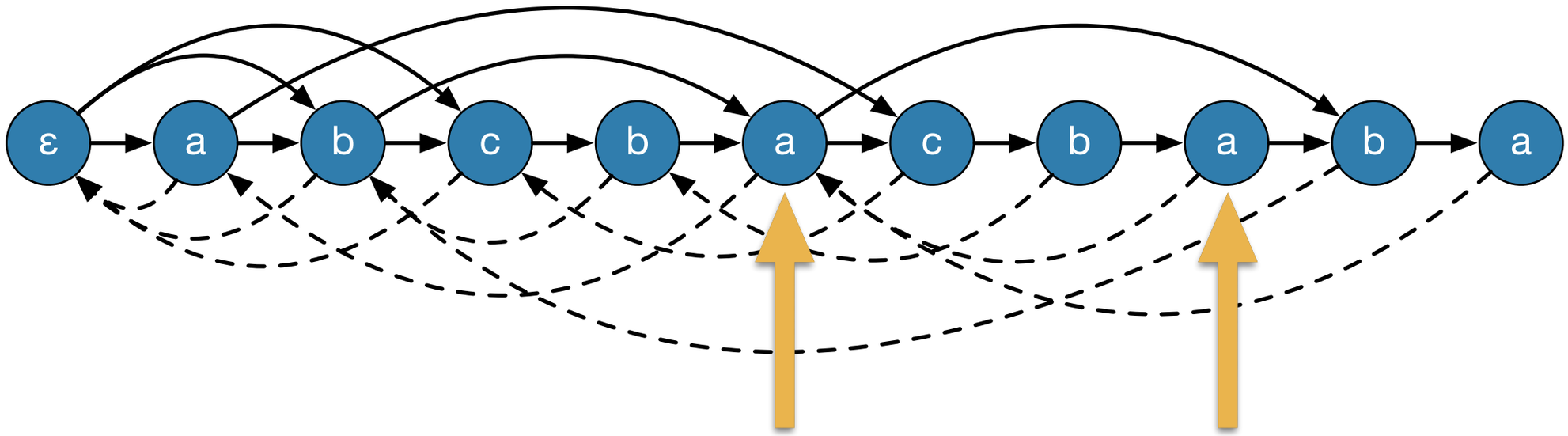
# Learning of Musical Structures in the Context of Improvisation

DYCI2 Meeting - Ircam
20/09/2017

Ken Déguernel[1,2]
Supervisors : Emmanuel Vincent[1], Gérard Assayag[2]

[1]Inria, F-54600 Villers-lès-Nancy, France
[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.**

G. ASSAYAG, G. BLOCH. *Navigating the oracle : a heuristic approach,* in Proceedings of the International Computer Music Conference, 2007, pp.405-412.

# 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.

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

global model

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

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

$C_t$: chord at time t

# Smoothing techniques

- Training corpora for improvisation are small.
- Smoothing techniques help avoiding zero-valued probabilities and overfitting.
  - Additive smoothing : every possible event appears δ times more than it actually appears in the corpus.
  - Back-off smoothing : interpolation with a lower order model.

    $P(X \mid Y) = \lambda P(X \mid Y) + (1-\lambda) P(X \mid Z)$ , with $Z \subset Y$

# Training the model

$$(B) \qquad P(M_t \mid X_{1:t}) = P(M_t \mid M_{t-1})$$
$$(M) \qquad P(M_t \mid X_{1:t}) = P(M_t \mid C_t)$$

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

Training on Charlie Parker's Omnibook (50 tunes with improvisation) :

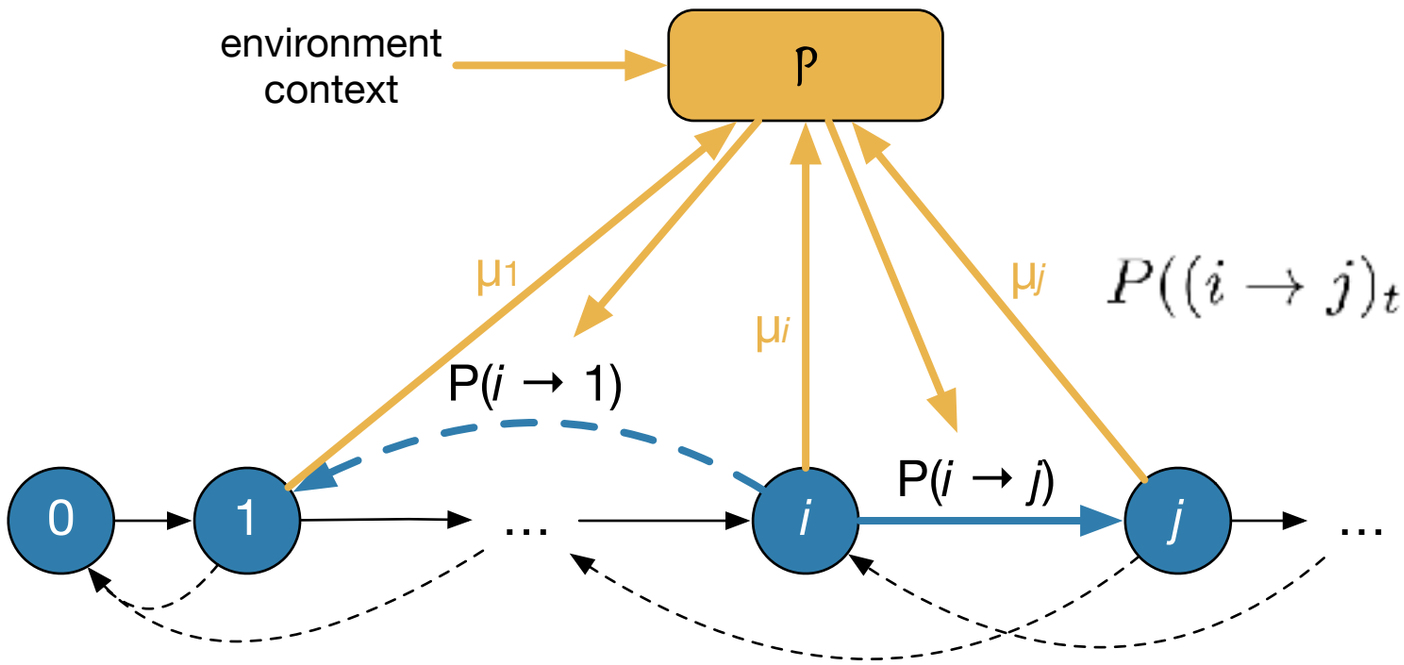
- Training corpus : 40 tunes and improvisations to train the sub-models.
- Validation corpus : 5 tunes and improvisations to train the interpolation and smoothing coefficients.
- Test corpus : 5 tunes and improvisations.

**Slight improvement, but lack of local context.**

---

Corpus available in Midi, MusicXml, MuseScore and PDF at : *http://repmus.ircam.fr/dyci2/ressources*

# Combining knowledge and context



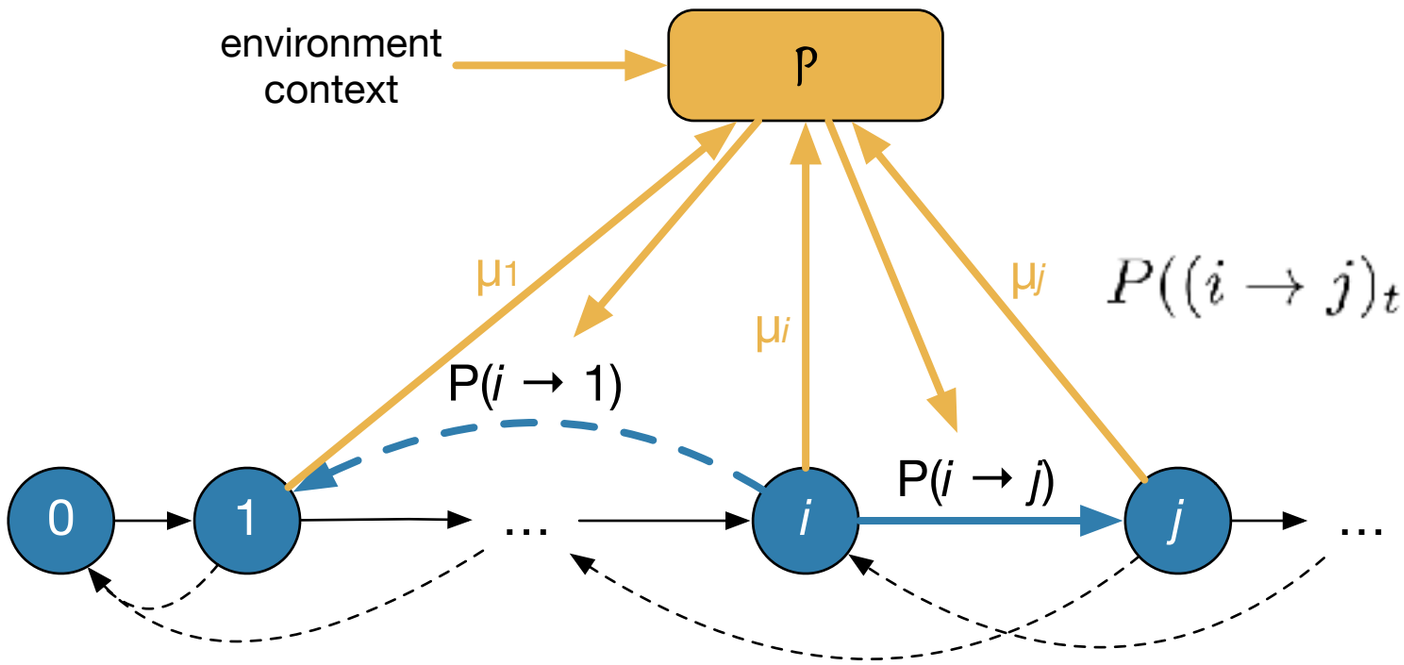

$$P((i \to j)_t | X_{1:t}) = \frac{P(M_t = \mu_j^M | X_{1:t})}{\sum_{k \in \mathrm{Att}(i)} P(M_t = \mu_k^M | X_{1:t})}$$

*"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**."*

---

M. CRISPELL. *Elements of improvisation* in Arcana : Musicians on Music, J. Zorn, ed., Hips Road/Tzadik, 2000, pp.190-192

# Combining knowledge and context




$$P((i \to j)_t | X_{1:t}) = \frac{P(M_t = \mu_j^M | X_{1:t})}{\sum_{k \in \mathrm{Att}(i)} P(M_t = \mu_k^M | X_{1:t})}$$

- **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.

K. DÉGUERNEL, E. VINCENT, G. ASSAYAG. *Using multidimensional sequences for improvisation in the OMax paradigm,* in Proceedings of the 13th Sound and Music Computing Conference, 2016, pp.117-122

# Experimentation and listening sessions

- Improvisation played on a chord chart with probabilistic model trained on melodic and harmonic data.

$$P(M_t \mid X_{1:t}) = P(M_t \mid M_{t-1}) \qquad P(M_t \mid X_{1:t}) = P(M_t \mid C_t)$$

- 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.

# Experimentation 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

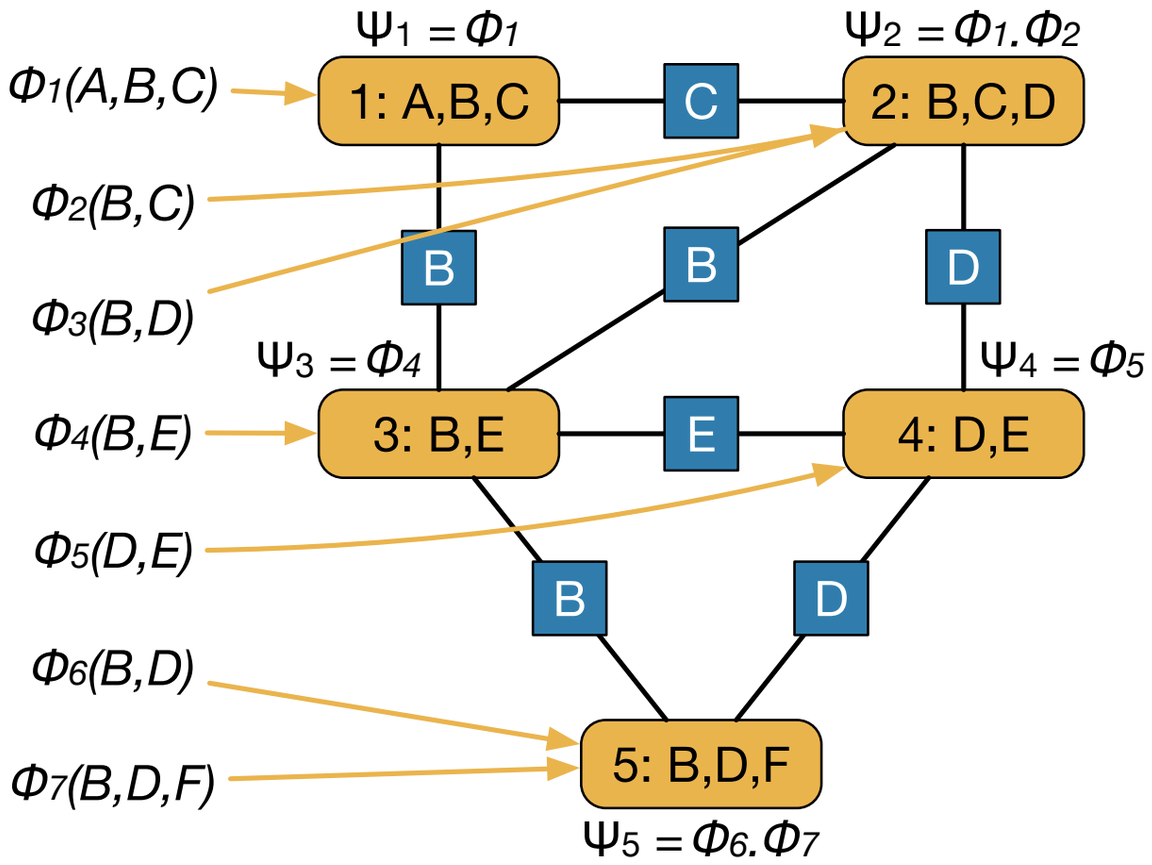
1. Multidimensional knowledge for improvisation
   1.1. Sub-model interpolation
   1.2. Combining knowledge and context
   1.3. Listening session

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)

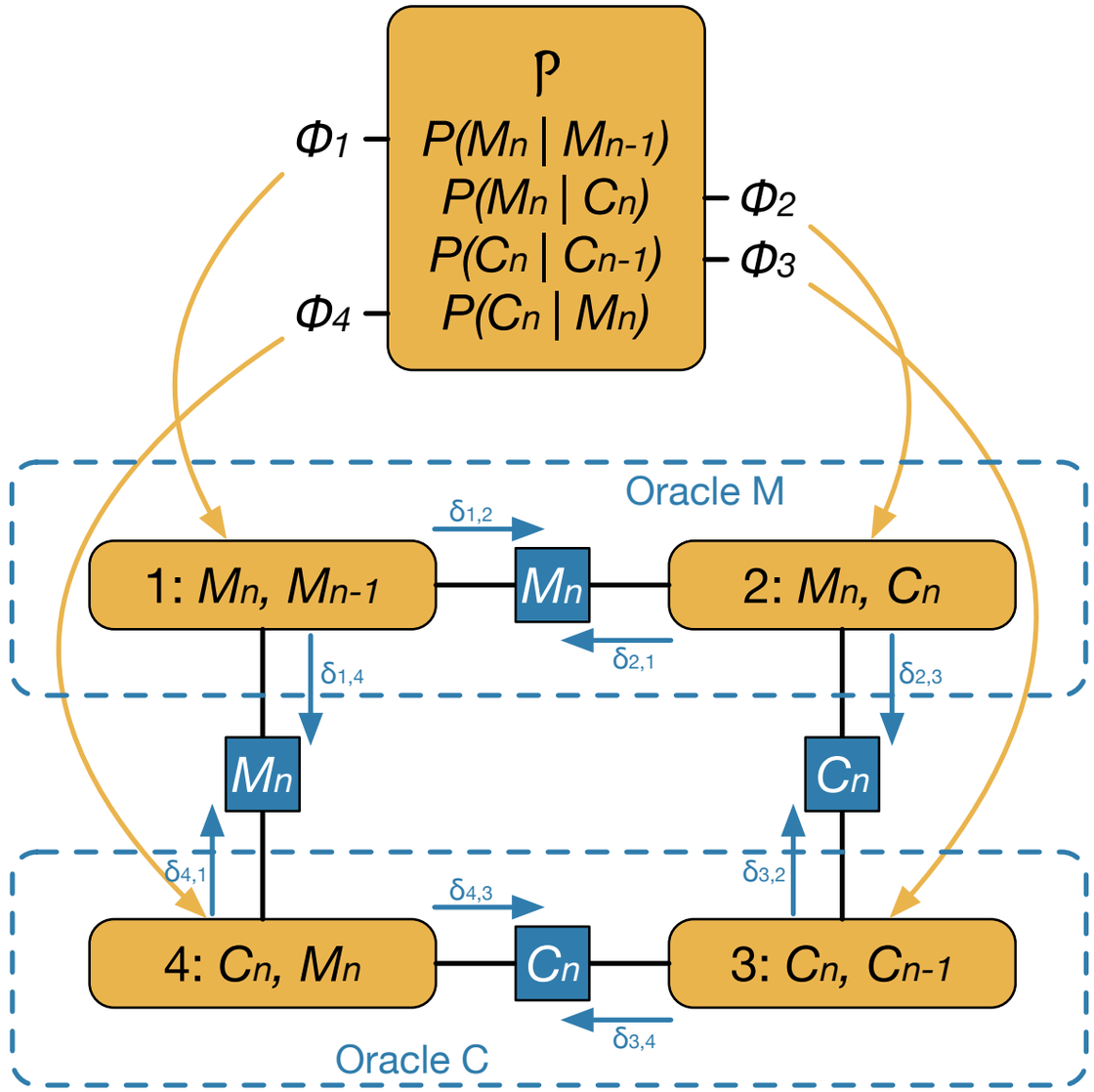
Construction rules :

- $\forall \Phi_i$ , $\exists\ C_j$ ; $Scope(\Phi_i) \subseteq C_j$
- $\forall (C_i, C_j)$ and variable A, $\exists$ *a unique path between* $C_i$ *and* $C_j$ *on which every cluster and sepset includes A.*

D. KOLLER, N. FRIEDMAN, *Probabilistic graphical models : principles and techniques*, MIT Press, 2009.

# 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.




K. DÉGUERNEL, E. VINCENT, G. ASSAYAG, *Probabilistic Factor Oracles for Multidimensional Machine Improvisation*, Submitted to Computer Music Journal.

# Message passing between dimensions/ musicians

1. Each oracle provides its attainable states.

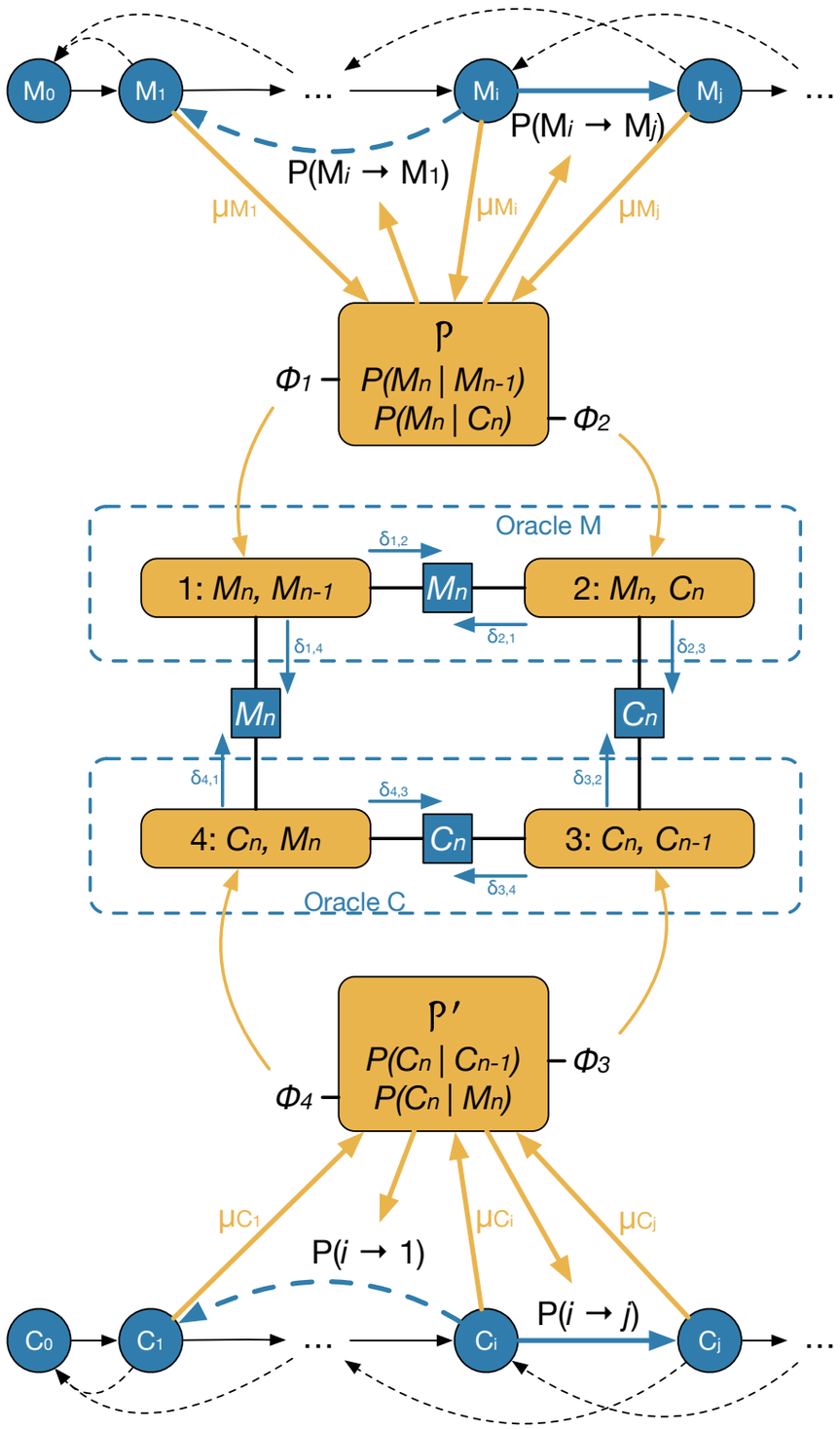
2. Initial potential of each cluster is computed accordingly.

3. Message passing with Belief Propagation algorithm.

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

5. Cluster Graph provides transition probabilities to the oracles.

6. Generation.



# 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) $Sentence \rightarrow NP + VP$

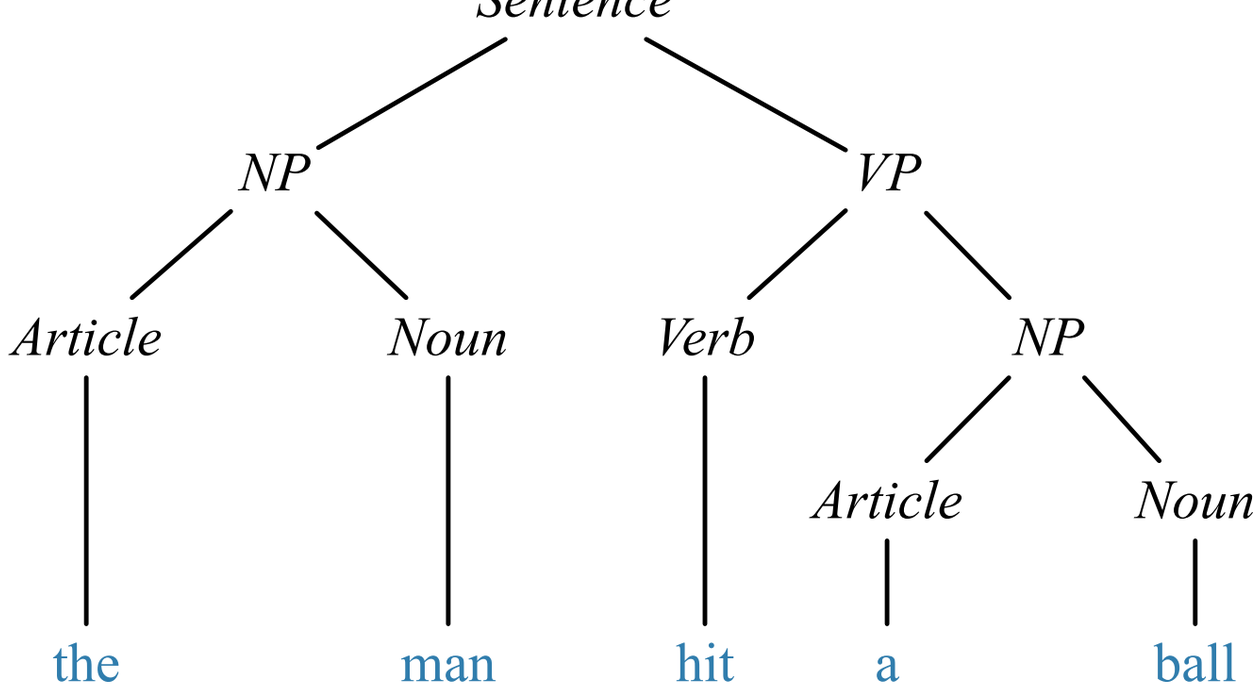
(ii) $NP \rightarrow Article + Noun$

(iii) $VP \rightarrow Verb + NP$

(iv) $Article \rightarrow$ a, the…

(v) $Noun \rightarrow$ man, ball…

(vi) $Verb \rightarrow$ hit, took…




N. Chomsky, *Studies on semantics in generative grammar.* Walter de Gruyter, 1996, vol. 107.

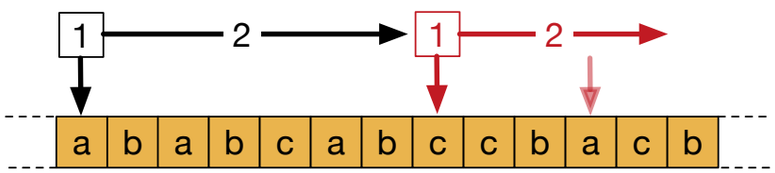
# Improvising on a scenario

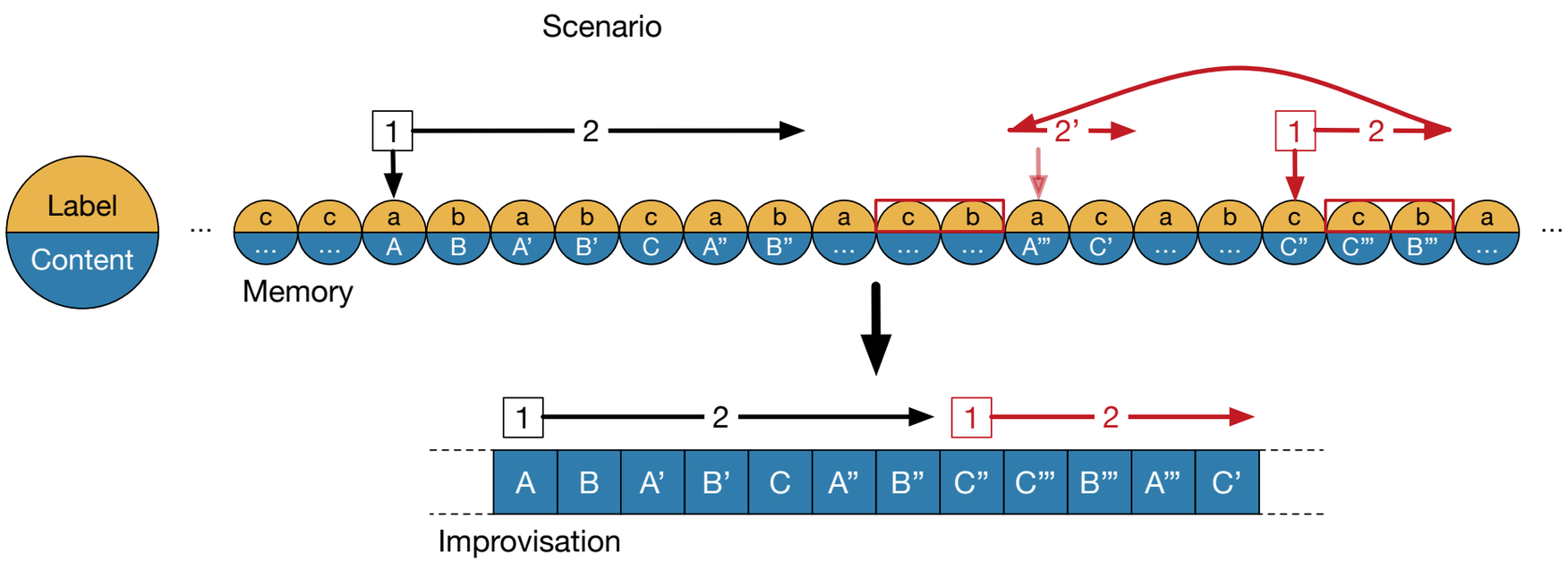
## *Anticipation Step:*

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)*

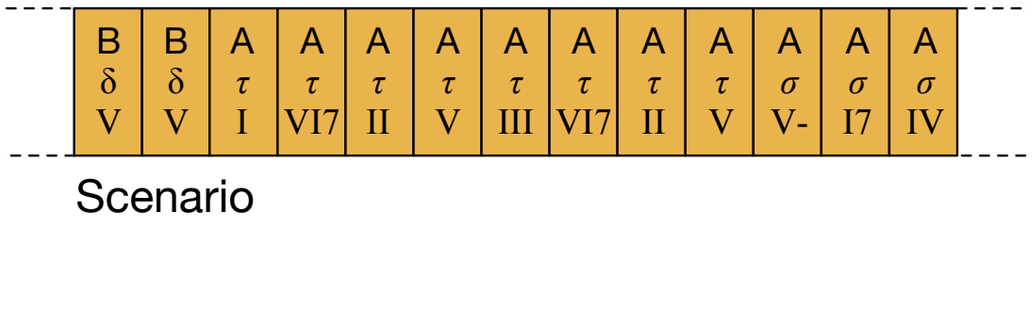


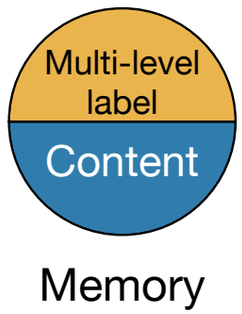




J. NIKA, M. CHEMILLIER, G. ASSAYAG, *ImproteK : introducting scenarios into human-computer music improvisation,* ACM Computers in Entertainment, 4 (2017).

# 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.

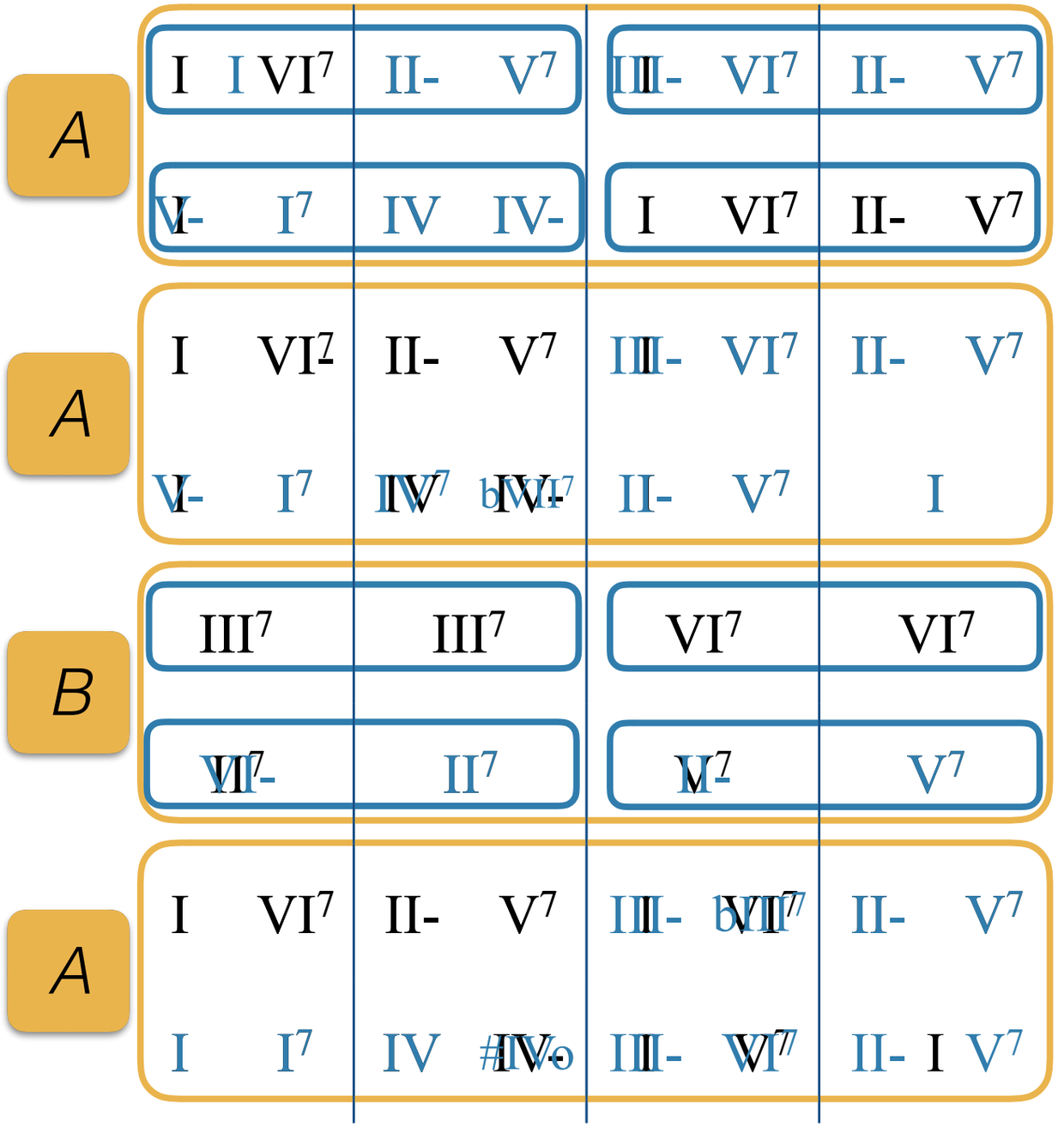
To compute a score of similarity between multi-level labels, a weight *Wi* is attributed to each level such that :

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

# 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 IV$^{th}$ degree.
- *B* section with dominant seventh chords following the circle of fifths.
- There exist many variations of *Rhythm Changes.*




Original *I Got 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)* $RC \rightarrow A_1 + A_2 + B + A$

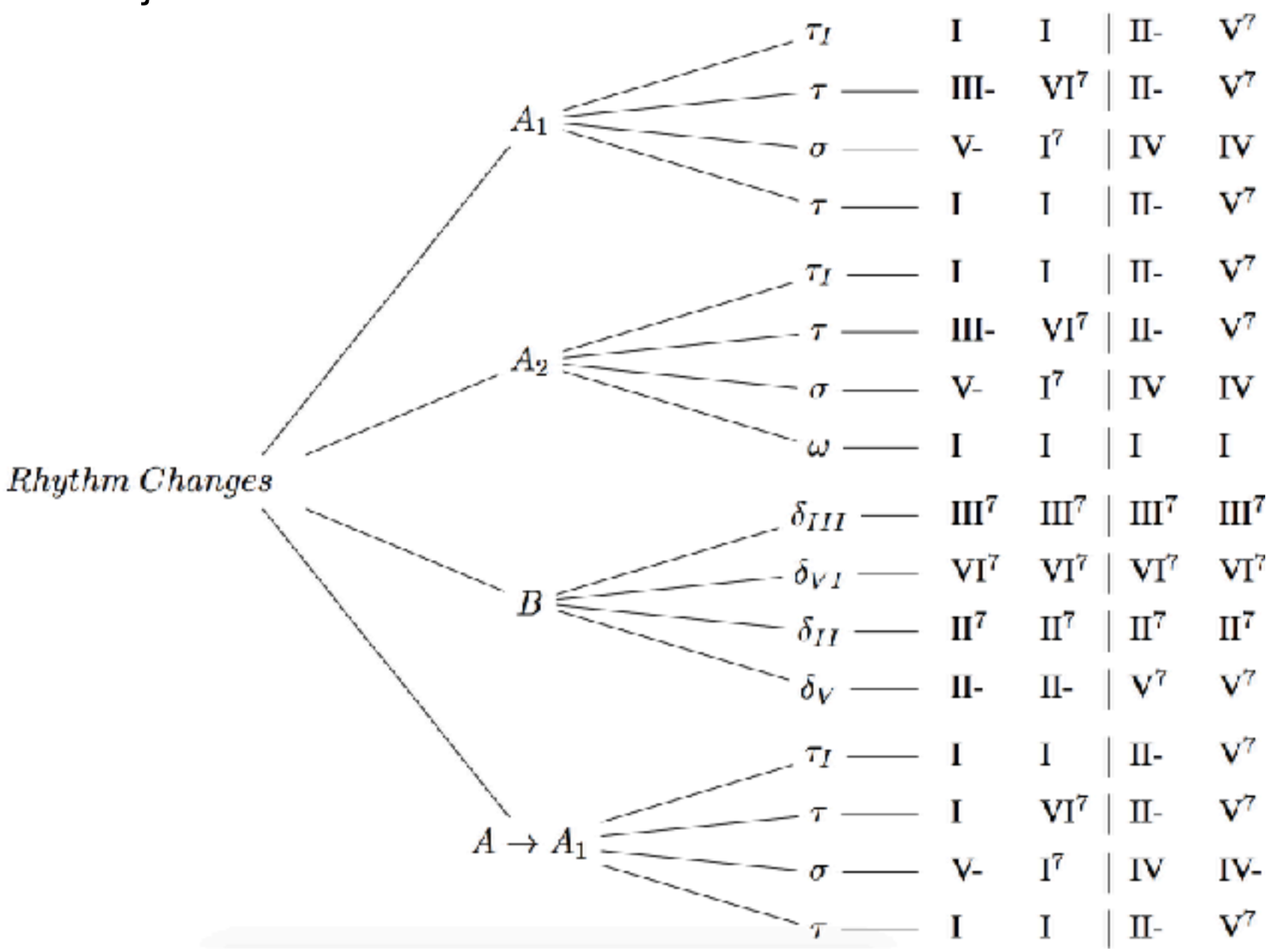
*(ii)* $A_1 \rightarrow \tau_I + \tau + \sigma + \tau$

*(iii)* $A_2 \rightarrow \tau_I + \tau + \sigma + \omega$

*(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’

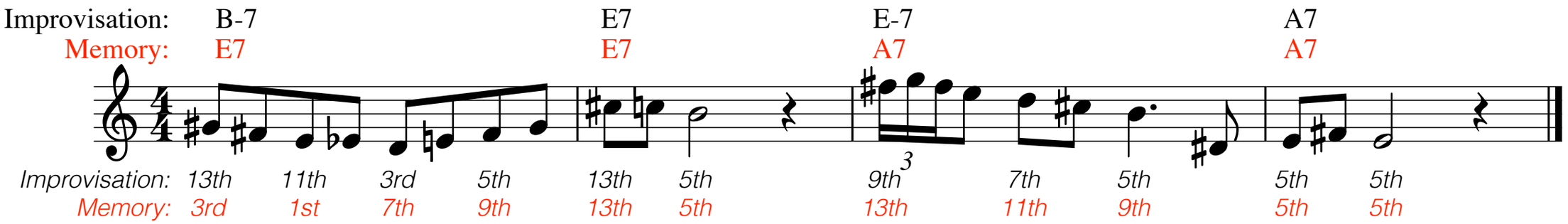
Both the phrase structure grammar and memory are trained on the ‘Rhythm Changes’ from *Charlie Parker’s Omnibook.*

- 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)*.

| | | | |
|---|---|---|---|
| I | II- $V^7$ | I | II- $V^7$ |
| $I^7$ | $IV^7$ | III- $VI^7$ | II- $V^7$ |
| I | II- $V^7$ | III- $VI^7$ | II- $V^7$ |
| V- I7 | IV IV- | I | I |
| $III^7$ | $III^7$ | $VI^7$ | $VI^7$ |
| VI- | $II^7$ | II- | $V^7$ |
| I | II- $V^7$ | I $VI^7$ | II- $V^7$ |
| I $I^7$ | IV IV- | I | II- $V^7$ |

# Experimentation on ‘Rhythm Changes’




Improvisation on *An Oscar for Treadwell*

Listening sessions with two professional jazz musicians :

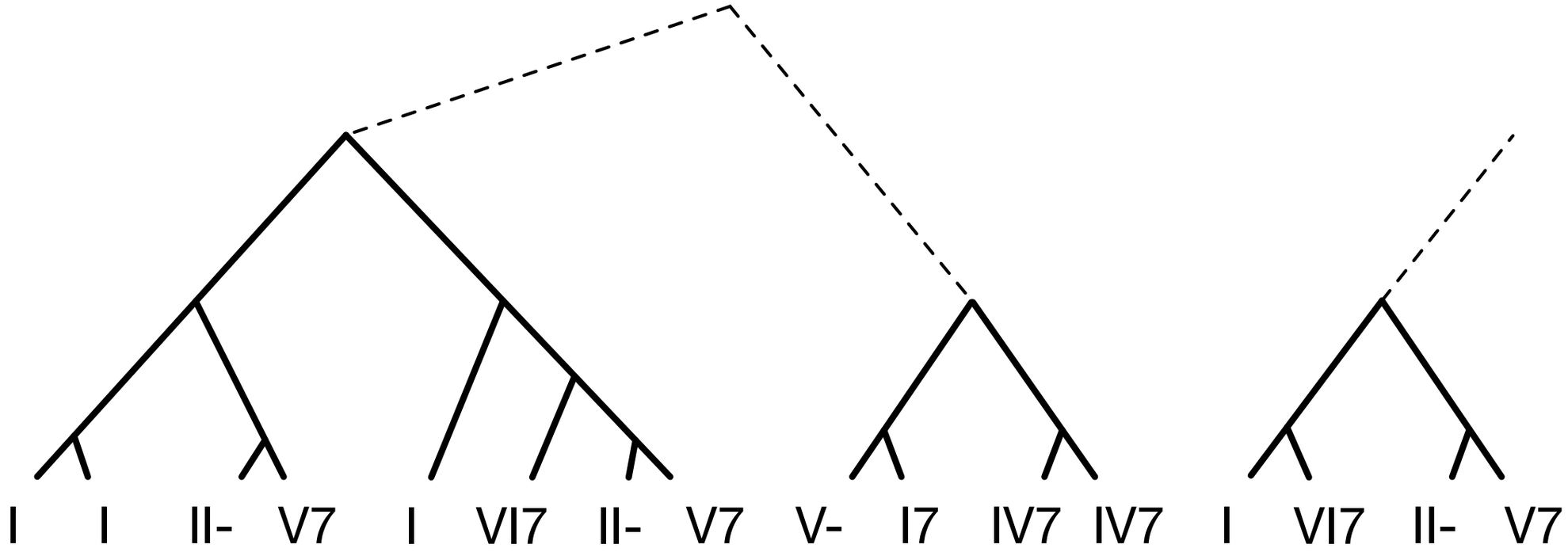
- Generated improvisations are better at following the structure,
- Add freedom when encountering new chords,
- Less fragmentation creating a better sense of consistency and fluidity,
- Inherent creativity by playing musical sentences on different chords than the original ones whilst keeping consistency.

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

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




I. ZITOUNI, K. SMAÏLI, J-P. HATON, *Beyong the conventional statistical language models: the variable-length sequences approach,* Interspeech (2000).

# 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


- [1] G. ASSAYAG, G. BLOCH. *Navigating the oracle : a heuristic approach,* in Proceedings of the International Computer Music Conference, 2007, pp.405-412.
- [2] S. A. RACZYNSKI, S. FUKAYAMA, E. VINCENT, *Melody harmonisation with interpolated probabilistic models,* Journal of New Music Research, 2013.
- [3] M. CRISPELL. *Elements of improvisation* in Arcana : Musicians on Music, J. Zorn, ed., Hips Road/Tzadik, 2000, pp.190-192.
- [4] K. DÉGUERNEL, E. VINCENT, G. ASSAYAG. *Using multidimensional sequences for improvisation in the OMax paradigm,* in Proceedings of the 13th Sound and Music Computing Conference, 2016, pp.117-122.
- [5] D. KOLLER, N. FRIEDMAN, *Probabilistic graphical models : principles and techniques*, MIT Press, 2009.
- [6] J. NIKA, M. CHEMILLIER, G. ASSAYAG, *ImproteK : introducting scenarios into human-computer music improvisation,* ACM Computers in Entertainment, 4 (2017).
- [7] K. DÉGUERNEL, J. NIKA, E. VINCENT, G. ASSAYAG. *Generating equivalent chord progressions to enrich guided improvisation : application to Rhythm Changes,* in Proceedings of the 14th Sound and Music Computing Conference, 2017.
- [8] I. ZITOUNI, K. SMAÏLI, J-P. HATON, *Beyong the conventional statistical language models: the variable-length sequences approach,* Interspeech, 2000.