LEARNING OF MUSICAL STRUCTURES IN THE CONTEXT OF IMPROVISATION

DYCI2 Meeting - Ircam
20/09/2017

Ken DÉGUERNEL\textsuperscript{1,2}
Supervisors: Emmanuel VINCENT\textsuperscript{1}, Gérard ASSAYAG\textsuperscript{2}

\textsuperscript{1}Inria, F-54600 Villers-lès-Nancy, France
\textsuperscript{2}IRCAM STMS Lab (CNRS, UPMC, Sorbonne Universités)
Improvising with symbols: the Factor Oracle

- Structure from formal language theory representing the evolution of an improvisation on a local context.
- One dimensional data. Online linear construction in time and space.
- Suffix links (dashed arrow) connect states with similar context.

*Can’t take the multidimensional and multi-level aspects of music into account.*

Contributions

• Use sub-model interpolation to represent multidimensional knowledge.

• Combine multidimensional knowledge of a cultural background and local musical context using probabilistic models and Factor Oracles.

• Create multidimensional improvisations with an interactivity model based on belief propagation and Factor Oracles.

• Extend the concept of scenario to take the multi-level structure into account and use it when generating guided improvisations.

• Test the methods with listening sessions with professional improvisers.
Outline

1. Multidimensional knowledge for improvisation
   1.1. Sub-model interpolation
   1.2. Combining knowledge and context
   1.3. Listening session

2. Interactivity between dimensions/musicians
   2.1. Graph representation of interactions
   2.2. Message passing between dimensions
   2.3. Listening session

3. Multi-level structures for improvisation
   3.1. Phrase structure grammar
   3.2. Multi-level improvisation
   3.3. Experiments

4. Conclusion
Bayesian approach with sub-model interpolation

Predict melody at time $t$, taking all previous musical information into account.

$$P(M_t \mid X_{1:t})$$

melody at time $t$  

set of all musical variables from times 1 to $t$

Approximation of this model with sub-model interpolation.

$$P(M_t \mid X_{1:t}) = \sum_i \lambda_i P_i(M_t \mid Z_{i,t}) \quad \text{with} \quad \sum_i \lambda_i = 1 \quad \text{and} \quad \lambda_i \geq 0 \quad \forall i$$

global model  

sub-models $Z_{i,t} \subset X_{1:t}$

Examples of sub-models: $P(M_t \mid M_{t-1}, \ldots, M_{t-n+1})$, $P(C_t \mid C_{t-1})$, $P(M_t \mid C_t)$
Smoothing techniques

• Training corpora for improvisation are small.

• Smoothing techniques help avoiding zero-valued probabilities and overfitting.
  
  • Additive smoothing: every possible event appears $\delta$ times more than it actually appears in the corpus.
  
  • Back-off smoothing: interpolation with a lower order model.

\[
P(X \mid Y) = \lambda P(X \mid Y) + (1-\lambda) P(X \mid Z) , \text{ with } Z \subset Y
\]
Training the model

\[
\begin{align*}
(B) \quad & P(M_t \mid X_{1:t}) = P(M_t \mid M_{t-1}) \\
(M) \quad & P(M_t \mid X_{1:t}) = P(M_t \mid C_t)
\end{align*}
\]

\[
P(M_t \mid X_{1:t}) = \alpha P(M_t) + \beta U(M_t) + \lambda_1 P(M_t \mid M_{t-1}) + \lambda_2 P(M_t \mid C_t)
\]

Training on Charlie Parker’s Omnibook (50 tunes with improvisation): 
- Training corpus: 40 tunes and improvisations to train the sub-models.
- Validation corpus: 5 tunes and improvisations to train the interpolation and smoothing coefficients.
- Test corpus: 5 tunes and improvisations.

Slight improvement, but lack of local context.

Corpus available in Midi, MusicXml, MuseScore and PDF at: http://repmus.ircam.fr/dyci2/ressources
“The development of a motive should be done in a logical, organic way […] based on intuition enriched with knowledge (from all the study, playing, listening, exposure to various musical styles, etc., that have occurred through a lifetime including all life experiences); the result is a personal musical vocabulary.”

Combining knowledge and context

- **Knowledge** represented by interpolated models trained on a large corpus. Takes multidimensionality into account.

- **Logical intuition** represented by a factor oracle trained on a local frame.
  - Acts as a set of constraints for the probabilistic models,
  - Enables the system to take into consideration a longer context.

Experimenteration and listening sessions

- Improvisation played on a chord chart with probabilistic model trained on melodic and harmonic data.
  \[
P(M_t | X_{1:t}) = P(M_t | M_{t-1}) \quad \quad P(M_t | X_{1:t}) = P(M_t | C_t)
\]

- Probabilistic model trained either on the Omnibook or on a classical music corpus (~1000 tunes).

- No rhythmic information (quarter notes only).

- Factor Oracle constructed on:
  - Anthropology (audio examples),
  - Donna Lee.

- No prior knowledge of the chord charts.
Experimenteration and listening sessions

“It’s way better than that annoying guy in jam session”.

• The melody works with the harmony. Global harmonic progressions and be-bop phrasing are clear.
  “It sounds like it practiced the chord chart, [...], it learnt harmony”.

• Tonal centres are obvious and it plays the full extent of the tonality.

• More logical than without probabilistic models despite some hazy moment.

• When trained on a classical music corpus, the local context is still strong. Charlie Parker’s idiomatic style is present. But it is more careful from a harmony point of view.
  “It is stricter on the harmony so it seems more realistic”.

  Musicology research to create more realistic avatars?

• It plays chords after chords. Lack of a global construction of the improvisation.
Outline

1. Multidimensional knowledge for improvisation
   1.1. Sub-model interpolation
   1.2. Combining knowledge and context
   1.3. Listening session

2. Interactivity between dimensions/musicians
   2.1. Graph representation of interactions
   2.2. Message passing between dimensions
   2.3. Listening session

3. Multi-level structures for improvisation
   3.1. Phrase structure grammar
   3.2. Multi-level improvisation
   3.3. Experiments

4. Conclusion
Cluster graph

\( \Phi_i \): factors
\( \Psi_i \): initial potentials (conditional or joint probabilities)
\( C_i \): clusters
\( S_{i,j} \subseteq C_i \cap C_j \): Sepsets (sets of random variables)

Construction rules:
- \( \forall \Phi_i, \exists C_j; \text{Scope}(\Phi_i) \subseteq C_j \)
- \( \forall (C_i, C_j) \) and variable \( A \), \( \exists \) a unique path between \( C_i \) and \( C_j \) on which every cluster and sepset includes \( A \).

Message passing between dimensions/musicians

• Sub-models are factors for the cluster graph.

• A subset of clusters represents the oracles.

• Oracles communicate via the cluster graph through messages.

Message passing between dimensions/musicians

1. Each oracle provides its attainable states.

2. Initial potential of each cluster is computed accordingly.


4. Estimation of $P(M_n)$ and $P(C_n)$.

5. Cluster Graph provides transition probabilities to the oracles.

Experimentation and listening sessions

“Yeah, yeah, totally!”

• Quite realistic.
  “It sounds like a wacky idea from the CNSM experimental improvisation class”.

• The harmony makes sense and the melody and harmony follow each other well. There is both a horizontal and vertical logic.

• Too much anticipation and not enough reactivity.
  “It is the point where they may know too much”.

• Once again, limitation to a local context. Lack of global construction.
Outline

1. Multidimensional knowledge for improvisation
   1.1. Sub-model interpolation
   1.2. Combining knowledge and context
   1.3. Listening session

2. Interactivity between dimensions/musicians
   2.1. Graph representation of interactions
   2.2. Message passing between dimensions
   2.3. Listening session

3. **Multi-level structures for improvisation**
   3.1. Phrase structure grammar
   3.2. Multi-level improvisation
   3.3. Experiments

4. Conclusion
Phrase Structure Grammar

Grammar based on constituent analysis, i.e. on a breakdown of linguistic functions within a hierarchical structure defining a syntactic structure.

**Example:**

(i)  \( \text{Sentence} \rightarrow \text{NP} + \text{VP} \)
(ii) \( \text{NP} \rightarrow \text{Article} + \text{Noun} \)
(iii) \( \text{VP} \rightarrow \text{Verb} + \text{NP} \)
(iv) \( \text{Article} \rightarrow \text{a, the} \ldots \)
(v) \( \text{Noun} \rightarrow \text{man, ball} \ldots \)
(vi) \( \text{Verb} \rightarrow \text{hit, took} \ldots \)

---

Improvising on a scenario

**Anticipation Step:**
Find an event in the memory sharing a common future with the scenario while ensuring continuity with the past of the memory. *(Prefix indexing)*

**Navigation Step:**
Retrieve the whole sequence or use the regularities in the memory to follow an equivalent non-linear path and thus digress from the original material. *(Factor Oracle structure)*

---

Improvising on a Multi-Level Progression

The scenario and the memory now have multi-level labels.

For both steps, we first look for matches with exact labels and if unavailable, open up the search to *equivalent labels*.

- Ensure consistency upon the global context,
- Favour places in the memory sharing similar multi-level labels,
- Can react to previously unmet chord changes as long as they share a similar role.

To compute a score of similarity between multi-level labels, a weight $W_i$ is attributed to each level such that:

$$\sum_{i \in \text{level}} W_i = 1.$$
Rhythm Changes

32-bar chord progression from George Gershwin’s *I Got Rhythm*.

- **AABA** Structure with a contrasting **B** section.

- **A** section with fast changing chords based on:
  - a series of turnarounds on the tonic,
  - a short modulation to the IVth degree.

- **B** section with dominant seventh chords following the circle of fifths.

- There exist many variations of *Rhythm Changes*.

---

Chris Potter - Anthropology (Charlie Parker), in Diggin’ In - Digging Out (1998)
Grammar for Rhythm Changes

The grammar was constructed based on a corpus with a professional jazz musician.

(i) \( RC \rightarrow A_1 + A_2 + B + A \)

(ii) \( A_1 \rightarrow \tau + \tau + \sigma + \tau \)

(iii) \( A_2 \rightarrow \tau + \tau + \sigma + \omega \)

(iv) \( A \rightarrow A_1, A_2 \)

(v) \( B \rightarrow \delta_{III} + \delta_{VI} + \delta_{II} + \delta_{V} \)

\( \tau, \sigma, \omega, \delta \) are learnt on a corpus.
Experimentation on ‘Rhythm Changes’

Both the phrase structure grammar and memory are trained on the ‘Rhythm Changes’ from Charlie Parker’s Omnibook.

• 40 derivations of ‘Rhythm Changes’ generated from the grammar and validated by professional jazzmen.

• On these we generated improvisations using 2 methods:
  • base generation model without multi-level information.
  • extended generation model with chord (0.3), functional (0.5) and sectional progressions (0.2).
Experimentation on ‘Rhythm Changes’

Listening sessions with two professional jazz musicians:

- Generated improvisations are better at following the structure,
- Add freedom when encountering new chords,
- Less fragmentation creating a better sense of consistency and fluidity,
- Inherent creativity by playing musical sentences on different chords than the original ones whilst keeping consistency.

Other examples available at: members.loria.fr/KDeguernel/smc2017
Hierarchical Tree Generated with Word Sequences Selection

Training on a corpus of scenarios.
Group couples sharing the highest mutual information.
Good results up to 4 bars.
No symbolic relations in the upper levels

\[ J(c_i, c_j) = \log \frac{N(c_i, c_j)T}{N(c_i)N(c_j)} \]

Conclusion

- Method combining multidimensional knowledge of a cultural background and local musical context using probabilistic model and Factor Oracles.
  *Unidimensional improvisations but with multidimensional knowledge*

- Interactivity model with belief propagation on a cluster graph and communication with Factor Oracles.
  *Multidimensional improvisation*

- Improvisation on multi-level scenarios with phrase structure grammar.
  *Unidimensional improvisation respecting the global form*

- Good feedback from musicians during listening sessions.
Plan for the end of thesis

• Resume work on learning of hierarchical structure
  • Merge symbols in upper levels sharing similar neighbourhood and contents,
  • Level selection,
  • Training on generated Rhythm Changes corpus and Realbook,
  • Generate improvisations on multi-level scenarios automatically analysed.

• Write the manuscript!
  • 3 chapters written… (Intro, musicology aspects, state of the art)
  • 4 to go.
  • Submission deadline : 08/01/2018
  • Defense : 06/03/2018
References


