

ANALYSE ET ESSAI DE RECONSTRUCTION DE STRUCTURES ARBORESCENTES DES GRILLES DE JAZZ

Patrice Thibaud Mathieu Giraud

Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRISTAL
F-59000 Lille, France

RÉSUMÉ

L'harmonie tonale s'organise autour d'enchaînements V/I et d'autres schémas harmoniques. Elle peut se modéliser par des arbres. La Jazz Harmony Treebank (JHT) regroupe ainsi 150 analyses de grilles de jazz sous forme arborescente. Nous analysons ici quelques aspects de la répartition des enchaînements d'accords et schémas suivant les arborescences de la JHT, et nous proposons un algorithme pour reconstruire ces arbres à partir de la grille. Selon les scénarios, cet algorithme regroupe autour de 60% des accords des grilles, avec une précision d'environ 70%, confirmant la prédominance de certains schémas harmoniques dans la structuration de ces grilles.

1. GRILLES

L'harmonie tonale s'organise autour de la gamme et d'enchaînements harmoniques. Elle peut se penser dans une structuration arborescente – que ce soit en considérant les différents niveaux d'analyse schenkérienne [13], ou, plus explicitement, dans les modèles introduits par Lerdahl et Jackendoff dans la *Generative Theory of Tonal Music* [9]. Rohrmeier propose plus précisément une organisation de l'harmonie tonale sous forme d'arbres modélisés par des grammaires, incluant des suites de *prolongations* et de *préparations* [11].

Jazz et grilles. Une représentation des suites d'accords sous forme de *grilles* est couramment utilisée afin de permettre au musicien d'accompagner, d'improviser et d'apprendre de manière plus aisée la suite de progressions harmoniques. En jazz, ces grilles sont souvent complétées de la ligne mélodique pour obtenir une *lead sheet*.

En 1942, George Godwin (directeur d'une station de radio) a imaginé les *Tune-dex*, cartes à destination des professionnels de la radio indiquant, au recto, des informations éditoriales que nous considérons aujourd'hui comme des *méta-données* (compositeur, parolier, droits d'auteur etc.), et, au verso, des *données* musicales succinctes, mélodie et accords [8].

Ces cartes n'étaient pas spécifiquement destinées au jazz, mais des musiciens s'en sont emparées afin de posséder un vaste répertoire pour jouer au débotté n'importe quelle chanson en faisant illusion (*to fake my way through*). Expurgées de leurs données éditoriales, les *lead sheets* ont

été assemblées, indexées (parfois thématiquement), naissance aux premiers *Fake books*, illégaux. L'émergence des Real Books quelques années plus tard et aujourd'hui le partage de grilles par l'intermédiaire de sites spécialisés fournissent de formidables sources à disposition des musiciens.

Jazz et MIR, Corpus. La recherche en informatique musicale (et, plus précisément, en MIR, music information retrieval) peut concerner de nombreux aspects sur le jazz autour des grilles, mais aussi des lead sheets, des solos, d'audio voire de métadonnées. On peut citer les travaux faits autour des transcriptions de solo de la Weimar Jazz Database [10] ou bien l'archive des enregistrements du Montreux Jazz Festival [4].

Les grilles sont un point particulier d'étude. Chemillier a proposé de générer des grilles suivant des grammaires formelles [2]. Shanahan et Broze [14] ont collecté 1200 grilles d'accords depuis le forum *irealb*¹ pour constituer un corpus de 1160 morceaux en format *jazz* (format apparenté au *kern* [7]). Sur 50 ans de jazz (entre 1920 et 1970) que traversent ce corpus, leur étude a notamment montré que les progressions de deux accords les plus couramment observés étaient ii-V, V-I et V/ii-ii [1]. En 2020, Harasim et al. ont analysé 150 de ces grilles (en 3/4 ou 4/4) sous forme d'arbres syntaxiques dans la JHT (Jazz Harmony Treebank) [6].

Chaque arbre est construit à l'aide d'une grammaire qui repère dans la grille les prolongations et les préparations d'accords (partant de la sous-dominante ou de la dominante), dégageant pour chaque morceau des plus hauts niveaux de représentation de l'enchaînement d'accords. On peut ainsi directement relier des accords loin les uns des autres dans la grille. La JHT a été réalisée manuellement, certes en voulant reconstruire prolongations et préparations [12], et également avec une relecture et validation, mais jusqu'à ce jour sans visée algorithmique.

Motivations. Nous étudions ici les arbres de la JHT, que ce soit par une analyse ou une ébauche de reconstruction. Nous avons conscience que différentes versions de grilles coexistent pour une même œuvre et que par ailleurs, le corpus sur lequel s'appuie la JHT repose sur des accords très souvent non enrichis (pas de 9^e, 11^e... ou de précision

1. <http://irealb.com/forums>

intervalles	P1	P4	M7	P5	M2	M6	m7	A4	m2	A8	m3	M3	A1	m6	d5	A4
feuilles consécutives	24.0	44.0	5.8	4.2	3.7	4.5	1.8	1.8	1.7	1.7	1.7	1.4	1.2	1.0	0.9	1.8
feuilles sœurs	43.5	43.0	5.6	1.5	1.4	0.7	0.1	0.1	0.9	1.4	0.2	0.3	0.8	0.4	0.2	0.1
nœuds internes frères	33.5	46.5	6.0	1.6	4.7	1.4	0.3	0.3	0.5	1.8	0.9	0.6	0.4	0.8	0.2	0.3
profondeur 1	90.0	4.7	–	–	–	–	–	–	–	–	3.3	–	–	2.0	–	–
profondeur 2	49.6	37.6	1.8	1.1	3.9	1.1	–	–	–	1.4	1.8	0.4	–	1.4	–	–
profondeur 3	43.1	33.4	6.2	2.6	6.8	2.6	0.4	0.6	–	1.0	1.2	0.6	–	1.4	0.2	0.6
profondeur ≥ 4	27.5	51.7	6.7	1.6	4.7	1.3	0.3	0.3	0.6	2.1	0.7	0.7	0.6	0.6	0.3	0.3

Table 1 – Présence dans la JHT des intervalles les plus fréquemment rencontrés en %. Les autres intervalles ont une fréquence inférieure à 0.2%

Algorithme 1 : Reconstruction des progressions harmoniques arborescentes.	
Entrées :	un tableau de feuilles $[f_1 \dots f_n]$, un scénario d'intervalles $[I_1 \dots I_m]$
Sorties :	un tableau d'arbre(s)
Initialiser A à $[f_1 \dots f_n]$	
pour chaque intervalle I du scénario faire	
si $I = P1^4$ alors	
pour toute suite de 4 arbres $[a_i, a_{i+1}, a_{i+2}, a_{i+3}]$	
avec $r = \rho(a_i) = \rho(a_{i+1}) = \rho(a_{i+2}) = \rho(a_{i+3})$, faire	
remplacer dans A ces 4 arbres par un seul arbre $((a_i, a_{i+1}), (a_{i+2}, a_{i+3}))$ de racine r .	
fin	
sinon	
pour toute suite de 2 arbres $[a_i, a_{i+1}]$	
avec $\rho(a_{i+1}) - \rho(a_i) = I$, faire	
remplacer dans A ces 2 arbres par un seul arbre (a_i, a_{i+1}) de racine $\rho(a_{i+1})$.	
fin	
fin	
fin	
Renvoyer A	
Pour un arbre a , $\rho(a)$ donne la tonique de l'accord porté par sa racine.	
Les parcours "Pour toute suite..." se poursuivent de gauche à droite tant qu'il y a des suites d'arbres vérifiant la condition.	

"même accord") sont généralement privilégiés avant les P4 (Table 1). Il y a même souvent des regroupements de P1 liant quatre accords identiques.

Ceci nous donne un indice fort sur les premières règles à appliquer si on souhaite reconstruire algorithmiquement un tel arbre.

Étude par profondeur. Afin de percevoir la structure a un plus haut niveau, nous avons analysé les intervalles entre tous les nœuds frères, qu'ils soient internes ou non, selon leur profondeur dans l'arbre, en considérant 0 comme la racine (Table 1). Le niveau 1 a une plus forte proportion de P1, ce qui traduit les prolongations fortes des accords correspondants aux tonalités des différents morceaux, y compris le retour à l'accord de I lorsque l'on reboucle sur le premier accord de la grille.

On retrouve des fréquences plus importantes pour les P4 lorsque la profondeur augmente, traduisant l'enchaînement des cycles de quinte couramment rencontrés dans le corpus.

3. VERS UNE RECONSTRUCTION ALGORITHMIQUE DES ARBRES

Prenant en considération les éléments précédents, nous proposons un algorithme de reconstruction d'un arbre, possiblement partiel, étant donné une grille d'accords. Les prochains paragraphes décrivent cet algorithme ainsi que l'évaluation de son efficacité.

Algorithme et scénarios d'intervalles. L'idée de base est de détecter les P4, mais également les P1 ainsi que les autres intervalles. L'algorithme regroupe ainsi au fur et à mesure les accords en testant les intervalles selon un scénario qui est une suite d'intervalles (Algorithme 1). L'algorithme maintient un tableau A de sous-arbres (initialement toutes les feuilles), et, pour un intervalle I du scénario, regroupe de gauche à droite les sous-arbres dont les racines respectent l'intervalle I . En sortie d'algorithme nous obtenons donc un tableau de sous-arbres, éventuellement réduit à un seul arbre lorsque tout a été regroupé.

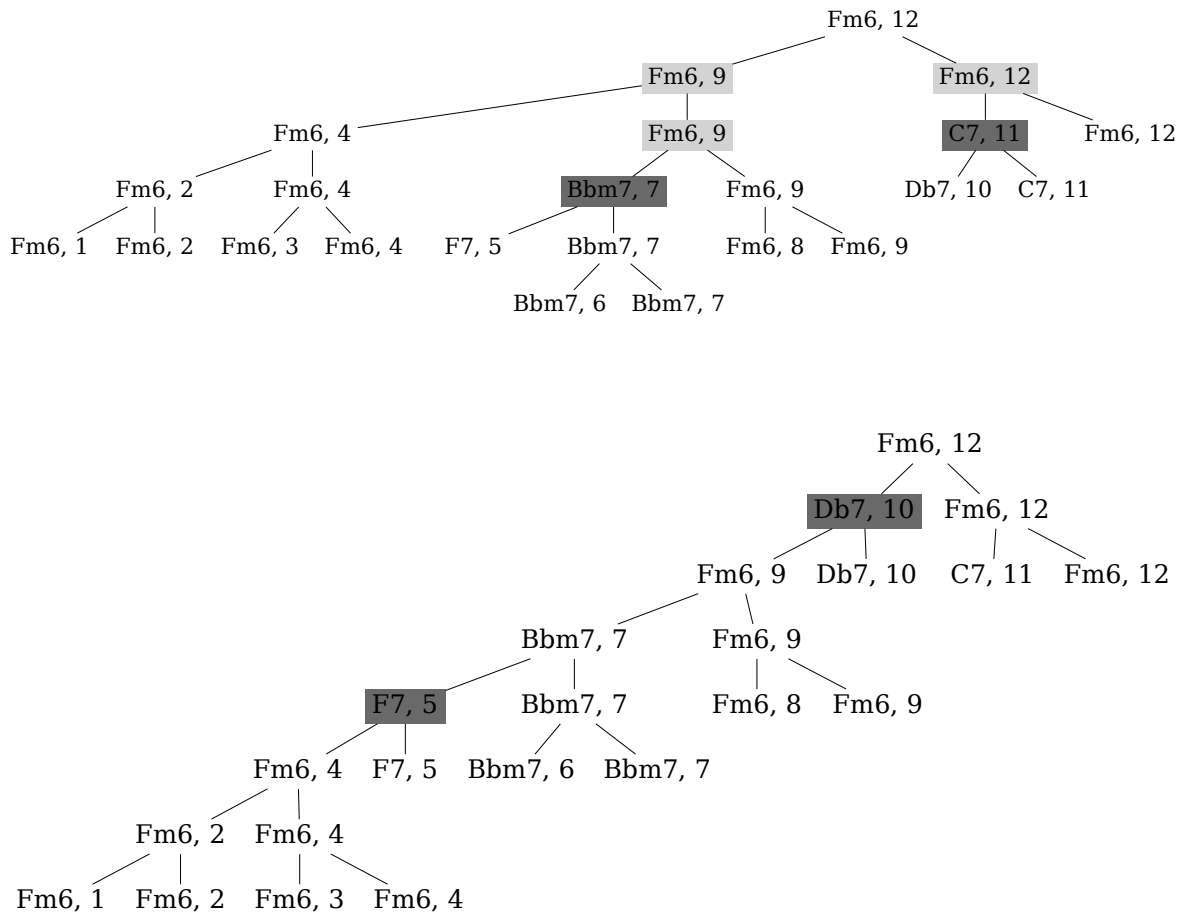


Figure 3 – Analyse arborescente de la grille d’*Interplay* (Bill Evans, 1963). (*haut*). Arbre proposé par la JHT (voir <https://github.com/DCMLab/JazzHarmonyTreebank/blob/master/tree-plots.md#interplay>), avec, pour chaque noeud interne, remontée depuis le fils droit de la dernière position dans lequel l’accord apparaît. (*bas*). Arbre reconstruit par le scénario 4, avec également remontée des positions depuis le fils droit. Deux noeuds (en grisé sombre) sont mal prédits, menant à une précision de 81.8%. D’autres regroupements auraient du être fait à la place de ces noeuds (noeuds grisés sombre dans l’arbre de la JHT), cette différence se propageant sur les noeuds parents (noeuds grisés clairs).

Nous avons retenu quatre scénarios :

- scénario 1 : P1*, P4, P1*
- scénario 2 : P1*, P4, P1*, P4, P1*, [P5, M2, M6], P1*
- scénario 3 : P1*, P4, M7, P1*, P4, [P5, M2, M6], P1*, P4, P1*
- scénario 4 : P1*, P4, M7, P1*, P4, [P5, M2, M6], OTHERS, P1*, P4, P1*

Tous les scénarios font régulièrement intervenir P1*, qui est en fait une suite P1⁴ - P1 : nous cherchons à rassembler d’abord s’ils existent quatre accords identiques deux à deux, puis à regrouper deux par deux ces accords. Ceci traduit la présence non négligeable dans les grilles analysées de P1, mais aussi de trois P1 consécutifs.

Nous avons également regroupé deux ensembles d’intervalles : [P5, M2, M6], de part leur fréquence ainsi que OTHERS, tous les autres intervalles : [A8, m3, m6, M3,

m2, A1, m7, A4, d5, A2, A3, d4, d1]. Les scénarios 1-3 ne prennent pas en considération la totalité des intervalles potentiellement rencontrés, et conduisent à des reconstructions partielles et donc à moins de noeuds que l’arbre de la JHT.

Évaluation. Nous évaluons l’algorithme en comparant les arbres prédits avec ceux de la JHT. Comparer des arbres peut se faire par distance d’édition [5]. Nous proposons ici de repérer les noeuds différents ou plus précisément les noeuds qui n’ont pas été regroupés de la même manière par notre algorithme.

On numérote dans l’ordre de la grille les feuilles puis on fait remonter ces positions à chaque fois du fils droit à son père et ceci jusqu’à la racine pour les 2 arbres (JHT et reconstruit). Un *noeud bien prédit* est un noeud ayant même accord *et* une même position.

Prenons pour exemple *Interplay*, blues mineur de Bill Evans – même si le caractère tonal du blues peut être contesté. Le scénario 4 reconstruit totalement un arbre (Figure 3). La plupart des regroupements sont corrects, mais deux nœuds internes sont mal prédits :

- (*F7,5*) au lieu de (*Bbm7, 7*) : le scénario regroupe par P1* (*Fm6,4*) et (*F7,5*), sans considérer le caractère mineur ou majeur des accords ;
- (*Db7,10*) au lieu de (*C7,11*) : le regroupement M7 aurait du ici être effectué avant le P4 vers (*Fm6, 12*).

Un arbre à n feuilles a $n - 1$ nœuds internes (racine comprise). Ainsi on peut comptabiliser pour les nœuds internes de notre arbre reconstruit (ou de nos sous-arbres), le nombre de nœuds bien prédits m (par rapport à ceux de la JHT) et ceux mal prédits m' . On définit ainsi la *précision* $p = \frac{m}{m+m'}$ et la *sensitivité* $s = \frac{m}{n-1}$ de l'algorithme sur un morceau.

Dans *Interplay*, il y a $n = 12$ feuilles et 11 nœuds internes, tous prédits, avec $m = 9$ nœuds internes bien prédits et $m' = 2$ mal prédits (Figure 3). On a donc ici $p = s = \frac{9}{11} \approx 81.8\%$.

scénario	precision moyenne	sensitivité moyenne
1	0.78	0.56
2	0.72	0.60
3	0.72	0.64
4	0.69	0.68

Table 2 – Sensitivité et de la précision moyenne de la reconstruction sur les 150 grilles de la JHT avec les différents scénarios.

Implémentation et Résultats. L'algorithme 1 a été implémenté en Python, en analysant les fichiers de la JHT (fournis au format *JSON*) et en utilisant *music21* [3] pour exploiter les accords et intervalles. Le code est disponible en licence libre GPLv3+ à l'adresse www.algomus.fr/code.

La Table 2 détaille les précisions et sensibilités moyennes de tous les scénarios. Le scénario 1 est le plus précis (78%)². De plus 28 morceaux (18.6% du total) dépassent les 90% de précision, les nœuds des sous-arbres reconstruits sont très proches de ceux rencontrés dans l'arbre de la JHT.

Les scénarios 2, 3 et 4 prennent en compte plus d'intervalles et prédisent donc plus de nœuds, d'où une sensibilité accrue (68% en moyenne pour le scénario 4), même si c'est parfois au prix d'une baisse de la précision (69%). Par exemple, dans *Interplay*, le nœud (*Db7, 10*) mal prédit

est ainsi la conséquence du M7 dans le scénario 4 et fait baisser la précision.

4. CONCLUSION ET PERSPECTIVES

Les arbres de la JHT révèlent la structuration harmonique des grilles. Les regroupements sont en particulier effectués suivant les P1 et P4, confirmant les prolongations et préparations dans les progressions harmoniques. Rassemblant successivement les accords selon ces intervalles les plus fréquents, l'algorithme proposé, selon les scénarios, arrive à traiter la majorité des accords avec une précision allant de 69% à 78%, et donne ainsi une ébauche de structuration de la grille. Certes, le but n'est pas uniquement de reconstruire automatiquement ces arbres car la vérification de leur pertinence reste à réaliser. Cependant, progresser dans ces algorithmes permet de progresser aussi dans la compréhension du langage musical et de sa modélisation.

Le choix initial de ne s'intéresser qu'aux intervalles entre les fondamentales pouvait paraître surprenant car il mettait de côté la fonction harmonique des accords. Dans plusieurs situations, il est très clair que les regroupements n'auraient pas été les mêmes. Les scénarios proposés ont tenté de tenir compte de cet écueil en intégrant des intervalles caractérisant quelques enrichissements courants en jazz (comme par exemple les substitutions tritoniques). De la même manière, le caractère mineur ou majeur d'un accord n'est pas nécessairement significatif dans l'analyse des enchaînements d'accords.

Il est plus difficile de regrouper les intervalles moins fréquents, en particulier lors de modulations ou de certaines substitutions. Les scénarios pourraient être affinés en ce sens. Plus généralement, il serait possible de générer et d'évaluer des scénarios dans une logique d'apprentissage machine, qui, vu la taille du corpus, serait délicate à effectuer sans sur-apprentissage.

L'évaluation serait aussi à approfondir. D'autres métriques sont possibles, notamment pour préciser la condition selon laquelle un nœud est bien prédit (remontée du fils gauche ou d'un intervalle). Plus généralement, on pourrait estimer une distance d'édition, mais c'est alors un défi de préciser quelles opérations d'édition seraient pertinentes ainsi que leur pondération.

Enfin, l'algorithme actuel ne prend en compte que la succession des accords. Tenir compte de la durée de ces accords, et, plus généralement, de la *forme* du morceau, permettrait probablement d'avoir des regroupements qui seraient structurellement plus pertinents.

Remerciements. Nous remercions Ken Déguernel pour les discussions sur cette étude ainsi que les relecteurs pour leurs remarques pertinentes.

2. Si l'on considérait aussi les feuilles, qui sont bien identifiées entre les arbres reconstruits à ceux de la JHT, la précision serait ici de 89.0% et la sensibilité de 78.1%.

5. REFERENCES

- [1] Yuri Broze and Daniel Shanahan. Diachronic changes in jazz harmony : A cognitive perspective. *Music Perception : An Interdisciplinary Journal*, 31 :32–45, 2013.
- [2] M. Chemillier. Improviser des séquences d’accords de jazz avec des grammaires formelles. In *Journées d’Informatique Musicale (JIM 2001)*, pages 121–126, 2001.
- [3] Michael Scott Cuthbert and Christopher Ariza. music21 : A toolkit for computer-aided musicology and symbolic music data. In *International Society for Music Information Retrieval Conference (ISMIR 2010)*, pages 637–642, 2010.
- [4] Alain Dufaux and Thierry Amsallem. The Montreux Jazz Digital Project : From preserving heritage to a platform for innovation. *Journal of Digital Media Management*, 7(4) :315–329, 2019.
- [5] Serge Dulucq and Hélène Touzet. Analysis of tree edit distance algorithms. In *Combinatorial Pattern Matching (CPM 2003)*, pages 83–95, 2003.
- [6] Daniel Harasim, Christoph Finkensiep, Petter Ericson, Timothy J. O’Donnell, and Martin Rohrmeier. The Jazz Harmony Treebank. In *International Society for Music Information Retrieval Conference (ISMIR 2020)*, pages 207–215. ISMIR, 2020.
- [7] David Huron. Music information processing using the Humdrum toolkit : Concepts, examples, and lessons. *Computer Music Journal*, 26(2) :11–26, 2002.
- [8] 99% invisible. *The Real Book*, episode 438, 2021.
- [9] Fred Lerdahl and Ray Jackendoff. *A Generative Theory of Tonal Music*. MIT Press, 1983.
- [10] Martin Pfeleiderer, Klaus Frieler, Jakob Abeßer, Wolf-Georg Zaddach, and Benjamin Burkhart, editors. *Inside the Jazzomat – New Perspectives for Jazz Research*. Schott Campus, 2017.
- [11] Martin Rohrmeier. Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1) :35–53, 2011.
- [12] Martin Rohrmeier. The syntax of jazz harmony : Diatonic tonality, phrase structure, and form. *Music Theory and Analysis (MTA)*, 7 :1–63, 2020.
- [13] Heinrich Schenker. *Der freie Satz*. Universal Edition, 1935.
- [14] Daniel Shanahan and Yuri Broze. iRealPro corpus of jazz standards, 2019.