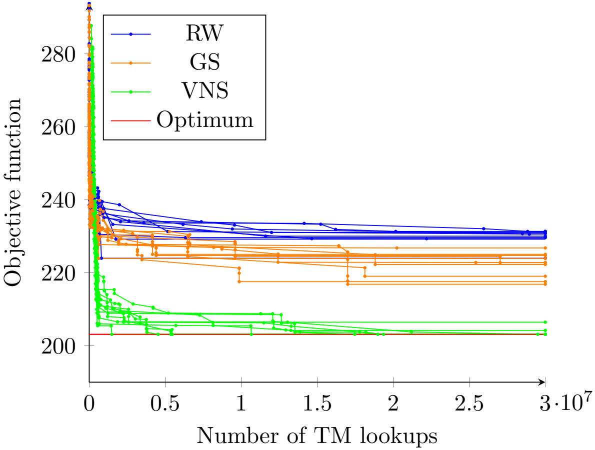
Random Walk

- 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

- Start with initial random piece
- 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

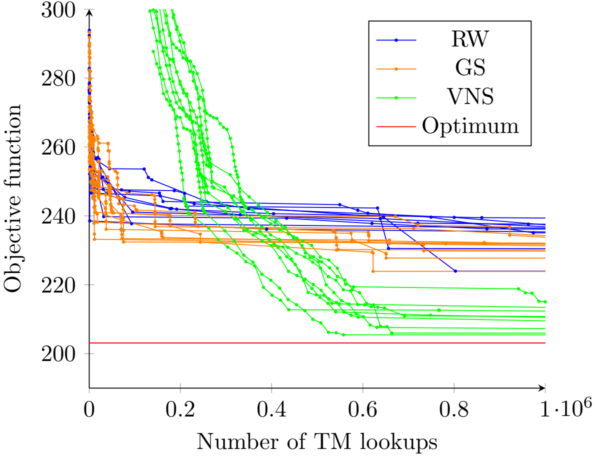
VNS vs Random Walk & Gibbs Sampling



→ VNS: $f(s)^{opt}$ found after an average of 15.8×10^6 TM lookups

→ GS & RW: optimum not found in any of the iterations

VNS vs Random Walk & Gibbs Sampling



Conclusion

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

University of Antwerp
Operations Research Group
ANT/OR

First species counterpoint generation with VNS and vertical viewpoints

D. Herremans[1], K. Sörensen[1], D. Conklin[2]
[1]University of Antwerp
[2]University of the Basque Country & IKERBASQUE

ORBEL 28, Mons