
Une analyse musicale automatique suivant une heuristique perceptive

Olivier Lartillot

*Ircam – Centre Pompidou
Place Igor-Stravinsky
75004 Paris
Olivier.Lartillot@ircam.fr*

RÉSUMÉ. Un modèle computationnel général de la musique est présenté, basé sur deux caractéristiques principales de la musique : une direction chronologique et une sélectivité locale. Le concept de motif musical est défini sous la forme d'une agrégation d'intervalles locaux successifs. Les motifs sont induits par analogie entre un contexte actuel et des contextes passés similaires qui sont réactivés par l'intermédiaire d'une mémoire associative. Les motifs sont des concepts qui se réalisent en des occurrences dans la partition. La partition est représentée sous forme d'un réseau de notes reliées en occurrences de motifs qui eux-mêmes forment des meta-motifs. Une telle modélisation informatique peut offrir à la musicologie une compréhension détaillée et explicite de la musique et peut suggérer aux sciences cognitives les conditions nécessaires d'une perception virtuelle de motifs musicaux.

ABSTRACT. This paper introduces a general computational model of music perception that relies on two main temporal characteristics of music: chronological direction and short-term selectivity. As a result, musical pattern is defined as an aggregation of successive local intervals. Patterns are induced by analogy between a current context and similar past contexts that are reactivated through associative memory. Here, patterns are conceived of as concepts that are actualized in the musical score. This score is represented as a network of notes, which are linked to pattern occurrences that themselves form meta-patterns of patterns. Such a computer modeling may offer to musicology a detailed and explicit understanding of music, and may suggest to cognitive science the necessary conditions for a virtual perception of musical pattern.

MOTS-CLÉS : musique, motif, intervalles, cognition, analogie, mémoire associative, réseau conceptuel.

KEY WORDS : music, pattern, interval, cognition, analogy, associative memory, conceptual network.

1. Analyse musicale et extraction de données

L'internet fourmille de ressources musicales, représentées en général sous forme sonore, mais également, de manière plus abstraite, en une description textuelle d'événements musicaux suivant le format standardisé MIDI. Cette dernière représentation, que l'on peut considérer comme une partition numérique, jouit des mêmes qualités d'abstractions que le texte écrit comparé au langage oral. Les ressources musicales numériques, elles aussi, demandent à être analysées et structurées. C'est ainsi qu'une nouvelle discipline dédiée à la recherche d'information musicale, ou *Music Information Retrieval*, œuvre à une meilleure organisation de la gestion des réseaux d'informations, une navigation à travers les contenus, une indexation et des outils de moteurs de recherche [LAR 02a] [LAR03].

Une des tâches les plus élémentaires que l'on attend d'un système d'analyse musicale automatique consiste en la découverte des motifs musicaux répétés, ou *Musical Pattern Discovery*. De nombreuses recherches ont tenté d'importer dans le domaine musical des méthodologies d'analyse de texte ou de génomes. Ces approches ont permis certaines avancées significatives. Cependant, les limitations auxquelles se heurtent dorénavant ces méthodes tiennent du fait même qu'elle ne peuvent saisir certaines dimensions spécifiques au phénomène musical.

La musique se démarque en effet du langage naturel ou du génome par son caractère polyphonique : la musique n'est pas une ligne temporelle simple, mais un flux entrelacé, qui ne peut pas, de manière générale, être scindé en voies monophoniques.

De plus, même si l'on considère la musique comme une séquence de symboles, ce séquençement est régi par une double paramétrisation□

- d'une part, la position des symboles au sein de la séquence ne se réduit pas à la simple succession. Le séquençement temporel de la musique est décrit dans un espace métrique, permettant de déterminer précisément les distances temporelles entre éléments□ l'*espace temporel*, représenté par l'axe horizontal de la partition.
- d'autre part, ces symboles eux-mêmes, indépendamment de leur position temporelle, se situent dans un autre espace métrique□ l'*espace des hauteurs*, représenté par l'axe vertical de la partition.

Cette double paramétrisation du séquençement n'existe ni dans la génomique, ni dans le langage naturel, et doit donc être considéré à sa juste valeur. En particulier, du fait de cette description précise suivant ces deux espaces — dont le produit forme l'*espace musical* — les répétitions musicales sont rarement exactes. Ainsi, dans cet exemple, figure 1, les deux motifs de trois notes peuvent être jugés similaires bien que non identiques□



Figure 1. Deux motifs assez similaires

L'approximation, telle qu'elle est considérée en algorithmique de texte ou en bioinformatique, est exprimée par des opérations de substitutions, d'insertions et de suppressions. L'approximation musicale, quant à elle, semble devoir être formulée en terme de ressemblance entre *intervalles* entre notes successives¹. Ces intervalles sont ainsi des vecteurs entre deux positions de cet espace bidimensionnel. Nous proposons une nouvelle approche de découverte de motifs musicaux basée sur une distance métrique entre intervalles.

2. Une approche temporelle

La musique étant un medium de communication humaine, une analyse automatique n'aura de sens que si elle met en œuvre des heuristiques d'ordre perceptuel [LAR 02b].

2.1. Une approche incrémentale

La musique n'est pas un objet simple, mais un objet temporel : il est perçu de manière progressive suivant un sens chronologique. Toute analyse musicale automatique qui imiterait les mécanismes perceptuels devrait alors suivre une approche incrémentale de la partition.

La *note* est l'objet musical élémentaire. Une note n est caractérisée par ses coordonnées dans l'espace musical, décrit ci-dessus. Une *partition* est un ensemble de notes, ordonné suivant une dimension chronologique. L'analyse de la partition consiste en une inférence progressive et incrémentale de connaissance. Un pointeur parcourt la partition le long de l'axe chronologique. Chaque incrément de l'algorithme correspond donc à un instant temporel de la partition — l'*instant pointé* — et aux notes qui lui sont associées.

2.2. Une sélectivité locale

Le temps se manifeste également sous un autre aspect, lié aux caractéristiques de la mémoire humaine : à chaque phase du parcours incrémental de la partition musicale, l'instant actuel est immédiatement accessible à notre perception, le passé récent reste disponible à la perception (c'est la « *réention* » au sens de Husserl [HUS 05]), alors que les instants éloignés ne peuvent être remémorés (en un « *ressouvenir* ») que de manière indirecte, suivant un mécanisme décrit plus bas. Un tel mécanisme cognitif peut être formalisé ainsi□

¹ L'idée de succession doit ici être considérée dans un sens large. Deux notes sont successives si l'une précède l'autre et sont « assez » proches l'une de l'autre.

- D'un point de vue spatial (la partition en tant qu'espace), au pointeur qui parcourt la partition est associée une fenêtre temporelle d'étendue limitée, qui sélectionne les notes appartenant au passé proche de l'instant pointé.
- D'un point de vue temporel, un tampon garde en mémoire les précédentes notes des instants récemment pointés.

À chaque incrément de ce parcours, une liaison particulière est établie entre les notes «retenues» et celles de l'instant actuel. Il s'agit non pas d'une simple liaison, mais d'une opération d'agrégation élémentaire des éléments du passé proche aux éléments présents, afin de construire progressivement une séquence. Est donc perçu un mouvement dirigé de la note retenue vers la note présente. Cette liaison élémentaire, qui joue ainsi un rôle primordial dans la perception musicale, sera appelée *intervalle local* (nous proposons ainsi une acception particulière du concept musical d'intervalle), et sera représentée non pas simplement sous la forme d'un couple de notes ($n1, n2$), mais plutôt sous la forme d'un vecteur de $n1$ vers $n2$.

2.3. Une mémoire associative

Les paramètres des éléments locaux — que ce soit les notes élémentaires ou les intervalles locaux — peuvent s'avérer être des propriétés caractéristiques qui, s'ils se manifestent de nouveau ultérieurement, rappelleront les anciennes instances similaires.

Nous proposons de modéliser ce mécanisme associatif de ressouvenir sous la forme d'un réseau sémantique. Nous reprenons ici à notre compte la modélisation cognitive des mécanismes inductifs — car il s'agit bien ici d'induction, c'est-à-dire d'inférence de nouvelles connaissances à partir de l'expérience — sous la forme d'activation de concepts au sein d'un réseau conceptuel [HOL 89]. Dans cette approche, les inductions sont proposées par le système lorsque les degrés d'activation dépassent un certain seuil.

Dans notre contexte musical, nous proposons une vision de la partition elle-même sous forme d'un réseau de relations. Les motifs musicaux similaires sont reliés entre eux par l'intermédiaire de relations conceptuelles. Cette émergence de relations de similarité étant progressive, il est nécessaire de mettre en place, tout d'abord, des mécanismes de mise en relation de similarité des intervalles locaux. La similarité s entre intervalles locaux, indépendamment de tout contexte, est simplement définie sous la forme d'une distance entre intervalles au sein de l'espace musical. Nous proposons de scinder le concept de distance en autant de distances que de paramètres de l'espace musical. La distance résultante peut être considérée comme multidimensionnelle.

L'association des intervalles similaires peut être implémentée de manière simple à l'aide de tables de hachage associant pour chaque valeur du paramètre considéré tous les intervalles associés.

3. La construction des motifs

Une fois que deux intervalles sont considérés comme similaires, les intervalles qui les succèdent au cours du texte peuvent également se ressembler. Dans ce cas, est perçue non pas un couple de deux similarités entre deux couples d'intervalles, mais une seule similarité entre deux successions d'intervalles.

C'est ici que l'on réalise l'importance du mécanisme d'agrégation d'intervalles en séquences. Les deux successions de deux intervalles successifs peut alors être rassemblées en une seule structure qui pourra de nouveau être étendue par de nouveaux intervalles. Une telle succession d'intervalles, en tant que construction pertinente et récurrente au sein de l'oeuvre, peut alors être nommée *motif*.

La similarité entre deux intervalles, considérés comme candidats respectifs de la prolongation de motifs similaires, bénéficie de l'existence du contexte induit par ces motifs. En d'autres termes, l'existence préalable de ces deux motifs similaires renforce la similarité entre les intervalles candidats à leur extension. On peut également supposer que plus cette structure sera longue et fréquente au cours du texte, plus des distorsions pourront être tolérées.

3.1. Les motifs abstraits

Des motifs considérés comme équivalents sont associés à un concept qui les subsume. Ce concept est appelé *motif abstrait*, et les motifs de la partition qui lui sont associés sont ses *occurrences*. Grâce à cette abstraction, toute nouvelle répétition sera considérée comme une simple nouvelle occurrence du concept. Par l'intermédiaire de sa liaison avec toutes ses occurrences, le motif abstrait conserve toute l'information relative aux caractéristiques de ses occurrences.

En raison du caractère incrémental de la perception musicale, les motifs sont découverts progressivement, intervalles par intervalles, de l'intervalle initial au motif complet. L'architecture de l'ensemble des motifs abstraits doit donc prendre en compte cette contrainte cognitive. Pour cela, l'ensemble des préfixes de tout motif est également considéré comme un motif, en tant que motif « en devenir ». Les préfixes du motif abstrait sont donc des motifs abstraits. L'ensemble de ces préfixes est ordonné par ordre croissant de taille, et forme une chaîne.

Progressivement, pour chaque nouvelle note du motif découverte, un nouveau motif abstrait intermédiaire est créé et considéré comme une continuation du motif abstrait précédent. Ainsi toute répétition ultérieure d'un préfixe de ce motif ne sera considérée que comme une nouvelle occurrence du motif abstrait intermédiaire associé.

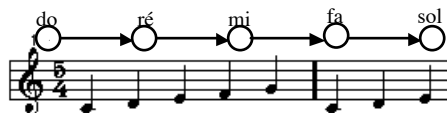


Figure 2. Un motif abstrait et deux occurrences, dont un préfixe.

3.2. L'arbre des motifs abstraits

Deux motifs de même préfixe partageront également le même motif abstrait intermédiaire, correspondant à ce préfixe, car ces motifs ne peuvent être discriminés si seul ce préfixe est présenté. Ainsi un motif abstrait peut accepter plusieurs continuations possibles. Il en résulte que l'ensemble des motifs abstraits peut être représenté sous la forme d'un arbre d'état, chaque état étant associé à une note, chaque branche à un motif abstrait : c'est l'arbre des motifs abstraits (a.m.a.).

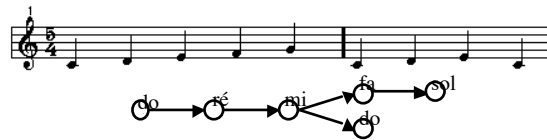


Figure 3. Deux motifs et l'arbre correspondant.

3.3. La chaîne d'occurrence de motif

Une occurrence d'un motif abstrait correspond à un parcours de la branche correspondante à ce motif dans l'a.m.a. Une interface supplémentaire doit mettre en relation la succession de notes dans la partition et le motif abstrait associé dans l'a.m.a. Cette structure se présente sous la forme d'une chaîne, dite chaîne d'occurrence de motif (c.o.m.), dans laquelle chaque état successif représente un état successif du parcours du motif. Ainsi chaque note de l'occurrence dans la partition est reliée à l'état correspondant dans la chaîne. De l'autre côté, chaque état de ce graphe est relié à l'état correspondant le long de la branche correspondante de l'a.m.a.

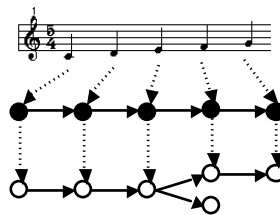


Figure 4. Un motif, son c.o.m. et l'a.m.a.

3.5. Motif cyclique

Si un motif est répété successivement plusieurs fois, ceux-ci deviennent eux-mêmes des objets élémentaires, appréhendables un peu à la manière d'une

succession de notes. Ainsi des intervalles locaux peuvent être considérés entre les occurrences des motifs.

À cette caractérisation classique du concept d'intervalle, il est nécessaire d'ajouter un paramètre supplémentaire, appelé intervalle motivique, relatif à la comparaison des deux occurrences du motif. Le degré de ressemblance entre les deux occurrences est exprimé sous la forme du motif abstrait le plus général les subsumant.

Les occurrences du motif pouvant ainsi être agrégées les unes aux autres, il en découle alors le concept de méta-motif de ces occurrences du motif de base (figure 5). Une occurrence de meta-motif de motifs peut être représenté également en terme de c.o.m., ou chaque état du graphe représente lui-même le c.o.m. d'un motif élémentaire. La découverte d'une occurrence d'un motif de motifs correspond alors à un parcours progressif du c.o.m. du méta-motif, ou chaque étape correspond au parcours du c.o.m. correspondant au motif élémentaire associé.

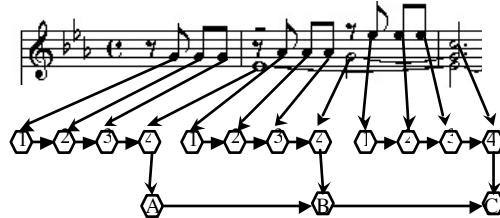


Figure 5. C.o.m. de meta-motif.

La comparaison entre meta-motifs nécessite la comparaison entre intervalles d'occurrences, mais également la comparaison de chaque occurrence respective dans la même position dans les deux méta-motifs.

3.6. Plusieurs niveaux de structuration

La musique ne présente pas une structuration motivique univoque. Il est possible de déceler, dans une section donnée, des motifs de divers tailles qui entrent en relation d'inclusion ou de chevauchement. La difficulté réside alors dans la mise en évidence des principes de compréhension de ces diverses structurations mis en jeu lors d'une simple écoute.



Figure 6. Jean-Sébastien Bach, *Prélude en do majeur, Clavier bien tempéré, début (simplifié)*.

Dans cet exemple, le motif principal de 8 notes est en fait lui-même structuré par une répétition parfaite d'un sous-motif de 3 notes. Chaque occurrence de ce

motif principal, bien qu'en proie à de multiples transformations, respecte scrupuleusement la structuration par répétition de deux sous-motifs. Remarquons en outre que le motif principal lui-même est toujours répété deux fois.

Le motif principal considéré ici ne peut être simplement caractérisé en terme de succession d'intervalles similaires. Il faut également pouvoir expliciter la structuration telle que nous l'avons expliquée ci-dessus, et — mieux encore — on doit attendre d'un système d'analyse informatique qu'il infère cette structuration automatiquement. Pour cela, à la description des motifs abstraits est ajoutée une caractérisation supplémentaire, relative à cette structuration interne, qui s'exprime sous la forme de c.o.m. apposée non pas, cette fois-ci, sur la partition, mais sur les branches de l'a.m.a.

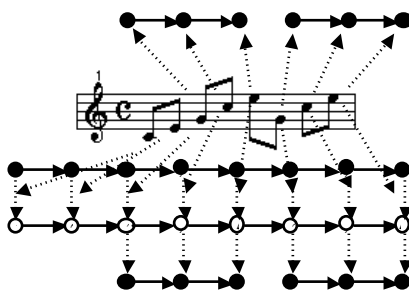


Figure 7. Un motif, les deux c.o.m. des petits motifs, la c.o.m. du grand motif, la branche dans l'a.m.a et les deux c.o.m. sur cette branche.

Cette mise en relation des divers motifs abstraits au sein même de l'a.m.a. permet de créer un réseau d'associations entre classes de motifs. Cette organisation de haut niveau peut avoir une influence sur la reconnaissance des motifs élémentaires. Des mécanismes protentionnels [HUS 05], c'est-à-dire des attentes [MEY06] sont ainsi générés par le système en cours d'analyse. Lorsque le système rencontrera une nouvelle occurrence d'un motif donné, il s'attendra à rencontrer également les motifs associés à ce motif, aux positions indiquées dans la branche correspondante à ce motif dans l'a.m.a.

4. Résultats.

L'algorithme présenté ci-dessus est en cours de conception et d'implémentation², sous la forme d'une librairie, appelée *OMkanthuss*, du logiciel de représentations musicales *Open Music* développé à l'Ircam [ASS 99]. Les premiers résultats, assez encourageants, tendent à prouver la pertinence de cette approche, mais montrent également qu'il reste encore beaucoup à faire avant de pouvoir espérer obtenir un algorithme performant. En outre, la recherche de motifs musicaux n'est qu'une

² Voir également le site internet consacré à ce projet, dont l'adresse est : www.ircam.fr/equipes/repmus/lartillot.

première étape d'un projet général d'analyse musicale, voire de découverte automatique de théorie musicale.

Ma thèse est dirigée par Emmanuel Saint-James (LIP6) et Gérard Assayag (Équipe Représentations Musicales, Ircam). Je tiens également à remercier Benoit Meudic pour ses remarques expertes.

5. Bibliographie

- [ASS 99] ASSAYAG G., RUEDA C., LAURSON M., AGON C., DELERUE O., «Computer Assisted Composition at Ircam: From PatchWork to OpenMusic », *Computer Music Journal*, vol. 23, n° 3, 1999, p. 59-72, MIT Press, Cambridge.
- [HOL 89] HOLLAND J. H., HOLYOAK K. J., NISBETH R. E., THAGARD P. R., *Induction: Processes of Inference, Learning, and Discovery*, MIT Press, Cambridge, 1989.
- [HUS 05] HUSSERL E., *Leçons pour une phénoménologie de la conscience intime du temps*, 1905, traduit par Henri Dussort, P.U.F, Paris, 1964.
- [LAR 02a] LARTILLOT O., «Integrating Pattern Matching into an Analogy-Oriented Pattern Discovery Framework », *International Conference on Music Information Retrieval ISMIR 2002*, Paris, 13-17 octobre 2002, Ircam – Centre Pompidou, Paris, p. 69-270.
- [LAR 02b] LARTILLOT O., «Generalized Musical Pattern Discovery by Analogy from Local Viewpoints », in LANGE, S., SATOH, K., SMITH C.H. (éds.): *Discovery Science, 5th International Conference DS 2002*, Lübeck, 24-26 novembre 2002, LNCS 2534, Springer, Berlin, p. 382-389.
- [LAR 03] LARTILLOT O., «Perception-Based Advanced Description of Abstract Musical Content », in *Digital Media Processing for Multimedia Interactive Services, 4th European Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 2003*, Londres, 9-11 avril 2003, World Scientific Press, Singapour.
- [MEY 56] MEYER L., *Emotion and Meaning in Music*, Chicago University Press, Chicago, 1956.