

Rapport d'activités 2015

Ken Déguernel

Directeurs de thèse : Emmanuel Vincent (Inria) et Gérard Assayag (Ircam)

5 février 2016

1 Contexte

Les systèmes d'improvisation actuels permettent de prendre en compte une information unidimensionnelle (typiquement une mélodie représentée par une succession de hauteurs) fournie en direct par un musicien ou lue dans un corpus afin de générer de nouvelles improvisations par une recombinaison du matériau musical.

Ces improvisations générées peuvent être, par exemple, dirigée par la connaissance antérieure d'un scénario à suivre [8] ou par une écoute active permettant au système de réagir en direct à son environnement [4].

Cependant, ces systèmes n'ont pas la possibilité de prendre en considération les corrélations existantes entre plusieurs dimensions musicales (hauteurs, harmonies, durées, intensités, timbres...).

2 Modèles probabilistes du langage

Dans [11], Raczynski et al. exploitent des méthodes issues de la modélisation probabiliste du langage pour effectuer des harmonisations de mélodies pour un corpus de musique classique. Dans [6], nous avons adapté ces méthodes pour une tâche de génération de mélodie capable de prendre en considération plusieurs dimensions musicales.

L'objectif est de prédire quelle mélodie jouer à un instant t connaissant l'ensemble de l'information musicale ayant précédé. Cette information musicale est multi-dimensionnelle, prenant en compte des variables musicales de natures différentes (hauteurs, harmonies...).

Cependant, une telle prédiction n'est pas possible

à calculer en pratique car le modèle global généré est d'une dimension trop importante. Pour palier ce problème, on effectue alors une interpolation de différents sous-modèles, plus simples, ne prenant en compte qu'une partie des variables musicales ; par exemple un modèle de n -gramme sur une dimension, ou des modèles d'interactions directes entre dimensions (par exemple, quelle note jouer à l'instant t connaissant l'harmonie à ce même instant?).

L'apprentissage ne pouvant être exhaustif de par la taille limitée des corpus d'apprentissage et les possibilités d'expressivité infinies de l'improvisation, on effectue un lissage des sous-modèles. Le but du lissage est de corriger l'estimation des probabilités estimées et d'éviter notamment d'avoir des probabilités d'événements nulles.

L'interpolation des sous-modèles peut être linéaire ou log-linéaire. Il sera également possible par la suite de considérer une modélisation par réseaux de neurones (permettant une généralisation du principe d'interpolation de sous-modèles).

Nous avons montré qu'utiliser une interpolation de sous-modèles permet au modèle globale d'avoir un meilleur pouvoir de prédiction pour la génération de mélodie. Cependant, cette amélioration semble assez faible. Nous avons donc décidé de combiner ces méthodes avec la structure d'oracle des facteurs.

3 Combinaison avec l'oracle des facteurs

L'oracle des facteurs est une structure issue de la bioinformatique et de la théorie des langages [1, 7]

ayant déjà fait ses preuves pour l'improvisation musicale dans des systèmes comme OMax [3, 2], ImproTek [9] ou PyOracle [12]. Cette structure permet de conserver la linéarité de ce qui a été appris et de créer des liens entre des zones dans la mémoire possédant un contexte commun.

Nous avons alors développé un modèle qui combine l'aspect probabiliste des modèles de langages qui prennent en compte les aspects multidimensionnels avec la mise en contexte de l'oracle des facteurs. Nous nous sommes inspirés d'un des éléments d'improvisation proposés par Marilyn Crispell dans [5] :

Le développement d'un motif doit être fait de manière logique, organique, ordonné (improvisation en tant que composition spontanée), pas cependant d'une manière préconçue, mais plutôt d'une manière basée sur l'intuition enrichie par des connaissances (de tout ce qui a été étudié, joué, écouté, des différents styles musicaux auxquels on a été exposé, etc. au cours de notre vie - y compris toutes les expériences personnelles); le résultat étant un vocabulaire musical personnel.

D'un côté, on crée un module probabiliste comprenant l'ensemble des différents sous-modèles que l'on souhaite prendre en considération ainsi que les valeurs des coefficients d'interpolation et de lissage nécessaire au modèle probabiliste global. Ce module peut être obtenu à partir d'un apprentissage effectué en différé sur un corpus important, mais peut également être appris (ou mis à jour) en temps réel. Dans la citation de Marilyn Crispell, il correspond à l'ensemble des connaissances acquises au cours de la vie de notre système.

D'un autre côté, on construit un oracle des facteurs dont la construction des états, des transitions et des liens suffixes ne dépend que d'une dimension choisie (typiquement la mélodie), mais dont les états contiennent l'ensemble de l'information musicale de l'instant qu'ils représentent. Dans la citation de Marilyn Crispell, cela correspond à la logique du contexte dans lequel le motif doit être développé. L'oracle est construit en direct avec les informations fournies en direct par des musiciens ou lues dans un corpus (généralement plus restreint que celui du module probabiliste).

À chaque étape de la navigation dans l'oracle, on va alors chercher des informations dans le module probabiliste afin d'orienter notre parcours. À partir d'un état, et connaissant les différents états accessibles (au sens du parcours dans un oracle des facteurs défini dans [2]) et l'information musicale qu'ils contiennent, on calcule pour chaque possibilité un score correspondant à l'interpolation des différents sous-modèles du module probabiliste. On peut alors prendre une décision informée du chemin à prendre, ou alors normaliser les scores des différentes possibilités pour obtenir des probabilités de transition et faire un choix aléatoire.

Une idée en cours de développement est de créer différents oracles correspondant à différentes musicales que l'on souhaite représenter. Le parcours de ces oracles est alors dirigé à la fois par le module probabiliste qu'ils peuvent partager, mais aussi par une communication entre les oracles par passage de messages en employant des algorithmes de *belief propagation*. Ainsi, les différents oracles effectuent un choix globale (de connaissance interne et externe) d'un parcours à effectuer.

Une autre piste à développer est la combinaison de modèles et de dimensions n'évoluant pas dans la même temporalité, les différentes dimensions musicales n'évoluant pas à la même vitesse, et au sein d'une même dimension les vitesses d'évolutions pouvant énormément varier. *Time is a blind guide...*

4 Activités annexes

- Création d'un corpus au format .xml de cinquante thèmes et improvisations (mélodies et harmonies) issues de l'Omnibook [10] de Charlie Parker.
- Démonstration d'OMax et présentation des axes de recherches lors de l'événement grand public *Science & You* à Nancy.
- Participation à la 12th *Sound and Music Computing Summerschool* à Maynooth, Irlande.
- Prise de contact et visite à l'EPFL pour un partenariat avec l'équipe responsable de la numérisation des archives du Montreux Jazz Festival, et récupération d'un échantillon de données pour le projet.

Références

- [1] Cyril Allauzen, Maxime Crochemore, and Mathieu Raffinot. Factor oracle : A new structure for pattern matching. *SOFSEM'99, Theory and Practice of Informatics*, pages 291–306, 1999.
- [2] Gérard Assayag and Georges Bloch. Navigating the oracle : A heuristic approach. In *Proceedings of the International Computer Music Conference*, pages 405–412, 2007.
- [3] Gérard Assayag and Shlomo Dubnov. Using factor oracles for machine improvisation. *Soft Computing*, 8-9 :604–610, 2004.
- [4] Laurent Bonasse-Gahot. An update on the SoMax project. Technical report, IRCAM, 2014.
- [5] Marilyn Crispell. Elements of improvisation. In John Zorn, editor, *Arcana : Musicians on Music*, pages 190–192. 2000.
- [6] Ken Deguernel. Apprentissage de structures multi-dimensionnelles pour l'improvisation musicale. Master's thesis, Inria Nancy Grand Est, 2015.
- [7] Arnaud Lefebvre and Thierry Lecroq. Computing repeated factors with a factor oracle. In *Proceedings of the 11th Australasian Workshop On Combinatorial Algorithms*, pages 145–158, 2000.
- [8] Jérôme Nika and Marc Chemillier. Imrotek, integrating harmonic controls into improvisation in the filiation of OMax. In *Proceedings of the International Computer Music Conference*, pages 180–187, 2012.
- [9] Jérôme Nika, José Echeveste, Marc Chemillier, and Jean-Louis Giavitto. Planning human-computer improvisation. In *Proceedings of the International Computer Music Conference*, pages 330–338, 2014.
- [10] Charlie Parker and Jamey Aebersold. *Charlie Parker Omnibook*. Alfred Music Publishing, 1978.
- [11] Stanisław A. Raczynski, Satoru Fukayama, and Emmanuel Vincent. Melody harmonisation with interpolated probabilistic models. *Journal of New Music Research*, 42(3) :223–235, 2013.
- [12] Greg Surges and Shlomo Dubnov. Feature selection and composition using pyoracle. In *Proceedings of the 2nd International Workshop on Musical Metacreation*, 2013.