

Proposition de thèse : *Modèles génératifs grammaticaux pour le traitement de la notation musicale*

Proposé par : Philippe Rigaux (philippe.rigaux@lecnam.net), Florent Jacquemard (florent.jacquemard@inria.fr) & Nicolas Audebert (nicolas.audebert@lecnam.net)

Contexte

Les modèles génératifs profonds sont appliqués avec succès à de nombreux cas d'apprentissage de données numériques, telles qu'images ou audio. Dans le domaine de la musique, ils peuvent également servir à la génération de représentations symboliques, dans le cadre de problèmes comme la génération automatique de musique ou la transcription [1-3]. Un défi important pour la génération de données symboliques structurées en général est d'obtenir des résultats bien formés.

C'est particulièrement vrai dans le cas de la musique. Il est en effet largement admis que la notation musicale représente, bien au-delà de simple suites de notes, une organisation hiérarchique de l'information mélodique et harmonique, induisant des dépendances non-locales entre les objets musicaux [4]. Une bonne représentation de cette information est essentielle à l'interprétation et à l'analyse de pièces.

Les auteurs de [5] étudient le cas des données symboliques structurées pouvant être représentées par des arbres syntaxiques d'une grammaire hors contexte, et proposent un traitement à l'aide d'un auto-encodeur variationnel (VAE), qui encode et décode directement ces arbres syntaxiques, pour une grammaire donnée, afin d'obtenir des données syntaxiquement valides - l'encodage est une représentation linéaire de calcul de la grammaire (pour produire un arbre). Cette approche a donné de bons résultats dans le domaine de la synthèse moléculaire.

Objectifs

Le but de cette thèse est l'étude de l'application de cette dernière approche au cas de la notation musicale.

On s'appuiera pour cela sur un modèle arborescent de la notation musicale que nous développons comme représentation intermédiaire pour le traitement de problèmes tels que la transcription musicale, la recherche d'information dans les bases de partitions et l'analyse de partitions [6-7]. Il conviendra de définir pour le cadre voulu des adaptations appropriées de ce modèle arborescent, des grammaires générant de tels arbres et de transformations d'arbres adéquates [8]. Un corpus d'arbres sera constitué à partir de partitions au format XML et de techniques d'augmentation de données, et appliqué pour l'entraînement de modèles.

Nous visons deux applications principales:

- la *génération de performances* musicales de piano (au format symbolique MIDI) à partir de partitions,
- et le problème inverse qui est la *transcription* d'enregistrements de piano MIDI en partitions [3].

On pourra utiliser dans ce cadre de larges bases de partitions piano au format XML et de performances MIDI correspondantes (qu'il s'agisse d'enregistrements de vrais musiciens sur claviers électroniques ou de performances artificielles). Le problème de la génération automatique

pourra aussi être approché, par exemple par imitation (d'autres partitions) ou pour produire un résumé, ou encore depuis une description du morceau voulu.

Candidature

Les candidats titulaires d'un master (ou équivalent international) en informatique, curieux de l'apprentissage automatique appliqué à la recherche d'informations musicales sont bienvenus. Ce projet demande de la créativité dans la réflexion et la programmation. Une expérience préalable en apprentissage automatique et des connaissances en notation musicale sont un atout important.

La thèse se déroulera au laboratoire Cedric du CNAM, Paris.

Références

[1] Jean-Pierre Briot, Gaëtan Hadjeres, and François Pachet. Deep Learning Techniques for Music Generation. Vol. 1 series Computational Synthesis and Creative Systems, Springer, 2020.

[2] Cheng-Zhi Anna Huang et al. Music Transformer. arXiv preprint arXiv:1809.04281 2018.

[3] Masahiro Suzuki. Score Transformer: Generating Musical Score from Note-level Representation. ACM Multimedia Asia. 2021.

[4] Stefan Koelsch, et al. Processing of Hierarchical Syntactic Structure in Music. Proceedings of the National Academy of Sciences 110.38, 2013.

[5] Matt J. Kusner, Brooks Paige, and José Miguel Hernández-Lobato. Grammar Variational Autoencoder. International conference on machine learning. PMLR, 2017.

[6] Francesco Foscarin et al. A Parse-based Framework for Coupled Rhythm Quantization and Score Structuring. 7th International Conference on Mathematics and Computation in Music (MCM 2019). see also <<https://qparse.gitlabpages.inria.fr>>

[7] Tiange Zhu et al. A Framework for Content-Based Search in Large Music Collections. Big Data and Cognitive Computing, 6 (1): 23, 2022.

[8] Florent Jacquemard, Pierre Donat-Bouillud, Jean Bresson. A Structural Theory of Rhythm Notation Based on Tree Representations and Term Rewriting. 5th International Conference on Mathematics and Computation in Music (MCM) 2015.