Guider / composer l’improvisation musicale homme-machine avec des scénarios temporels

Jérôme Nika

Ircam
Équipe Représentations Musicales
Take advantage both:
of the prior knowledge of a temporal structure
of the live environment

Use a guiding **scenario** and learned **memory** scheme

Introduce **anticipatory behaviors** in the music generation processes

Combine **planning** and **reactivity**

Concentrate on **pulsed music and idiomatic** improvisation

PhD Thesis (May 2016)
Directors: M. Chemillier, G. Assayag

ImproteK system
Scenario

Memory

knows

guides

learns
researches

Musician’s improvisation

Machine’s improvisation

2. Generation model

Introducing temporal structures in music generation processes
Algorithmic approach (1)

Continuity with the future of the scenario
Find a prefix of the current suffix of the scenario in the memory.
Algorithmic approach (2)

Continuity with the past of the memory
Musical memory learned in a Factor Oracle automaton (Crochemore et al.) —> Suffix links
index \( T \)  
\[ \ldots \ a\ b\ c\ w\ x \ldots \ldots \ldots \ldots \ a\ b\ c\ y\ z \ldots \]

Scenario (\( S \))

Research phase \( \phi_n = S_T \):
- Store the positions and the lengths of the prefixes of \( S_T \) found in \( M \)
  
  \[
  \ldots \ldots \ a\ b\ c\ w\ x \ldots \ldots \ldots \ldots \ a\ b\ c\ .\ .
  \]

\[
\begin{array}{|c|c|c|}
\hline
\text{Memory (} M \text{)} & \text{length} = 5 & 2 & 3 \\
\hline
\end{array}
\]
- Store \( B_T \) (computed during the research)

Later, research phase \( \phi_m = S_{T'} \):
- Using \( B_T \) —> length of the common prefix between \( S_T \) and \( S_{T'} \): \( LCP(m, n) = 3 \)
- Prefixes of \( S_{T'} \) in \( M \): partial results obtained from research phase \( \phi_n = S_T \):

\[
\ldots \ldots \ldots \ a\ b\ c\ w\ x \ldots \ldots \ldots \ldots \ a\ b\ \ldots \ldots \ldots \ldots \ a\ b\ c\ ?
\]

\[
\begin{array}{|c|c|c|}
\hline
\text{Memory (} M \text{)} & \text{length} = 5 & 2 & 3 \\
\hline
\text{LCP}(m, n) = 3 < 5 & 2 & 3 \\
\hline
\end{array}
\]

Results \( \phi_m \):
- Prefix of \( S_{T'} \) in \( M \) length = 3
- Prefix of \( S_{T'} \) in \( M \) length = 2
- Prefix of \( S_{T'} \) in \( M \), length = ?
—> research on a shorter suffix: \( S_{T'} + LCP(m, n) \)

Rent party
Rémi Fox (saxophone), Jérôme Nika (ImproteK),
Repetitions for a performance at Montreux Jazz Festival 2015.

Scenario: chord chart - Memory: live audio
ImproteK music improvisation software, with Rémi Fox

Rent party
Rémi Fox (saxophone), Jérôme Nika (ImproteK),
Repetitions for a performance at Montreux Jazz Festival 2015.

Scenario: simple chord progression
II: Cm7 Bb7 I AbMaj7 Bb7 :II (Rent party, Booker T. Jones)
Memory: live audio segmented and labeled online —> same chord progression
Scenario: Chord progression: *The Man I Love*

Memory:
- Billie Holiday singing *The Man I Love*
- Edith Piaf singing *Mon dieu* and *Milord*
- Elisabeth Schwarzkopf singing *Mi tradì quell’alma ingrata* (Mozart, Don Giovanni), and *Tu che del gel sei cinta* (Puccini, Turandot).
Combining long-term planning and reactivity
Guided improvisation modeled as dynamic calls to an offline model

External event affecting a date \( \geq \) current date

Dynamic score

External events \( \rightarrow \) Reaction \( \rightarrow \) Modification of reactive inputs \( \rightarrow \) Query \( \rightarrow \) Rewriting anticipations

Generic framework:

- External events: operator controlling the system, composed reactivity rules, external listening module…
- Reactive inputs: scenario itself or secondary parameters.

Related work:

4. Playing with (the sound of) the musicians

Recording, sequencing, synchronizing, rendering
Synchronizing the rendering of dynamically evolving sequences generated from live inputs with an external non-metronomic beat source
Synchronized rendering = Tempo estimation (temporal variable Antescofo) + Phase vocoder
Continuity:
The temporal variable $T$ is updated, and the new event to play is **contiguous** to the previous one.

**Hierarchy**
- Reactive process listening to the updates of $T$
- Synchronisation loop process
- Control / rendering process

**Running processes**
- Ideal case: $T$ updated at predicted date
- Case 1: $T$ updated before predicted date
- Case 2: $T$ updated after predicted date

Discontinuity:
The temporal variable $T$ is updated, and the new event to play is **not contiguous** to the previous one.

**Hierarchy**
- Reactive process listening to the updates of $T$
- Synchronisation loop process
- Control / rendering process

**Running processes**
- Case 1: $T$ updated before predicted date
- Case 2: $T$ updated after predicted date

---

**A** launches a child process B

**A** kills B

**Estimated period of variable $T$ at time $T_i$**

**Update the period of the periodic process $L$**

**Render from x to y** (dates in the memory)
Adaptive multimedia recorder / sequencer / renderer
Used as an application case to design the new scheduling / rendering engine of OpenMusic
(Bouche, Bresson, Projet EFFICACe)

5. Scenarios, scenarii...

```
blue
```

"head"

Genericity and "meta-composition"
| Idiomatic alphabet (example) | Content-based alphabet (example) |
1-Define an alphabet

<table>
<thead>
<tr>
<th>Idiomatic alphabet (example)</th>
<th>Content-based alphabet (example)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C Maj7</td>
<td>D m7</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1-Define an alphabet

<table>
<thead>
<tr>
<th>Idiomatic alphabet (example)</th>
<th>Content-based alphabet (example)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C Maj7, D m7, E m7, F 7, ...</td>
<td>L_i, B_j, PM_k</td>
</tr>
</tbody>
</table>

2-Define (possibly different) equivalences to compare:

- events in $M$
  - $X$ m7 $\sim X$ m7, $Y$ m7 $tr$ $X$ m7
  - $L_i \ldots \sim L_i \ldots, L_i \ldots tr L_j \ldots$

- labels in $M$ / in $S$
  - $X$ m7 $\sim X$ m7, $Y$ m7 $tr$ $X$ m7
  - $L_i \ldots \sim L_i \ldots, L_i \ldots tr L_j \ldots$

### 1-Define an alphabet

<table>
<thead>
<tr>
<th>Idiomatic alphabet (example)</th>
<th>Content-based alphabet (example)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C Maj7, D m7, E m7, F 7, ...</td>
<td>L_i, B_j, PM_k</td>
</tr>
</tbody>
</table>

Loudness class i  
Brightness class j  
Playing Mode k

### 2-Define (possibly different) equivalences to compare:

- events in $M$  
  - $X\ m7 \sim X\ m7, Y\ m7 \sim X\ m7$
  - $L_i \ldots \sim L_i \ldots, L_i \ldots \sim L_j \ldots$

- labels in $M$ / in $S$  
  - $X\ m7 \sim X\ m7, Y\ m7 \sim X\ m7$
  - $L_i \ldots \sim L_i \ldots, L_i \ldots \sim L_j \ldots$

### 3-Define associated transformations for the contents

- $Y\ m7 \sim X\ m7 \xrightarrow{tr} tr(Y)$  
  - $tr = \text{Transposition}$

- $L_i \ldots \sim L_j \ldots \xrightarrow{tr} tr(I)$  
  - $tr = \text{Add gain}$

---

1-Define an alphabet

C Maj7, D m7, E m7, F 7, ...

2-Define (possibly different) equivalences to compare:

- events in M
  X m7 ~ X m7, Y m7 ~ tr X m7, L_i ... ~ L_i ...

- labels in M / in S
  X m7 ~ X m7, Y m7 ~ tr X m7, L_i ... ~ L_i ...

3-Define associated transformations for the contents

Y m7 ~ X m7, tr(Y) with tr = Transposition

L_i ... ~ L_j ..., tr L_i ... ~ L_j ...

4-Compose a fixed or dynamic scenario

Dm7, Dm7, G7, G7, C Maj7, C Maj7

Playing mode 1

Playing mode 2

Loudness

Brightness
Scenario example 3:
Composed form using an abstract alphabet

Alphabet = Abstract alphabet A,B,C,...
Scenario example 3: Composed form using an abstract alphabet

Abstract alphabet A, B, C,…

Scenario

2 voices: solo and accompaniment

Constraints for the accompaniment voice: memory restricted to the occurrences of A1 and first measures of B1 to get a repetitive but evolving result
Generative improvisation #1
Rémi Fox (saxophone), Jérôme Nika (ImproteK),
Repetitions for a performance at Montreux Jazz Festival 2015.

Scenario example 3: Composed form using an abstract alphabet
3 different voices

Scenario: Composed form using an abstract alphabet
Memory: Live audio recorded, segmented, and labeled online
6. Method and validation

Collaborations with expert musicians
"Let the music(ians) (pl/s)ay"
Performances with ImproteK since 2013
"The rhythm the rhythm and your memory in my head three years after"

Scenarios: texts by writers from the beat generation

BEAT GENERATION 
En présence de William S. Burroughs - Secret Heroes
Grande Salle Centre Pompidou, June 22.
Musical focuses of the collaborations

Bernard Lubat and "La Compagnie Lubat": "Jazzed-up song", jazz, scat, and free improvisation.
Long-term collaboration to design the first models and prototypes: recombining and phrasing; downstream controls; reduction and multiplication; "hybridization"; rhythm, synchronization, groove.

Jovino Santos Neto: Brazilian music and jazz.
improvisation using an online musical memory; harmonization, arrangement; "hybridization"; rhythmic phrasing.

Kilema, Velonjoro, and Charles Kely: Marovany zither and jazz.
Contrametricity; rhythmic articulation.

Louis Mazetier: Stride piano.
Mixing offline and online memory; scenario defined on an alphabet describing harmony and macrostructure; secondary generation parameters.

Michelle Agnes Magalhaes: Composed contemporary improvisation.
improvisation using an online musical memory; non-idiomatic composed improvisation; content-based alphabet; scenario: discretized multimodal profile of audio descriptors.

Rémi Fox: Funk, jazz, and structured "generative improvisation".
improvisation using an online musical memory; interface and controls: duet between a musician and an operator; rhythm, synchronization, groove; definition of an alphabet, a scenario and constraints.

Hervé Sellin: Jazz and "deconstruction of the idiom".
"Hybridization"; mixing offline and online memory; improvisation plans; "music-driven" vs. "event-driven" interactions; video improvisation.

Benoît Delbecq: Structured improvisation.
improvisation using an online musical memory; interface and controls: duet between a musician and an operator; rhythm, synchronization, groove; definition of an alphabet, a scenario and constraints.
Approach: methodology and validation

- **Intense fieldwork** with Marc Chemillier (EHESS), ethnomusical approach.

- "Participant observation": numerous interactions with musicians fully integrated to the incremental development of the models and architectures.

- **Step-by-step validation** of the successive versions of the models and architectures with musicians:
  - Through work sessions, residences, performances.
  - Gathering assessments from experts of the concerned idioms to validate and/or refine the scientific and technical choices.

- **Filmed listening sessions and interviews** (hours of video): 2 objectives
  - Get some information to understand the musician’s conception of improvisation and motivate the following development steps.
  - Validation of the successive versions of the models.

Outline of a work session with an expert musician:

- Improvisations using the current version of the system (filmed)
- Semi-structured interview (filmed)
- Listening sessions (filmed)
- Semi-structured interview (filmed)

---


Some considerations from listening sessions and interviews

- Validation
  - "Groove", "It does not play mechanically".
  - "A real dialog", "Two musicians playing together".
  - Going from simulation to stimulation.
    - "Here, I played something that is not usual for me".
  - Anticipatory behavior
    - "Inspiring ‘music-driven’ reactions instead of only ‘event-driven’ reactions".
  - Hybridization:
    - "There are some strange areas, the notes are strangers to the harmony, but in a great way."
  - Composed scenarios: (not so) paradoxical interest of the scenario:
    - "Since the system knows the structure I can improvise freely".
  - Depending on the projects and the musicians, validation: "sounds human" / "‘machinic’ virtuosity"
    - Reflects different expectations regarding the system and human-computer improvisation in general.

- Limitations & suggestions
  - Difficulty to develop a musical narration over several occurrences of the scenario.
    - "A real saxophonist would never do that, insisting that much on the first phrase."
    - Interface / Controls ? Reactive listening ? Planning at a higher level ?
  - Hybridization using an offline memory: select the corpus carefully!
    - "Undesired patchwork", "It sits on the fence".
  - Working on structure and sound with the scenario?
  - Current implementation dedicated to pulsed improvisation.
Some conclusions from the listening sessions and interviews
Some conclusions from the listening sessions and interviews
Work in progress…

• Merge the paradigms "free" / "reactive" / "scenario"
  (Omax / Somax / ImproteK)

• Combine scenarios and reactive listening

• Automatic discovery and short-term inference of scenarios

• Anticipatory behaviors / *predicted* short-term structures

• Vertical chaining of agents working on different alphabets

• Enhance the possibilities regarding music composition