



HAL
open science

**Modélisation anticipative des systèmes musicaux.
Reconnaissance, génération, synchronisation et
programmation synchrone temps réel en informatique
musicale**
Arshia Cont

► **To cite this version:**

Arshia Cont. Modélisation anticipative des systèmes musicaux. Reconnaissance, génération, synchronisation et programmation synchrone temps réel en informatique musicale. *Revue des Sciences et Technologies de l'Information - Série TSI: Technique et Science Informatiques*, 2012, 31 (3), pp.311-335. 10.3166/tsi.31.311-335 . hal-00699290

HAL Id: hal-00699290

<https://inria.hal.science/hal-00699290>

Submitted on 20 May 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modélisation anticipative des systèmes musicaux

Reconnaissance, génération, synchronisation et programmation synchrone temps réel en informatique musicale

Arshia Cont*

* *Ircam CNRS UMR STMS*
1 Place Igor Stravinsky, 75004 Paris, France
arshia.cont@ircam.fr

RÉSUMÉ. Cette article étudie l'anticipation musicale comme un élément de conception pour des applications temps réel en d'informatique musicale. Modélisation anticipative est présenté comme un principe de conception pour modéliser des systèmes artificiels. Nous proposons des modèles anticipatifs concernant trois préoccupations principales de l'attente musicale : quoi attendre? comment attendre? quand attendre? Chaque sujet est abordé en appelant une littérature varié en géométrie de l'information, apprentissage actif et automatique, et synchronisation des flux en temps réel pour des signaux musicaux. La conception anticipative est la facture commune entre toutes ces propositions avec les prémices de réduire la complexité computationnelle de modélisation pour des problèmes complexes dans l'informatique musicale.

ABSTRACT. This paper studies musical anticipation as a design element for realtime computer music applications. Anticipatory modeling is presented as a design principle for modeling artificial systems. We propose anticipatory models and applications concerning three main preoccupations of expectation : What to expect? How to expect? and When to expect? Each question is addressed within different literatures such as information geometry, active and on-line learning, and stream synchronisation in realtime and for music signals. The anticipatory design concept is common in all propositions with the premise of reducing the structural and computational complexity for complex problems in computer music.

MOTS-CLÉS : Informatique musicale, modélisation anticipative, géométrie de l'information, apprentissage automatique, suivi de partition, programmation synchrone.

KEYWORDS: Computer music, Anticipatory Modeling, Information Geometry, Music Information Retrieval, Online Learning, Score Following, Synchronous Programming.

1. Introduction

L'*attente musicale* joue un rôle important dans la perception et la cognition de la musique. De nombreux chercheurs dans différents domaines le soulignent depuis plus d'un siècle. Dès que nous commençons à écouter de la musique, dans une salle de concert, sur une chaîne stéréo ou en multicanal, les phénomènes d'attente entrent en jeu. Aujourd'hui, le rôle d'*attente musicale* sous toutes ses formes est largement reconnu en psychologie auditive et en physiologie du système limbique (Huron, 2006). Beaucoup d'affects musicaux comme *la surprise*, *le frisson*, *la tension*, etc. sont liés au rapport qui se tisse entre la forme musicale et nos attentes qui dépendent notamment de nos connaissances. L'attente, dans ce sens, implique des représentations mentales, lesquelles sont constamment examinées et réactualisées par nos expériences musicales quotidiennes. Différents facteurs entrent en jeu dans la constitution de ces représentations mentales : le contexte culturel, le genre et schémas musicaux, la mémoire véridique ou conscient, ou même des éléments extra-musicaux. Le rôle de l'attente musicale n'est pas limité à celui de la perception d'écoute mais joue un rôle central dans l'acte de création. Meyer (1956), dans son livre "*Emotion and Meaning in Music*" souligne cet aspect en arguant le fait que le contenu principal de l'émotion musicale survient au travers de la chorégraphie d'attente créée par les compositeurs. Des cas explicites sont examinés et analysés dans le livre récent de David Huron (Huron, 2006).

Malgré son importance cognitive, l'attente musicale n'a pas été suffisamment prise en compte en informatique musicale (ni dans les approches computationnelles, ni pour l'extraction des données musicaux). Aujourd'hui la plupart des modèles computationnels de musique favorisent des architectures entraînées par *prédiction*. Tandis que de point de vue cognitif, la prédiction n'est pas l'unique résultat de l'attente musicale, mais juste un sous-produit de ce processus cognitif. Beaucoup d'effets d'attente musicale pourraient être unifiés dans un cadre global sous le terme de *comportement anticipatif*, défini comme un comportement qui non seulement dépend du passé et présent, mais aussi d'une prédiction, une attente, ou un croyance de future (Butz *et al.*, 2003a).

Nous commençons cet article en section 2 pour des survols séparé des cadres cognitifs et computationnelle de notre approche. Ceci nous ramène à une définition des *systèmes anticipatifs* comme une cadre de conception, tout en exposant leurs propriétés nécessaires. Section 3 à 5 vont exposer trois cadre computationnelles, liés à trois questions centrales de l'anticipation : *Quoi ?*, *Comment ?* et *Quand ?*. Section 3 propose un cadre mathématique de la géométrie de l'information pour quantifier et qualifier l'information contenu dans un flux continu et temps réel du son. Ce cadre nous permettra de gérer les séries temporelles et continues, avec des entités d'informationnelles et en utilisant leurs propriétés géométriques sur un espace riemannien. En particulière, nous montrons deux exemples applicatifs pour démontrer la puissance de ce type d'abstraction pour résoudre des problèmes souvent considérés complexes dans le domaine d'extraction des données musicales. Section 4 étudier la question de *comment* anticiper en interaction avec un environnement extérieur. Nous étudions

cette question dans le cadre d'apprentissage interactif et en-ligne et exposons des résultats dans le cadre d'une application d'improvisation automatique. Section 5 aborde la question de synchronisation des flux et en temps réel en utilisant une conception anticipative. En particulière, nous discutons comment l'accès à l'information et temps pourrait permettre aux artistes et programmeurs d'écrire des processus temporels à l'instar de l'écriture musicale. Sur ce volé, nous exposons *Antescofo*, un système de suivi de partition temps réel couplé avec un langage synchrone pour l'information musicale.

Enfin, cet article est un résumé étendu d'une thèse (Cont, 2008b) et a comme but d'introduire les concepts de modélisation anticipative et démontrer leurs puissances pour résoudre des problèmes complexes. Les lecteurs en recherche de plus de détails pourraient consulter le manuscrit. Section 3 est également exposée dans (Cont *et al.*, 2011). Section 4 est également exposé en plusieurs versions dans (Cont *et al.*, 2010; Cont *et al.*, 2007), et une version plus détaillée de section 5 est publié dans (Cont, 2010; Cont, 2008a).

2. De la modélisation d'anticipation à la modélisation anticipative

L'*anticipation* en tant que processus, diffère de la *prédiction* et de l'*attente*. L'*attente* est une forme de connaissance mentale ou corporelle de l'arrivée vraisemblable d'un événement ou une classe d'événement (Huron, 2006). La *prédiction* est un processus qui estime le future à partir du passé ou à partir d'une régularité environnementale dans le passé. Nous distinguons l'*anticipation* dans notre contexte computationnelle, en le designant comme un processus qui agit sur la base de la prédiction et/ou le retour de la croyance prédictif du système sur elle-même. Avant de spécifier notre approche de modélisation anticipative, nous regardons la littérature en psychologie de l'attente musicale pour en retirer des concepts clés pour une modélisation computationnelle.

2.1. Modélisation de l'anticipation musicale

La terminologie de l'*attente musicale* dans le domaine de la perception musicale est large et touche plusieurs aspects et sous-domaines de la recherche. L'idée ici est d'étudier la psychologie de l'attente pour en prendre en compte les aspects spécifiques de l'attente utiles pour la modélisation. Trois domaines d'études ont des conséquences de modélisation explicites : l'apprentissage auditif, la représentation mentale et la mémoire auditive. En guise de résumé :

- Le facteur déterminant pour l'apprentissage d'un régularité est sa *stabilité* environnementale.
- Les auditeurs sont sensibles à la statistique des événements dans leur environnement sonore, a qui met en évidence l'aspect statistique de l'apprentissage auditif (Saffran *et al.*, 1999; Loui *et al.*, 2006).

– L'exposition à un environnement stable donne lieu aux *attentes*, sous forme de *représentations mentales*.

– Les représentations mentales (ou attentes) ne sont pas exactes, ni infaillibles.

– Pour chaque phénomène sonore, il existe plusieurs représentations qui rivalisent et collaborent entre elles.

– Dans l'acte de perception, ces différentes représentations sont différenciellement favorisées en fonction de leurs succès *prédictifs* et de l'expérience individuelle de chaque individu. L'interaction avec l'environnement extérieur donne lieu à un renforcement et à une récompense.

Nous regardons à présent les *modèles* proposés dans la littérature pour l'*attente musicale*. Ces modèles viennent de domaines différents et peuvent être regroupés sous trois catégories : *théorie musicale*, *apprentissage automatique* et *théorie de l'information*. Au sein du premier groupe, nous étudions les modèles concernant l'*attente mélodique* et notamment les modèles proposés par Narmour (1992) et Margulis (2005). Les approches computationnelles venant de la théorie musicale ont la plupart du temps comme but d'imposer un cadre computationnel souvent simple au déroulement des flux d'information musicaux afin de répliquer des phénomènes liés à l'attente musicale provenant d'une analyse de la partition musicale. Ces modèles sont souvent biaisés par les préférences théoriques de leurs auteurs, parfois liées à un style particulier de musique. Le deuxième groupe en revanche n'a aucun présupposé quant aux données musicales et essaye d'apprendre les relations qui donnent lieu à la représentation mentale de l'attente à l'aide de l'apprentissage automatique. La troisième approche a donné lieu à une littérature récente (Dubnov, 2008). Elle utilise la théorie de l'information, l'apprentissage automatique implicite ainsi que les faillibilités des modèles appris à l'aide de leurs succès prédictifs.

Étant donné nos observations ci-dessus nous favorisons, pour le but de ce projet, les deux dernières approches d'apprentissage automatique et théorie d'information pour étudier les comportements anticipatifs musicaux. Les études cognitives autour de l'attente musicale suggèrent une approche pluraliste de ce phénomène, qui aborde directement la complexité des comportements musicaux sans une intellectualisation excessive du phénomène lui-même.

2.2. Modélisation anticipative

Nous questionnons la légitimité d'un modèle universel qui engendre le phénomène d'anticipation. Nous croyons que l'anticipation, comme définie auparavant, est un phénomène cognitif qui peut être explicitement considéré dans le processus de modélisation computationnelle et qui donne lieu aux comportements complexes musicaux pour les approches computationnelles dans la musique et le son. Pour cela, nous renversons l'approche de modélisation d'anticipation en modélisation anticipative avec le but de concevoir des *systèmes anticipatifs* musicaux. Un système anticipatif est défini comme un système qui contient un modèle prédictif de lui-même et/ou de son envi-

ronnement, qui permet des changements instantanés de son état actuel en accord avec sa prédiction afférente à un moment dans le futur (Rosen, 1985). Au contraire de la modélisation d'anticipation, la modélisation anticipative ne tente pas d'expliquer le phénomène universel d'anticipation, mais tente de fournir des cadres computationnels pour concevoir des modèles qui anticipent. Cette définition a donné naissance à une littérature sur le sujet avec une grande concentration sur des applications en robotique.

Une telle approche globale de modélisation nécessite une étude des outils computationnels pour arriver aux buts désirés. Le schéma global de la modélisation anticipative est celui de l'*apprentissage adaptatif* avec une interaction *en ligne* et en *temps réel* avec un environnement extérieur. Ces interactions ont pour effet de mettre à jour les *croiances* du système de son environnement dans les *mémoires* du système (ou ses *représentations mentales*) et de guider sa décision vers le futur. D'après (Butz *et al.*, 2003b), nous distinguons quatre approches schématiques de modélisation anticipative basé sur la conception des processus de décision markovien partiellement observable et démontrés en figure 1 :

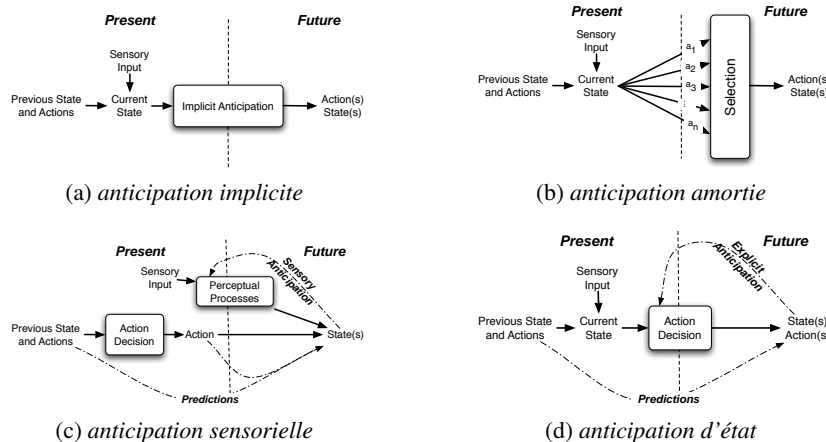


Figure 1 – Modèles et schéma de conception anticipatifs d'après Butz *et al.* (2003b)

Anticipation implicite : Des systèmes contenant une entrée sensorielle et sortie décisionnelle où la prédiction du système n'est pas explicitement considéré pour influencer les prises de décision à présent. Dans la plupart de ces modèles, la prise de décision est basé sur une prédiction implicite basé sur le passé et le présent du système (figure 1a).

Anticipation amortie¹ : Si la prédiction est considéré en tant que amortissement des actions possible, sans une considération explicite des états qui en résultent, il s'agit d'un système d'anticipation amortie (figure 1b).

Anticipation sensorielle : La prédiction des états future du système peut influencer les traitement sensorielles de l'entrée (ou module d'observation en terme com-

putationnelle). Cette conception computationnelle imite une fonction psychologique liée à l'*attention préparatoire* et assume la disposition d'un modèle prédictive explicite (figure 1c).

Anticipation d'état : Quand le comportement (artificielle) d'un système est influencer par la prédiction d'état complet du système dans le future, il s'agit de la forme plus explicite de modélisation anticipative (figure 1d).

Modélisation anticipative a des implications spécifiques qui sont détaillées et considérées tout au long du manuscrit et dans les sections qui suivent, et contiennent :

1) **Disponibilité de l'information :** Le premier facteur dans un modèle anticipatif est l'accès rapide aux informations musicales *pertinentes* dans le contexte de l'interaction, de l'apprentissage et de la décision comportementale du système. Ce thème est le sujet d'étude de la partie *II* de la thèse et de la section 3 de ce résumé.

2) **Apprentissage en ligne et interactif :** Si l'anticipation est l'attente dans l'action qui engendre la représentation mentale des stabilités d'un environnement extérieur, une modélisation anticipative implique une forme d'apprentissage automatique qui est *interactif* et aussi *en ligne*. L'apprentissage dans ce contexte est d'acquérir et de mettre à jour des représentations mentales (le sujet de la section 3) et d'apprendre les comportements anticipatifs issus des interactions et des échanges avec l'environnement. Ce point est étudié sous deux formes différentes dans les sections 4 et 5 de ce résumé (parties *III* et *IV* du manuscrit original).

3) **Modélisation et interaction multimodale :** Étant donné l'aspect concurrentiel des représentations mentales issues de l'attente, une modélisation anticipative est par nature *multimodale*. Les sections 4 et 5 représentent deux modèles différents en destinés à deux applications musicales différentes.

3. Géométrie de l'information musicale

La première considération pour un modèle anticipatif est la forme de la représentation, son stockage et l'accès à l'information dans un contexte interactif avec l'environnement. Cette considération est fort commune dans la conception des systèmes d'extraction de données musicales² et d'informatique musicale. Une question fréquente qui se pose pour chaque modélisation est donc : comment les informations sont présentées au système et quelle est leur pertinence. La littérature d'extraction de données musicales (ou *MIR*) se concentre la plupart du temps sur une construction des espaces métriques basée sur des mesures de *similarité* entre les descripteurs audio. La *métrique* utilisée dans ce domaine varie beaucoup selon les chercheurs et les applications, souvent sans aucune considération explicite pour la validité mathématique des espaces construits. Parmi les approches existantes, la *temporalité* des signaux musicaux est la plus approximative en termes de modélisation. Par exemple, dans la plupart des modèles proposés dans la littérature *MIR* (destinées à la classification entre autres) une

2. "Music Information Retrieval" (*MIR*)

série temporelle de données musicales est considérée comme un “sac de descripteurs”³ avant toute considération computationnelle, perdant l’aspect fort temporel des structures musicales. Par ailleurs, la théorie de l’information originale de Shannon (1948) fournit peu de réponses à notre préoccupation quant à la représentation et la fidélité d’informations musicales. Ces théories estiment *l’incertitude* des signaux. Pressing (1999) montre l’inconsistance des mesures élémentaires de la théorie de l’information pour exprimer la complexité des structures musicales. À cet égard, il est intéressant de remarquer que Warren Weaver dans son introduction à l’article phare de Shannon, constate que le travail de ce dernier n’aborde qu’une partie de la théorie du “traitement de l’information”. Cependant, il exige que le problème soit étudié sur trois fronts : *technique, sémantique, et influent*.

Ici, nous proposons un cadre mathématique et compréhensif basé sur la littérature récente de la *géométrie de l’information* (Amari *et al.*, 2000) et en particulier sur la géométrie de Bregman (Nielsen *et al.*, 2007a). Cette approche réunit plusieurs littératures souvent utilisées dans les domaines de MIR notamment le traitement des signaux musicaux, la théorie de l’information, l’apprentissage statistique et la géométrie différentielle. Après l’introduction des cadres mathématiques, nous redéfinissons les termes souvent utilisés dans la littérature de MIR et en particulier les notions de similarité, de divergence, les modèles etc. Nous poursuivrons notre proposition avec deux algorithmes basés sur les principes de la géométrie de l’information musicale, qui permettent d’extraire automatiquement les structures musicales d’un flux d’audio, et avec un algorithme rapide qui permet d’effectuer une recherche musicale dans une base de données. Tous les algorithmes et les méthodes proposés dans cette section sont destinés à une utilisation en ligne et en temps réel, et sont en conséquence idéaux pour une modélisation interactive.

3.1. Cadre théorique

Dans notre cadre théorique les données sont présentées comme des *points* dans une *variété Riemannienne*. La variété utilisée est une géométrie des structures statistique exponentielle ayant comme norme la mesure de l’information de Fisher, donnant lieu à une forme canonique et duale de connexion géométrique (Amari *et al.*, 2000). Pour compléter la géométrie informationnelle, nous introduisons la notion de *divergence* entre les points et les modèles statistiques, notion basée sur la famille canonique des divergences de Bregman (Bregman, 1967). Les divergences de Bregman ont les particularités suivantes :

1) Il existe une *bijection* entre la famille des probabilités exponentielles et les divergences de Bregman (Banerjee *et al.*, 2005).

2) La variété peut être définie directement par les divergences canoniques de Bregman donnant lieu à une variété plate et une forme de connexion affine (Zhang, 2004), et donc utilisable dans un cadre computationnel.

3. “Bag of features”.

3) Les divergences canoniques de Bregman engendrent des “distances” déjà introduites dans la littérature de reconnaissance des formes et de traitement du signal notamment Kullback-Leibler, Itakura-Saito, la norme Euclidienne etc.

Ayant deux points \mathbf{p} et \mathbf{q} dans la variété $\mathcal{X} \in \mathbb{R}^d$, la divergence de Bregman $D_F(\cdot|\cdot) : \mathcal{X} \rightarrow \mathbb{R}$ associée à une fonction strictement convexe et différentiable F est défini comme suite :

$$D_F(\mathbf{p}|\mathbf{q}) = F(\mathbf{p}) - F(\mathbf{q}) - \langle \nabla F(\mathbf{q}), \mathbf{p} - \mathbf{q} \rangle \quad [1]$$

où $\nabla F = \left[\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_d} \right]$ et $\langle \mathbf{p}, \mathbf{q} \rangle$ est un produit scalaire.

Cette fonction sur une variété informationnelle génère des éléments géométriques comme les *bulles de Bregman*, les *centroids de Bregman*, l'*information de Bregman*, les *bissecteurs*, *géodésique*, le plus important étant une dualité structurelle entre l'espace des paramètres d'attente et naturelle à l'aide d'une transformation de Lègènder (définitions détaillées dans (Cont *et al.*, 2011; Cont, 2008b)).

Dans notre établissement d'une géométrie des signaux musicaux, nous assumons que les signaux arrivent d'une manière incrémentale (en ligne/temps réel) et sont soumis à une analyse fréquentielle sur les fenêtres d'analyse. Dans cette approche, chaque point sur la variété d'information est une distribution fréquentielle qui peut être convertie sans perte de généralisation à une forme de distribution Multinomiale.

Suivant Dubnov (2008), nous faisons la distinction entre deux sources d'information musicale L'un liée aux *données*, et l'autre liée aux *modèles* qui établissent la structure géométrique, décrite ci-dessous. Dans chaque approche, nous cherchons à établir des méthodes pour évaluer la pertinence de l'information provenant de l'aspect *influent* ou relationnel de l'information.

3.1.1. Géométrie des données

L'ensemble des données (ou points dans notre géométrie) engendre une information structurelle continue si le processus à partir duquel les données sont générées peut être considéré comme stationnaire. Dubnov (2008) propose une mesure du *taux d'information* basée sur l'information mutuelle partagée entre les données musicales du passé et le présent, et prouve que c'est équivalent à la mesure de “Spectral Flatness” souvent utilisée dans le traitement du signal. Dubnov *et al.* (2006) montre également que cette mesure correspond à l'expérience des auditeurs ainsi qu'aux structures des sons naturels. Nous montrons mathématiquement⁴ que la mesure de Dubnov est automatiquement déduite dans notre cadre de géométrie de l'information en tant que *Information de Bregman* correspondant à une divergence d'Itakura-Saito sur la puissance du spectre.

4. Théorème 4.3 du (Cont, 2008b)

3.1.2. Géométrie des modèles

Pour atteindre la structure réaliste des informations musicales, nous abordons le problème de stationnarité en introduisant le concept des *modèles*. Un modèle est défini comme un morceau audio qui subit une stationnarité (ou quasi-stationnarité) interne. En d'autres mots, un sous-ensemble des points $\mathbf{x}_i, \dots, \mathbf{x}_{i+k}$ constitue un modèle θ_j s'il peut être considéré stationnaire sous un modèle probabiliste $P(\mathbf{x}_i, \dots, \mathbf{x}_{i+k} | \theta_j)$. Plus précisément, un modèle θ_i dans une variété duale de Bregman est défini comme un sous-ensemble des points temporellement continus qui forment une *sphère de Bregman* $B(\boldsymbol{\mu}_i, R_i)$ avec centre $\boldsymbol{\mu} \in \mathbb{R}^d$ et radius R_i , définie comme

$$B_r(\boldsymbol{\mu}_k, R_k) = \{\mathbf{x} \in \mathcal{X} : D_F(\mathbf{x}, \boldsymbol{\mu}_k) \leq R_k\}.$$

Nous proposons un cadre théorique dans lequel les modèles peuvent être automatiquement extraits par des signaux arrivant dans un système temps réel. Ceci nécessite en particulier l'existence d'un espace métrique pour comparer la pertinence de l'information entre les modèles et les signaux. Nous basons notre espace sur une définition de la *similarité* dans laquelle la quantité de similarité n'est pas obtenue par une analyse des contenus mais par *la différence* entre les modèles. De plus, un espace métrique est caractérisé par sa symétrie ainsi que par l'existence d'inégalité triangulaire de la distance considérée. Pour atteindre cet espace métrique, nous proposons la divergence Bregman liée à la géométrie sous considération (qui pour nous est le fameux Kullback-Leibler comme divergence associée à une variété Multinomiale) avec les considérations suivantes : les divergences de Bregman ne sont pas a priori symétriques. D'après Nielsen *et al.* (2007b) nous proposons une symétrisation définie comme un problème d'optimisation, à la fois pour former les sphères de Bregman et aussi pour calculer la divergence. Également, pour approcher l'inégalité triangulaire nous montrons mathématiquement qu'une divergence de Bregman approche l'inégalité dans le cadre d'un problème de maximum de vraisemblance, ce qui est le cas pour nous à cause de la dualité structurelle des variétés dans laquelle le problème d'optimisation des sphères symétrique est équivalent à un problème de maximum de vraisemblance convexe (voir (Cont *et al.*, 2011) pour les détails).

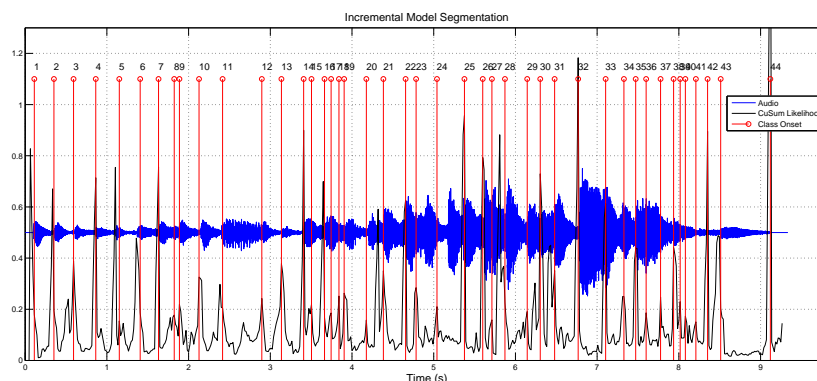
Suite à ces considérations, les *modèles* sur une variété continue peuvent être formés incrémentalement à l'aide d'un algorithme de détection de changement telle que *CUSUM* (Basseville *et al.*, 1993) avec un simple seuil d'information (défini comme *gain d'information*) pour segmenter un flux audio en morceaux continus et quasi stationnaires. La figure 2 montre un exemple de segmentation non-supervisé où les 7 premiers modèles indexés coïncident avec les premières notes dans la partition musicale.

3.2. Audio Oracle : Découverte automatique des structures musicales

Pour beaucoup d'applications, il est souhaitable et nécessaire d'avoir accès aux structures temporelles des flux musicaux comme la régularité, les répétitions et les



(a) La partition d'extrait audio



(b) Résultats de segmentation avec $R = 0.15$

Figure 2 – Segmentation vers les *modèles* du premier thème de la Sonate N.1 de Beethoven, extrait d'un enregistrement de Friedrich Gulda (1950-58)

recombinaisons d'éléments intérieurs. Suite à notre cadre de géométrie de l'information, nous proposons un algorithme nommé *Audio Oracle* pour l'extraction automatique et en ligne de la structure temporelle à long terme. *Audio Oracle* apprend une structure de graphe dont chaque sommet correspond à une *donnée* (point) ou à un *modèle* (donc deux cadres d'utilisation) dans leur ordre temporel d'arrivée, et les arêtes correspondent aux liaisons structurelles. La figure 3 montre des diagrammes schématiques d'un *Audio Oracle* imaginaire dont chaque boîtier correspond à un bout de son avec des arêtes montrant la régularité structurelle du flux. En particulière, figure 3b montre les liaisons temporelle de structure linéaire de la figure 3a dans une visualisation étendu en formant un ensemble d'arbre temporelle et en liant les régularités structurelles temporelles. *Audio Oracle* est l'extension d'un algorithme d'origine symbolique nommé *Factor Oracle* (Allauzen *et al.*, 1999) destiné à l'analyse de texte et des séquences ADN. La particularité d'*Audio Oracle* est son apprentissage incrémental ainsi que sa complexité linéaire dans le temps et l'espace. En plus, les arêtes

uniques de structure (appris et construits en ligne) correspondent aux suffixes ou préfixes *la plus longs* observés dans le passé.

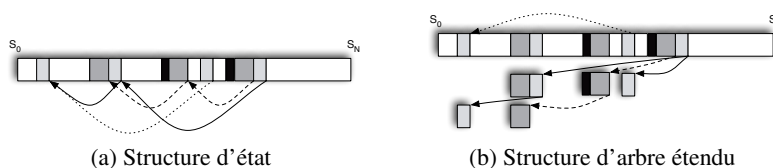


Figure 3 – Structure schématique d'un *Audio Oracle*.

La définition d'*Audio Oracle* dans le cadre d'une variété informationnelle décrite auparavant permet l'accès aux structures dans un espace métrique basé sur la similarité et assure la continuité des structures découvertes sur le long terme. Nous abordons la découverte des structures audio sur *les données* et *les modèles* dont les premiers adressent les microstructures dans un flux sonore (comme les sons naturels) et les seconds les structures long terme (comme dans les morceaux de musique). Figure 4 montre un *Audio Oracle* appris sur un son naturel d'oiseau qui, énoncé deux fois par l'animal et appris sur une analyse de coefficients fréquentielle de Mel (MFCC). Chaque sommet correspond à une trame d'analyse et les arêtes révèlent la régularité structurelle. Figure 5 montre un exemple d'*Audio Oracle* sur un morceau entier de

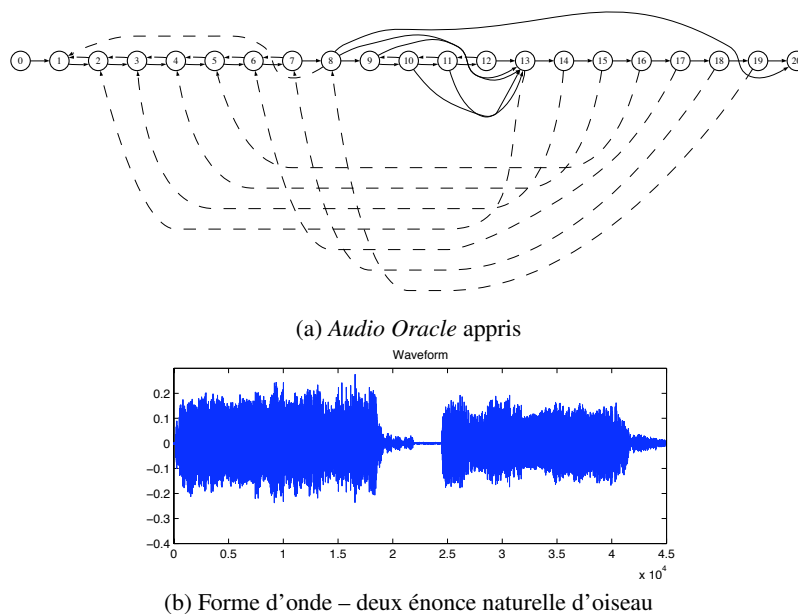


Figure 4 – Exemple d'*Audio Oracle* basé sur les *données*.

musique (Mouvement 3 de Sonate 1 pour Piano de Beethoven⁵), basé sur *des modèles*. Dans cet exemple, une analyse structurelle venant de la partition (de forme A, B, ...) est superposée aux résultats. La figure du milieu montre le rapport temporel des arêtes (envoi vers le temps passé) et la figure en bas montre la taille maximum de rappel associé à chaque rappel.

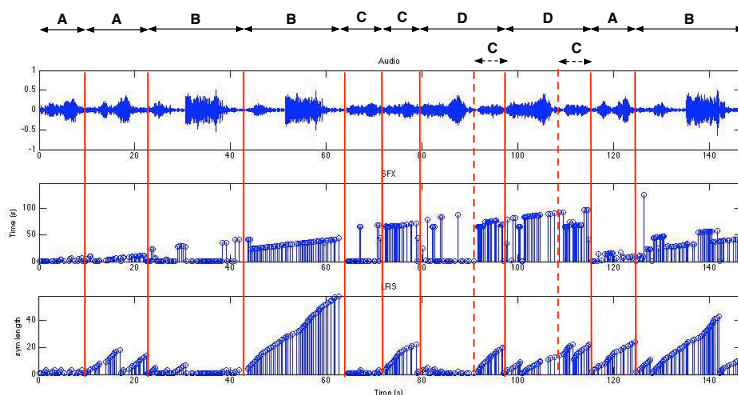
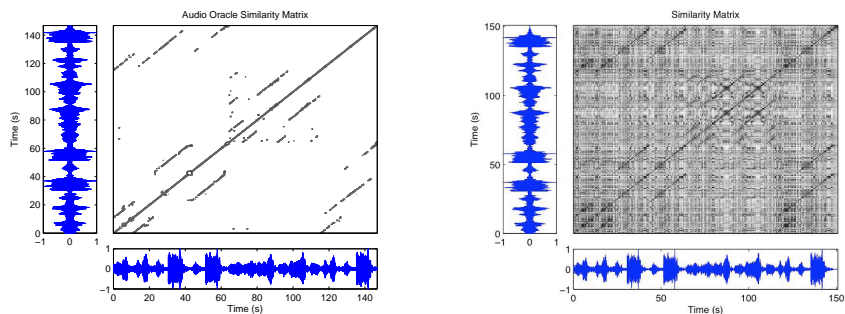


Figure 5 – Structure d'Audio Oracle : Forme d'onde avec analyse structurelle venant de la partition musicale (haut), rappel temporelle des arêtes (milieu), et la taille des rappels (bas)

La structure d'état d'Audio Oracle de la figure 5 contient 440 états appris sur un flux audio de 9500 trames et révèle la structure répétitive du morceau en question. Elle peut également être visualisée comme une matrice de similarité montrant le lien direct entre les trames comme démontré en figure 6a. Ceci est une matrice entre les 440 états où les valeurs correspond à la mesure de similarité entre les modèles sélectionnés. Pour comparaison, figure 6b montre une matrice de similarité classique d'après (Foote *et al.*, 2003) (la référence en matière de calcul de similarité multimédia) entre tous les 9500 trames d'audio et en utilisant une distance Euclidean sur une représentation MFCC des trames. À noter que la matrice de similarité dans figure 6b est sur une base très exhaustive (9500×9500 trames) et son calcul n'est possible qu'en temps différé. Celui de figure 6a est obtenu en temps réel et d'une manière incrémentale et contient une matrice parcimonieuse de 440×440 .

Mis à part l'utilité d'Audio Oracle pour la découverte automatique des structures audio, les structures apprises donnent un accès rapide aux sous-parties audio basée sur leur intersimilarité. Cette caractéristique est particulièrement utile pour considération dans les applications d'extraction et de recherche audio, ainsi qu'en apprentissage automatique des agents.

5. Interprété par Friedrich Gulda, Sony DECCA, date d'enregistrement 1950 – 1958.



(a) Matrice de similarité basé sur oracle d'audio de la figure 5 (b) Matrice de similarité audio d'après (Footo *et al.*, 2003)

Figure 6 – Comparaison des matrices de similarités sur l'exemple de la figure 5, produit par *Audio Oracle* (gauche) et calcul de similarité traditionnelle.

3.3. Guidage : Recherche rapide et assemblage des données

Ayant un accès rapide aux structures audio, il est tout à fait légitime de demander *comment* récupérer l'information pertinente dans un flux audio ou même dans une base de donnée musicale par une requête externe. Pour aborder cette question, nous introduisons *Guidage*, un algorithme de recherche audio basé sur des requêtes sur les bases de donnée audio et qui est capable de rassembler les bouts de chaque fichier audio en les concaténant pour reconstruire un nouveau flux audio qui ressemble à la requête en question. *Guidage* est basé sur un simple schéma de programmation dynamique et destinée à une utilisation en temps réel. La particularité de *Guidage* dans son usage est le fait qu'il est capable de rassembler des bouts d'audio dont la combinaison est inexistante dans la base de donnée pour atteindre une requête demandée. Dans cette perspective, *Guidage* aborde le problème de la sélection d'unité qui est le cœur de beaucoup des systèmes comme la synthèse concaténative et l'extraction des données. Sur le plan algorithmique, la particularité de *Guidage* est son non-usage des trames audio et son exploitation en navigation directe sur les structures temporelles des données en utilisant des *Audio Oracles*. Dans ce sens, *Guidage* utilise des structures *Audio Oracles* comme métadonnée audio. Aussi nous montrons que *Guidage* est *scalable* en terme de calcul et stockage des données.

Ayant une requête audio et une cible, *Guidage* cherche des formes continues dans *Audio Oracle* de la cible qui contient des parties plus longue de requête. Le résultat de *Guidage* est en effet un ensemble d'arbres de sommets, dont chaque sommet correspond à un état d'oracle sur la cible utilisée pendant l'assemblage et la synthèse sonore. Comme avant, *Guidage* peut tourner sur les *modèles* ainsi que sur les *données*. Figure 7 montre un arbre résultant du *Guidage* tourné sur le premier thème de la première sonate de Beethoven (figure 2) comme requête et la sonate entière comme cible. Cet arbre montre plusieurs chemins de reconstructions possibles dont le chemin

original (en gris) ainsi que les réapparitions du thème en deux reprises au long de la pièce. Parmi tous ces chemins, l'un correspond à une reconstruction avec une similarité maximum à la séquence temporelle de la requête, qui peut être resynthétisée à une séquence audio. La figure 8 montre des résultats sonores obtenus lors d'une séance de *Guidage* sur les *données* d'une base de donnée de boucles percussives, dont la figure 8b montre la forme d'onde de la requête, la figure 8a la forme d'onde de la cible avec des sous-parties utilisées pendant la concaténation en rouge (gris), et la figure 8c le résultat de la synthèse concaténative aligné sur la requête⁶.

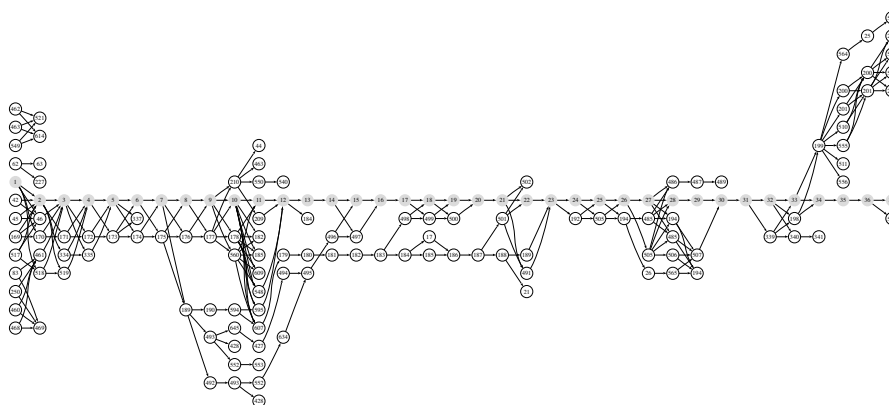


Figure 7 – Exemple d'arbre résultant de *Guidage* sur les *Audio Oracle* basé sur les *modèles* – la requête étant la première thème du 1re Sonate pour Piano de Beethoven (Figure 2), et la cible l'enregistrement entier de la Sonate. Lecture de gauche, le chemin grise étant le chemin de reconstruction optimale.

4. Apprentissage actif des comportements anticipatif

Ayant accès aux structures d'information musicale, nous nous concentrons autour de l'apprentissage automatique des comportements anticipatifs décrits comme un problème d'apprentissage d'une stratégie de planification pour des prises de décision futures dans un environnement. Pour atteindre les comportements réactifs et proactifs décisionnels, l'apprentissage doit être interactif et adaptatif. Dans la littérature, le domaine de la recherche qui approche ce but est celui de la modélisation stochastique de la musique, celui de l'improvisation automatique et d'imitation du style. Parmi ces dernières, nous sommes particulièrement intéressés par les approches *agnostiques* dans lesquelles aucune connaissance musicologique n'est demandée ni a priori ni intégré. Donc le système apprend et accumule toute sa connaissance par une

6. Pour voir et entendre plus de résultat : <http://cosmal.ucsd.edu/arshia/index.php?n=Main.Guidage>

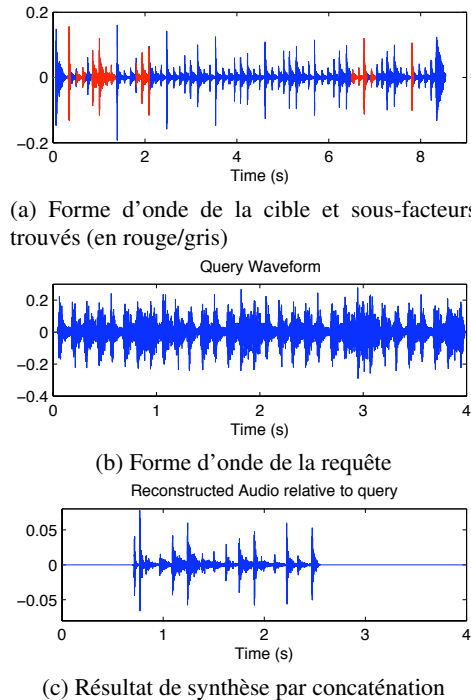


Figure 8 – Exemple d'extraction de donnée avec la synthèse de *Guidage* sur les données

interaction implicite ou explicite avec un environnement extérieur. Ces systèmes diffèrent en grande partie par les approches variées pour modéliser la *mémoire* musicale du système et différentes approches de planifications pendant la génération. Les deux problèmes impliquent l'utilisation d'apprentissage automatique, la première pour découvrir les structures existantes dans une base de donnée afin de les réutiliser pendant la génération, et la deuxième pour pouvoir décider *les meilleures actions* pendant la génération automatique. Ces deux problèmes sont parfois adressés implicitement dans la conception des systèmes. L'approche dominante aujourd'hui pour les modèles de mémoire est l'utilisation des variantes des chaînes de Markov. Le problème de planification en revanche est souvent indirectement abordé, soit en utilisant *une marche aléatoire* sur les modèles appris contrôlés par un musicien ou un utilisateur (Dubnov *et al.*, 1998; Pachet, 2002), ou par un échantillonnage des formes (Pearce *et al.*, 2004). Un problème souvent rencontré dans les systèmes existants est la difficulté de gérer les données multidimensionnelles (la caractéristique de base des signaux musicaux), et d'atteindre la complexité en long terme observée dans une pièce de musique à partir de la génération automatique.

Nous introduisons un cadre d'apprentissage automatique qui adresse explicitement les deux problèmes décrits ci-dessus ainsi qu'une partie de la complexité des données musicales. Notre approche est basée sur *apprentissage par renforcement* et *apprentissage actif*. Apprentissage par renforcement est un manière direct d'aborder le problème d'apprentissage pour atteindre des comportement spécifique à un environnement. L'apprenti est appelé *un agent*, qui interagit avec *l'environnement* désigné comme tout ce qui est extérieur. L'agent et l'environnement interagissent continuellement, l'agent entreprend des actions et l'environnement répond à ces actions en représentant des nouvelles situations à l'agent. L'environnement donne aussi lieu à des *récompenses*, des signaux numériques que l'agent essaie à maximiser à chaque interaction. Dans ce sens-là, l'agent est interactif et s'adapte aux situations comportementales de chaque cycle d'interaction. Un problème d'apprentissage par renforcement peut être formalisé avec trois signaux en aller-retour entre le système et son environnement : un pour représenter les *actions* prises par l'agent, un autre signal représentant les *états* courants du système, et un troisième signal représentant les cible comportementale du système (*récompenses*) (Sutton *et al.*, 1998). L'apprentissage est fait en simulant des épisodes d'actions et en mettant à jour, au travers d'expériences itérées, un comportement décisionnel appelé *stratégie*. Ce cadre d'apprentissage aborde le problème d'apprentissage automatique dans un environnement changeant, mais n'aide pas l'agent à *explorer activement* son environnement et exige souvent un temps suffisamment long pour atteindre des états *pertinents* dans chaque interaction. Ce problème est une préoccupation du domaine d'apprentissage actif où l'agent à l'aide d'un *oracle* choisit lui-même les états pertinents pour une mise-à-jour automatique. Dans ce cadre, les récompenses sont remplacées par les *guides* qui non seulement fournissent au système des signaux de récompenses mais aussi dirigent apprentissage vers les états pertinents dans la mémoire du système.

Notre système d'apprentissage constitue deux comportements interactifs différents démontrés en figure 9. Dans la première le système est en interaction avec un environnement à l'extérieur qui peut être un musicien, une partition musicale observée en séquences etc. Dans la deuxième le système est en interaction avec lui-même et donc en écoute introspective. Ce deuxième mode est utilisé pendant les générations.

L'architecture du système ressemble à celle du *dyna* (Sutton, 1991), augmenté par une architecture multi-agent collaborative et compétitive. Chaque agent dans cette conception contient un aspect spécifique de représentation mentale de l'environnement, un apprenti, et un preneur de décision. L'architecture globale a trois composants principaux qui sont :

1) *Modèles de mémoire* : qui constituent des représentations par état de la connaissance jusqu'à présent ainsi que les actions. Nous utilisons les *Audio Oracle* présentés précédemment comme des modèles de mémoire du système.

2) *Sélection Active* : responsable des transactions entre l'environnement et les agents génératifs, et *guidage* des apprentis vers les états pertinents. Naturellement, nous adoptons l'algorithme *Guidage* avec quelques modifications pour ce composant du système.

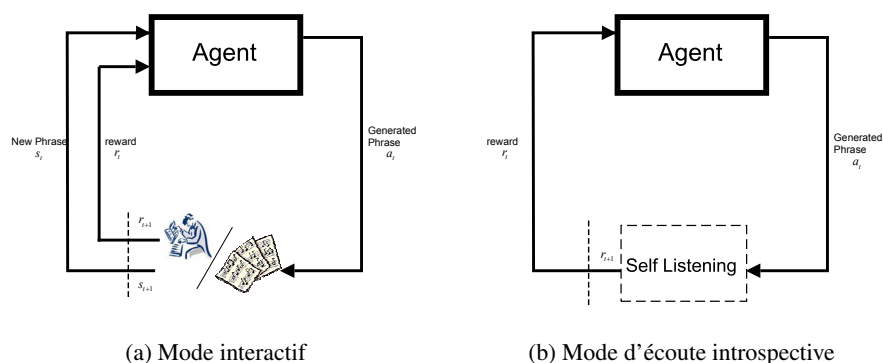


Figure 9 – Les modes d'opérations d'apprentissage actif

3) *Apprenti Anticipatif*: responsable d'apprentissage des *stratégies* de chacun des agents. Cet apprentissage est anticipatif car nous sommes intéressés d'apprendre des valeurs anticipatoires associées à chaque pair état/action dans les modèles et pas nécessairement à leur valeur instantanée prédictive.

Dans chaque cycle d'interaction il y a deux types d'apprentissage : un pour apprendre et mettre à jour des modèles de mémoires (*Audio Oracles* de chaque agent), et apprendre les valeurs stratégiques d'espace des paires état/action. Ce deuxième ressemble à un simple algorithme de *Q-Learning* (Sutton *et al.*, 1998), avec un extensions vers un apprentissage multi-agents compétitif et collaboratif, basé sur la mémoire, et aussi guidé par une *sélection active*. À chaque instant de l'apprentissage (ou génération) le comportement d'ensemble du système est le résultat de la collaboration entre les agents (dont chacun responsable d'un aspect d'information environnementale) en suivant *un agent de comportement* choisi au début de chaque épisode à partir d'un processus de sélection compétitive entre agents. La compétition entre agents est gérée par une distribution de Boltzmann sur les valeurs apprises précédemment, et la collaboration influence les autres agents par l'agent de comportement à l'aide d'une méthode d'*Importance Sampling*.

Le but de cette conception est de montrer qu'un apprentissage anticipatif et interactif est capable de capturer le comportement anticipatif long terme et aussi de générer de la structure complexe en présence de peu de données d'apprentissage. Pour cela, figure 10 montre la partition résultante du système dans un environnement multiagent destiné à l'improvisation automatique et l'imitation du style musical. L'apprentissage a eu lieu sur les données extraites de la partition de l'*Invention à deux voix n°3* de J. S. Bach, en mode interactif (figure 9a), et la génération en mode d'écoute introspective (figure 9b). Un oeil de musicien peut identifier des formes musicales et temporels long terme (et basique) implicitement appris sur la musique en question⁷.

7. Pour entendre des résultats : <http://cosmal.ucsd.edu/arshia/index.php?n=Main.Improvisation>

Improvisation Session after learning on Invention No.3 by J.S.Bach

Figure 10 – Résultat d’une session d’imitation de style après une session d’apprentissage actif sur la troisième *Invention à deux voix* de J.S.Bach

5. Synchronisation Anticipative

Un des sujets importants pour lequel anticipation vient aider la modélisation est la *temporalité* comme dans un problème de synchronisation temps réel. La synchronisation temps réel est aussi un des soucis principal des systèmes interactifs de la musique informatique (et même de l’art numérique). Dans un contexte computationnel, le problème de synchronisation en musique informatique est de trouver la position du musicien lors qu’il joue une partition écrite dans la partition pour que l’ordinateur puisse se comporter comme un deuxième musicien muni de sa propre partition numérique. Les systèmes abordant cette problématique sont classiquement appelés les *suiveur de partition*. Le suivi de partition se trouve au cœur de la musique électronique temps réel lors de l’écriture de la pièce ainsi que sa réalisation. Dans ce sens, suivi de partition constitue un élément principal pour prendre en considération *l’écriture du temps et de l’interaction* dans la musique informatique en temps réel.

Lors d’une session de concert ou de répétition entre deux musiciens humains, un élément clé de synchronisation est l’anticipation de chaque musicien sur son propre comportement ainsi que sur le comportement futur de son collègue. Pour une synchronisation automatique entre un environnement extérieur (musiciens) et la machine, nous proposons un modèle anticipatif multi-agents capable d’imiter la même anticipation temporelle entre les humains. Le modèle proposé est basé sur des chaînes

hybrides Markov/Semi-Markov cachées. Les chaînes hybrides sont construites à partir d'une partition musicale et contiennent plusieurs sémantiques temporelles souvent utilisées dans l'écriture musicale. La synchronisation donc se résume à trouver l'état le plus probable dans la chaîne à partir d'une observation audio en temps réel. Ce problème est un problème classique d'inférence. La particularité de notre approche réside dans une conception basée sur deux agents audio et tempo qui tournent en parallèle et qui sont *couplés*. L'agent du tempo existe sur l'échelle d'événement et est basé sur le modèle cognitif des structures métriques musicales de Large *et al.* (1999). Ce modèle est développé de plus dans notre conception pour fournir les paramètres de tempo prédictif et les fonctions probabilistes de survie en temps réel à l'algorithme d'inférence. L'agent audio en revanche vit sur l'échelle continue. Ce schéma est donc un modèle anticipatif sensoriel car les probabilités de vraisemblance sont influencés dynamiquement par le tempo prédit, et en retour l'agent de tempo est directement affecté par la décision obtenue par l'agent d'audio. La figure 11 montre l'architecture globale du système.

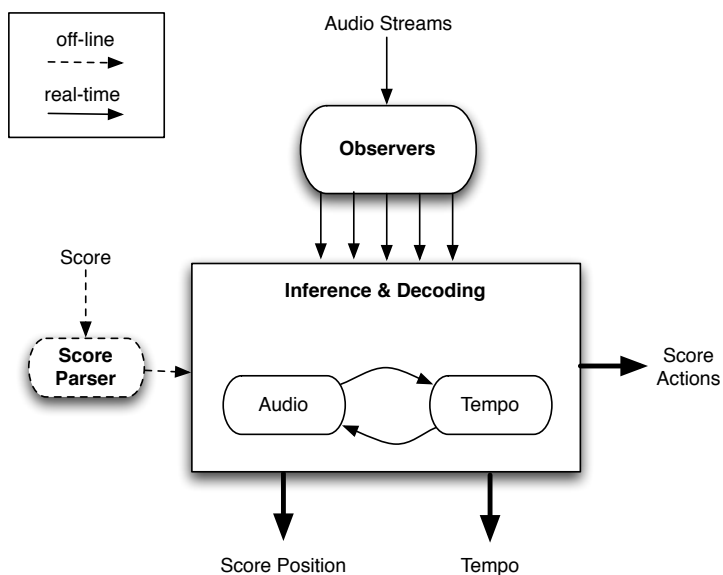
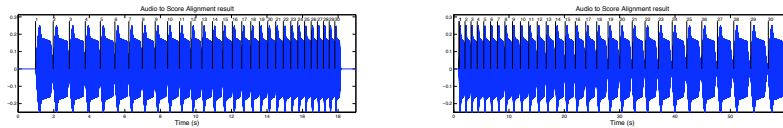


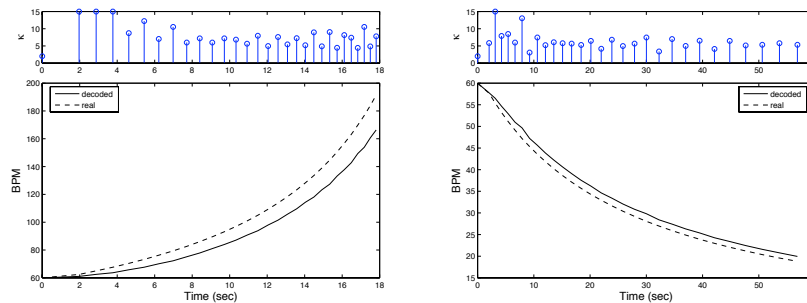
Figure 11 – Architecture générale de synchronisation anticipatif

L'agent du tempo est basé sur un modèle adaptatif de filtre de Kalman étendu. Dans cette conception, les filtres sont la version circulaire des Gaussiennes (ou distribution *von mises*) et leurs paramètres de variance (κ) sont adaptés à l'erreur de prédiction à chaque instant d'activation. Le résultat de cet agent à chaque activation est donc le tempo prédit et une estimation de survie des événements futur utilisés directement dans la formulation d'inférence. La figure 12 montre deux exemples sonores de la même partition rythmiquement homogène ("noire" pour toutes les notes) mais accé-

léré pour un et ralenti pour l'autre, et les résultats du décodage de tempo (en BPM) pour les deux.



(a) Forme d'onde d'audio synthétique à partir d'une partition homogène accéléré (gauche) et ralenti (droit)



(b) Tempo estimée et tempo réel pour l'accélération et le ralentissement

Figure 12 – Exemple du décodage de tempo pour le changement continu du tempo

Le système proposé est entièrement polyphonique et testé sur des bases de données de morceaux de musique pré-alignés pour évaluer la précision de synchronisation en temps réel. Les résultats d'évaluation donnent une précision totale de 91.49% sur un ensemble de plus de 30 minutes de musique variée en style et instrumentation.

5.1. Vers une langage synchrone temps réel pour la composition musicale

L'application d'une telle système de synchronisation entre un musicien et une partition musicale en temps réel, peut s'éteindre de l'*accompagnement automatique* (où le musicien joue la partie solo et l'ordinateur la partie orchestre) jusqu'à la *composition musicale* avec des dispositifs informatique temps réel (comme des effets et synthèses sonores et spatiale). Dans un utilisation créatif telle que la composition musicale, les éléments d'accompagnement électronique ne sont pas statique et peuvent engendrer des programmes dynamique à la fois en terme de résultats finale, mais aussi en terme de temporalité à l'image d'une écriture musicale classique. Ces éléments sont aussi par nature concurrentielle à l'image de la polyphonie musicale.

L'idée d'une écriture musicale en informatique au temps de composition, et l'évaluer en temps réel lors de la performance musicale, est d'une manière étonnante proche aux paradigmes de programmation réactif et synchrone (Halbwachs, 1993). Une par-

tion musicale peut être vu comme un programme déterministe, évalué en temps réel (performance) malgré toutes les variabilité temporelle et événementielle d'une interprétation musicale. Ce qui compte et la synchronisation des événements comme indiqués dans la partition et leurs ordonnance temporelle.

L'accès au temps anticipatif décrit dans notre système de synchronisation fournit une possibilité de pouvoir définir l'ordonnance des programme électronique relatif à un horloge musicale (le *tempo*) destinée à une évaluation temps réel et couplé avec le musicien. L'intégration de la système de synchronisation et reconnaissance décrit ci-dessus avec un langage synchrone pour la composition musicale est implémenté sous le nom d'Antescofo⁸. Figure 13 montre une simple partition d'Antescofo, avec la partie instrumentale (destinée à être jouée par un instrumentiste) avec des programmes synchrones et concurrents comme des boîtes. À l'intérieur de chaque bloque de programme est une série de message ordonné en temps musicale, qui constitue un processus de musique électronique en temps réel. La longueur de chaque bloque est relatif au tempo de la partie instrumentale et le scheduler d'Antescofo prends en charge la synchronisation du contenu avec le musicien en temps réel.

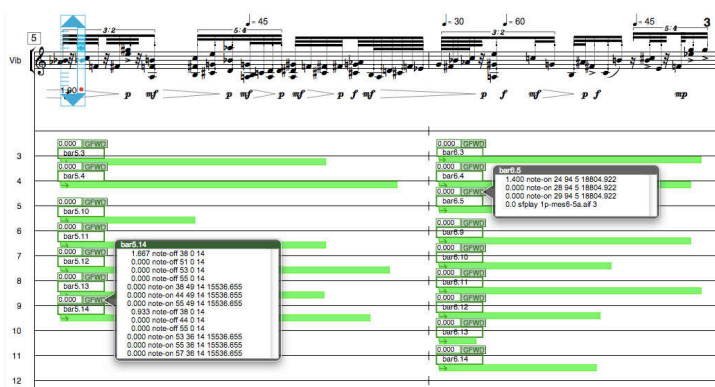


Figure 13 – Une partition d'Antescofo montrant la partie instrumentale et l'accompagnement (partie électronique temps réel) comme boîtes de programmes concurrents.

Antescofo est développé pour les environnements de programmations graphiques *MaxMSP* (Cycling74, 2011) et *PureData* (Puckette, 1997), et depuis a été utilisé dans plusieurs concerts de musique contemporaine, jouant la partie électronique des œuvres mixte de musique contemporaine, et notamment avec l'orchestre Philharmonique de Los Angeles, l'orchestre Écossais de la BBC, Ensemble Intercontemporain, Orchestre de Radio France, et Opéra La Scalla parmi d'autres. Antescofo est aujourd'hui le système de base du suivi de partition et écriture synchrone temps réel à l'IRCAM pour

8. <http://repmus.ircam.fr/antescofo>

plusieurs pièces existant comme celles de Macro Stroppa, Pierre Boulez, et Philippe Manoury, avec de nombreux projets de productions à venir⁹.

6. Conclusion

Notre perception est un processus temps réel par nature et appelle notre puissance d'anticipation à tous moments pour prendre des décisions. Ce travail a comme but d'introduire des cadres de modélisations pour introduire ce phénomène cognitif dans la boucle de la conception informatique. Nous distinguons entre prédiction, l'attente et anticipation et nous avons introduit quelques cadres de modélisation en section 2. Sections 3 à 5 abordent plusieurs problèmes complexes en traitement du signal et apprentissage automatique sous l'oeil de systèmes anticipatifs.

Le cadre de la géométrie de l'information introduit en section 3 aborde des problèmes complexes tels que la qualification de l'information continue et en temps réel pour la reconnaissance automatique des structures régulières et requête de l'information. La géométrie de l'information musicale est aujourd'hui un champ de recherche à part entière qui pourrait résoudre la divergence actuelle entre les aspects symboliques et signaux de l'information continue.

Section 4 propose un cadre d'apprentissage actif en interaction avec un environnement extérieur pour apprendre des représentations "mentales" qui constitue la croyance anticipative d'un système numérique. Les systèmes similaires à celui exposé prennent en générale une base énorme d'information pour apprendre des modèles et sont en particulier basée sur la notion de prédiction. Nous montrons comment en introduisant le concept d'anticipation dans la conception, le temps et espace d'interaction et apprentissage pourrait être réduit.

Le système *Antescofo* introduit en section 5 constitue aujourd'hui une référence pour la synchronisation temps réel et programmation synchrone des pièces mixte pour électronique et instrument de musique. Le succès relatif d'*Antescofo* dans sa communauté est lié à une modélisation anticipative, proche à la logique humaine, qui le permet non seulement de raisonner en temps réel, mais aussi de l'écrire.

Cet article et aussi son manuscrit original (Cont, 2008b), ne sont qu'un début pour la modélisation des systèmes temps réel pour la musique. Les questions posées et abordées, malgré leurs importances dans leurs domaines respectifs, représentent une validation de principe de modélisation anticipative et une étude de faisabilité sur des données réelle et complexe. Nous espérons continuer sur cette voie, et espérons également de pouvoir montrer que la musique constitue un domaine d'étude autosuffisant pour étudier la complexité et la cognition temporelle.

9. Pour suivre les événements musicaux liés à *Antescofo* : <http://repmus.ircam.fr/Antescofo/events>

7. Bibliographie

- Allauzen C., Crochemore M., Raffinot M., « Factor Oracle : A New Structure for Pattern Matching », *Conference on Current Trends in Theory and Practice of Informatics*, p. 295-310, 1999.
- Amari S., Nagaoka H., *Methods of information geometry*, vol. 191, Oxford University Press, 2000. Translations of mathematical monographs.
- Banerjee A., Merugu S., Dhillon I. S., Ghosh J., « Clustering with Bregman Divergences », *Journal of Machine Learning Research*, vol. 6, p. 1705-1749, 2005.
- Basseville M., Nikiforov I. V., *Detection of abrupt changes : theory and application*, Prentice-Hall, Inc., NJ, USA, 1993.
- Bregman L. M., « The Relaxation Method of Finding the Common Point of Convex Sets and Its Application to the Solution of Problems in Convex Programming », *USSR Computational Mathematics and Mathematical Physics*, vol. 7, p. 200-217, 1967.
- Butz M., Sigaud O., Gérard P., « Anticipatory Behavior : Exploiting Knowledge About the Future to Improve Current Behavior. », *Anticipatory Behavior in Adaptive Learning Systems*, Springer-Verlag, chapter 1, p. 1-10, 2003a.
- Butz M., Sigaud O., Gérard P., *Internal Models and Anticipations in Adaptive Learning Systems : Foundations, Theories, and Systems*, n° 2684 in LNCS, Springer-Verlag, chapter 6, p. 86-109, 2003b.
- Cont A., « ANTESCOFO : Anticipatory Synchronization and Control of Interactive Parameters in Computer Music », *Proceedings of International Computer Music Conference (ICMC)*, Belfast, August, 2008a.
- Cont A., *Modeling Musical Anticipation : From the time of music to the music of time*, PhD thesis, University of Paris 6 and University of California in San Diego, October, 2008b.
- Cont A., « A coupled duration-focused architecture for realtime music to score alignment », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, n° 6, p. 974-987, 2010.
- Cont A., Assayag G., Dubnov S., Bloch G., *The Structure of Style*, Springer Verlag, chapter Interaction with Machine Improvisation, p. 219-246, August, 2010.
- Cont A., Dubnov S., Assayag G., « Anticipatory Model of Musical Style Imitation using Collaborative and Competitive Reinforcement Learning », in B. M.V., S. O., P. G., B. G. (eds), *Anticipatory Behavior in Adaptive Learning Systems*, vol. 4520 of *Lecture Notes in Computer Science / Artificial Intelligence (LNAI)*, Springer Verlag, Berlin, p. 285-306, 2007.
- Cont A., Dubnov S., Assayag G., « On the Information Geometry of Audio Streams with Applications to Similarity Computing », *IEEE Transactions on Audio, Speech and Language Processing*, January, 2011.
- Cycling74, « Max/MSP realtime graphical programming environment », 2011.
- Dubnov S., « Unified View of Prediction and Repetition Structure in Audio Signals With Application to Interest Point Detection », *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, n° 2, p. 327-337, 2008.
- Dubnov S., Assayag G., El-Yaniv R., « Universal Classification applied to Musical Sequences », *Proc. of ICMC*, Michigan, p. 322-340, 1998.

- Dubnov S., McAdams S., Reynolds R., « Structural and affective aspects of music from statistical audio signal analysis », *Journal of the American Society for Information Science and Technology*, vol. 57, n° 11, p. 1526-1536, 2006. Special Topic Section on Computational Analysis of Style.
- Foote J., Cooper M., « Media Segmentation using SelfSimilarity Decomposition », *Proceedings of SPIE Storage and Retrieval for Multimedia Databases*, vol. 5021, p. 167-175, 2003.
- Halbwachs N., *Synchronous Programming of Reactive Systems.*, Kluwer Academics, 1993.
- Huron D., *Sweet Anticipation : Music and the Psychology of Expectation.* MIT Press, 2006.
- Large E. W., Jones M. R., « Dynamics of Attending : How People Track Time-Varying Events », *Psychological Review*, vol. 106, n° 1, p. 119-159, 1999.
- Loui P., Wessel D., Kam C. H., « Acquiring New Musical Grammars : a Statistical Learning Approach », *28th Annual Conference of the Cognitive Science Society*, Cognitive Science Society, Vancouver, Canada, p. 1711-1716, 2006.
- Margulis E., « A model of melodic expectation », *Music Perception*, vol. 22, n° 4, p. 663-714, 2005.
- Meyer L. B., *Emotion and Meaning in Music*, Univ. of Chicago Press, 1956.
- Narmour E., *The Analysis and Cognition of Melodic Complexity : The Implication-Realization Model*, The University of Chicago Press, 1992.
- Nielsen F., Boissonnat J.-D., Nock R., « On Bregman Voronoi Diagrams », *Proc. 18th ACM-SIAM Sympos. Discrete Algorithms*, 2007a.
- Nielsen F., Nock R., « On the Centroids of Symmetrized Bregman Divergences », *Arxiv.org*, 2007b.
- Pachet F., « The Continuator : Musical Interaction With Style », *Proc. of International Computer Music Conference*, Gotheborg, Sweden, September, 2002.
- Pearce M., Conklin D., Wiggins G., « Methods for Combining Statistical Models of Music », in U. K. Wiil (ed.), *Computer Music Modelling and Retrieval*, p. 295-312, 2004.
- Pressing J., « Cognitive complexity and the structure of musical patterns », *Proceedings of the 4th Conference of the Australasian Cognitive Science Society*, Newcastle, 1999.
- Puckette M., « Pure data », *Proc. Int. Computer Music Conf.*, Thessaloniki, Greece, p. 224-227, September, 1997.
- Rosen R., *Anticipatory Systems*, vol. 1 of *IFSR International Series on Systems Science and Engineering*, Pergamon Press, Oxford, 1985.
- Saffran J. R., Johnson E. K., Aslin R. N., Newport E. L., « Statistical learning of tonal sequences by human infants and adults. Cognition », *Cognition*, vol. 70, p. 27-52, 1999.
- Shannon C. E., « A mathematical theory of communication », *The Bell System technical journal*, vol. 27, p. 379-423, 1948.
- Sutton R. S., « DYNA, an Integrated Architecture for Learning, Planning and Reacting », *Working Notes of the AAAI Spring Symposium on Integrated Intelligent Architectures*, 1991.
- Sutton R. S., Barto A. G., *Reinforcement Learning : An Introduction*, MIT Press, 1998.
- Zhang J., « Divergence function, duality, and convex analysis », *Neural Comput.*, vol. 16, n° 1, p. 159-195, 2004.

SERVICE ÉDITORIAL – HERMES-LAVOISIER
14 rue de Provigny, F-94236 Cachan cedex
Tél. : 01-47-40-67-67
E-mail : revues@lavoisier.fr
Serveur web : <http://www.revuesonline.com>