

Rapport de soutenance Zoe Fayet

Zoe Fayet a présenté de manière synthétique et pédagogique son travail très novateur sur la recherche par le contenu dans des corpus de partitions musicales -

Le jury a apprécié la mise en valeur de la généralité de l'approche pour l'étude de séquences temporelles -

Ses réponses aux très nombreuses et difficiles questions ont illustré son recul et sa maîtrise du sujet. Son travail approfondi s'est concrétisé par une réalisation pratique - Il ouvre sur de nombreuses perspectives sur le domaine étudié et plus largement sur tout domaine lié aux séquences temporelles.

C2
G. JOMIER

B. DEFUDE

D. LAURENT
P. Régaut

C. de MOUZA

G. GUYOT

RAPPORT SUR LA THESE DE Zoé FAGET

Titre de la thèse : « Un modèle pour la gestion des séquences temporelles synchronisées.
Application aux données musicales symboliques »

La thèse de Zoé FAGET s'est déroulée au sein du laboratoire LAMSADE à Paris-Dauphine sous la direction de Philippe RIGAUX. Elle se situe dans le domaine de la gestion de séquences temporelles synchronisées avec application aux données musicales symboliques. Cette thèse est découpée en cinq chapitres plus une introduction et une conclusion, le tout formant 96 pages. Le premier chapitre de la thèse présente un état de l'art large sur les thématiques abordées durant la thèse (séquences temporelles, similarité, indexation). Le deuxième chapitre présente un modèle et une algèbre pour représenter et manipuler des séquences temporelles alors que le troisième chapitre décrit le langage de requêtes proposé. Un index spécifique permettant de résoudre efficacement des requêtes exactes et approchées est ensuite exposé dans le quatrième chapitre. Enfin, le dernier chapitre décrit la plateforme Neuma sur laquelle les propositions ont été déployées.

Le chapitre I présente tout d'abord l'état de l'art sur la modélisation et la gestion de séquences temporelles. Il décrit également un ensemble de distances permettant de faire de la recherche par similarité en présentant notamment celles liées au temporel ainsi que les distances n-gram, reprises ensuite au chapitre IV. Les méthodes d'indexation utilisables dans le contexte de l'étude sont ensuite analysées. La présentation est complète, couvre un large champ et montre bien la bonne maîtrise du domaine de Z. Faget. On aurait cependant aimé avoir un peu plus de détails sur la partie indexation ainsi que sur les modèles et langages proposés pour l'interrogation de bases de données de capteurs.

Le chapitre II commence la partie contributions par la définition d'un modèle logique destiné à la gestion de séquences temporelles et une algèbre associée. L'idée proposée ici est d'étendre le modèle relationnel par l'introduction d'un type de données spécifique appelé séquence temporelle. Une telle séquence peut prendre ses valeurs dans un domaine simple ou complexe, ce qui permet alors de représenter du synchronisme. Cet ajout d'un type spécifique, quoique simple, permet effectivement de représenter des séquences temporelles de manière abstraite et donc indépendamment d'un domaine d'application. L'introduction de ce type particulier nécessite l'extension des opérations de base de l'algèbre relationnelle, à savoir sélection, projection et produit cartésien. Une algèbre temporelle est ensuite définie pour manipuler les séquences temporelles au travers de trois opérateurs, composition (permet de transformer une séquence via une fonction), addition et dérivation-agrégation. Ces opérateurs sont très généraux et sont paramétrés par une (ou des) fonction(s).



Quelques exemples supplémentaires auraient sûrement amélioré la lisibilité de la présentation un peu aride. On aurait également aimé avoir une réflexion sur l'expressivité de l'algèbre proposée. Le choix d'avoir peu d'opérateurs paramétrés par des fonctions, permet d'avoir une algèbre « compacte » mais va également poser des problèmes à l'optimisation puisque une partie éventuellement importante du traitement peut être « caché » dans les fonctions.

Le chapitre III propose un langage d'expressions algébriques construit au dessus de l'algèbre proposée. Une grammaire est proposée ainsi qu'une traduction des expressions du langage vers l'algèbre définie précédemment. Une implantation en OCaml au dessus de MySQL a permis de valider l'interprète ainsi défini. Il manque pour l'instant un optimiseur pour ce langage.

La problématique d'indexation des sources musicales symboliques est abordée dans le chapitre IV. L'objectif est de construire un mécanisme d'indexation spécifique aux données musicales symboliques qui permette des recherches efficaces tant pour des requêtes exactes qu'approchées. Pour les recherches approchées, la distance visée est celle d'édition mais elle est connue comme étant difficile à indexer. L'idée retenue ici est de construire un index sur la base d'une borne inférieure de la distance d'édition et pour cela la distance n-grams est retenue. L'index retournera donc un sur-ensemble des résultats qu'il conviendra de filtrer dans un deuxième temps pour enlever les faux positifs. L'approche retenue consiste à reprendre le principe du AS-index développé dans l'équipe pour filtrer des documents et de l'adapter au contexte musical. Il repose sur des signatures algébriques et les propriétés sous-jacentes de ces structures. Des algorithmes de recherche exacte et approchée utilisant l'index construit sont ensuite présentés. Il aurait été bon néanmoins de mieux expliciter ce qui est spécifique au MS-index de ce qui provient du AS-index.

Le chapitre se termine par un ensemble de résultats expérimentaux permettant de valider l'indexation proposée tant en recherche exacte qu'approchée. Les collections musicales utilisées sont des collections récupérables en ligne qui représentent environ 7500 fichiers et de l'ordre de 1.2 Gbytes. Les premières mesures portent sur le temps de construction et la taille du MS-index. La taille de l'index est faible relativement aux données indexées. Elle dépend de la taille des n-grams, mais celle-ci joue surtout sur le temps de construction de l'index, sachant que celui-ci reste cependant toujours petit. L'indexation diminue les temps de recherche exacte de manière importante relativement à une recherche séquentielle (facteur d'accélération de l'ordre de 8 à 80 selon les jeux de données et les types de recherche).

Le gain obtenu en recherche approchée est encore plus important (facteur d'accélération jusqu'à 280). Les expérimentations montrent également que la distance n-grams est finalement une très bonne approximation de la distance d'édition pour les collections testées.

Les résultats expérimentaux sont donc très positifs. Cependant, une comparaison avec d'autres approches d'indexation aurait été intéressante pour avoir une idée plus précise de la réelle accélération apportée par le MS-index. Les types de matériel et de logiciels utilisés pour les expérimentations seraient également les bienvenus.

Le chapitre V présente la plateforme NEUMA qui a pour but de fournir du contenu musical symbolique à disposition de communautés d'experts via une plateforme collaborative. Une architecture à base de services a été définie pour NEUMA dans laquelle les contributions proposées dans cette thèse s'intègrent naturellement (modélisation de sources musicales symboliques, langage de requête et indexation). Nous n'avons pas d'indication sur le statut de la plateforme, mais sa mise en exploitation permettra sans doute de pouvoir valider les contributions de cette thèse par l'usage (par exemple, quels sont les types de recherche préférés des experts).

La conclusion dresse le bilan de ce travail et donne quelques perspectives.

Sur la forme, le document est de bonne qualité même si on aurait aimé avoir plus d'exemples illustratifs pour présenter les contributions. Sur le fond, la démarche de recherche est complète puisqu'elle va de l'analyse du problème, la formalisation des contributions, une implantation (y compris au sein d'un système plus large NEUMA) et des validations expérimentales. Les contributions apportées peuvent s'appliquer à de nombreuses applications manipulant des séquences temporelles synchronisées. Il reste du travail d'expérimentation à faire pour être tout à fait convaincu du caractère générique des contributions, mais le résultat obtenu est d'ores et déjà satisfaisant. Sur le plan des données musicales, les résultats obtenus pour les recherches tant exactes qu'approchées sont très prometteurs.

Ce travail a déjà fait l'objet de plusieurs publications dans des conférences du domaine au niveau national et international, ce qui situe bien son apport.

Cette thèse ouvre des perspectives intéressantes qui sont esquissées dans la conclusion.

En conclusion, le travail effectué est de bonne qualité et je donne un avis favorable à la soutenance de ce mémoire par Zoé FAGET pour obtenir le titre de docteur en informatique de l'Université de Paris-Dauphine.

Evry, le 20 novembre 2011

Bruno DEFUDE
Professeur, HDR
Institut TELECOM, TELECOM SudParis
UMR CNRS 5157 SAMOVAR



RAPPORT SUR LE MEMOIRE DE THESE
Un modèle pour la gestion des séquences temporelles synchronisées.
Application aux données musicales symboliques
présenté par Zoe FAGET

Le travail décrit dans la thèse s'intéresse au problème de la représentation et de l'interrogation de séquences temporelles, en considérant plus précisément les données musicales, vues comme des données symboliques. La thèse contient deux contributions principales, à savoir d'une part la définition d'un langage d'interrogation de données contenant des séquences temporelles, et d'autre part, la définition d'une méthode d'indexation pour la recherche exacte ou approchée de séquences temporelles.

Outre une introduction et une conclusion, le mémoire est structuré en cinq chapitres respectivement consacrés à (i) un état de l'art concernant différents sujets liés aux séquences temporelles, (ii) la modélisation des séquences temporelles, (iii) la définition d'un langage de requêtes pour les séquences temporelles, (iv) une méthode d'indexation pour la recherche exacte ou approchée de séquences temporelles, et (v) la description d'un prototype implémentant les concepts développés dans les chapitres précédents.

Dans le chapitre 1, la notion de séquence temporelle ainsi que les domaines d'applications sont tout d'abord introduits, puis un bref historique des bases de données temporelles est proposé. Après avoir remarqué que les bases de données temporelles ne permettent pas de traiter les séquences temporelles, l'auteur consacre une section aux bases de données de séquences temporelles dans laquelle les principaux systèmes et langages introduits dans la littérature sont rappelés. La section suivante est consacrée aux bibliothèques numériques, dans lesquelles la notion de séquence temporelle joue un rôle important. Les deux sections suivantes sont consacrées à des aspects plus techniques concernant la recherche approchée de séquences : les notions de *précision* et de *similarité* sont rappelées, et les différentes distances entre séquences utilisées dans la littérature sont détaillées. Ces distances sont la distance euclidienne, la distance d'édition, la distance *Dynamic Time Warping* (ou DTW), la distance *Longest Common Subsequence* (ou LCSS), et les distances *n*-grams. Les travaux plus spécifiques sur les données musicales font l'objet de la section qui suit, puis les différentes approches d'indexation de séquence (exacte et approximative) sont présentées et commentées. Le chapitre se termine par une section permettant de récapituler les problématiques abordées et de positionner et motiver la contribution de la thèse par rapport à l'existant.

Ce chapitre fait état d'une recherche bibliographique importante, et également du souci de l'auteur pour motiver son travail, non seulement dans le domaine d'application visé (la musique), mais également dans tout autre domaine concerné par les séquences temporelles.

Le chapitre 2 est consacré à la définition d'une extension de l'algèbre relationnelle permettant la prise en compte des séquences temporelles. L'utilité d'une telle algèbre, à partir de laquelle un langage d'interrogation pourra être défini, est tout d'abord motivée par de nombreux exemples, liés ou non aux données musicales. Dans l'extension présentée, une relation peut être définie sur des attributs standard et sur des attributs permettant la prise en compte de séquences temporelles, vues comme des fonctions d'un ensemble isomorphe à l'ensemble des entiers naturels, noté T , dans un domaine complexe (produit cartésien de domaines

atomiques). Les opérations de base que sont la sélection, la projection, le produit et la jointure sont naturellement étendues et une opération testant si une séquence temporelle est vide ou non est introduite. Trois opérations spécifiques aux séquences temporelles sont ensuite introduites :

- la composition, combinant une séquence temporelle et une fonction de T dans T ,
- l'addition qui, liée à une opération o définie sur les valeurs atomiques des éléments des séquences, permet d'appliquer o élément par élément aux séquences argument,
- la dérivation, combinant dans un premier temps une séquence temporelle et une suite de fonctions de T dans T , puis permettant d'appliquer une fonction d'agrégation à la séquence résultat.

Les notions introduites dans ce chapitre sont illustrées par des exemples qui aident à la compréhension. Toutefois, ce chapitre est un peu bref car, d'une part l'opération de dérivation mériterait plus de commentaires, et d'autre part, aucune propriété de l'algèbre n'est étudiée. De plus, la conclusion du chapitre évoque l'*expressivité* du langage pour motiver l'introduction de cette algèbre, mais aucun élément formel sur cette propriété n'est donné...

Le chapitre 3 présente le langage de requête associé à l'algèbre du chapitre précédent. Dans ce langage, inspiré de SQL, une requête se présente sous la forme

from	liste de tables
let	variable := expression
construct	expressions
where	condition de sélection

et une liste de fonctions permettant la manipulation des séquences temporelles est donnée. La grammaire du langage est décrite puis il est mentionné que toute requête de ce langage peut être exprimée dans l'algèbre du chapitre précédent. Ce point, ainsi que la réciproque (faisant l'objet d'une simple affirmation page 54), mériteraient plus de détails. De nombreux exemples, principalement, mais pas uniquement du domaine musical illustrent le langage de requête, montrant en particulier son expressivité. Le chapitre se termine par une brève description de l'implémentation du langage, dont l'architecture repose sur une interface avec MySQL.

Ce chapitre, qui constitue une contribution intéressante, mériterait d'être développé afin de donner plus de détails sur, d'une part l'équivalence entre le langage et l'algèbre associée, et d'autre part sur les aspects de performance.

Le chapitre 4 est consacré à la seconde contribution de la thèse, à savoir la recherche exacte ou approchée de séquence dans une base de données contenant des séquences. Dans les deux cas, l'approche repose sur la proposition d'une technique d'indexation, appelée MS-index (*Musical Signature Index*), qui est une extension aux séquences musicales d'une technique d'indexation de chaînes de caractères, appelée AS-Index, et proposée récemment par l'équipe dans laquelle la thèse s'est déroulée. Dans le cas de la thèse, quatre types de recherche sont tout d'abord considérés selon qu'une recherche est exacte ou transposée, et avec ou sans rythme. Une séquence est représentée par une suite de couples (intervalle, durée), où un intervalle représente la différence entre deux *événements musicaux* (note, silence,...), vus comme des éléments d'un corps fini. Pour chaque n -gram à une position donnée dans la base, l'index stocke la position, la note la plus basse, l'identifiant du descripteur indexé et trois signatures associées au n -gram. Il est ensuite montré comment utiliser cet index pour la recherche d'un motif P dans les quatre cas évoqués ci-dessus. Un résultat important est que seuls le premier et le dernier n -grams de P sont recherchés via l'index. Les algorithmes de recherche dans les cas d'une recherche transposée sans rythme et d'une recherche exacte sans rythme sont détaillés.

Le cas de la recherche approchée est ensuite abordé, la notion de similarité étant définie à partir de la distance n -gram, dont le lien avec la distance d'édition est rappelé. L'idée générale est de rechercher dans la base, en se servant de l'index, tous les n -grams avec leur position qui sont présents dans le motif recherché. En appliquant une fenêtre glissante dont la taille peut être majorée, pour chaque position de fenêtre, on détermine d'une part, la distance minimale entre le motif recherché et le contenu de la fenêtre, et d'autre part, la plus longue sous-chaîne réalisant cette distance minimale.

Les résultats expérimentaux correspondant aux deux types de recherche (exacte et approchée) montrent de très bonnes performances. Dans le cas de recherche exacte, l'analyse porte sur le temps de construction et la taille de l'index, puis sur le temps de recherche en fonction de la taille des n -grams, et de la taille des données. Dans le cas de la recherche approchée, outre la comparaison des temps de recherche entre la stratégie proposée et une recherche exhaustive, le lien entre distance n -gram et distance d'édition est quantifié : il est montré empiriquement que les deux distances ont des valeurs très proches, montrant ainsi que la distance n -gram est une bonne approximation de la distance d'édition, tout en ayant de bonnes propriétés qui permettent l'indexation.

Le chapitre 5 décrit la plateforme appelée Neuma, qui a été développée dans le cadre de la thèse pour une communauté d'experts en musique, dans le but de mettre à leur disposition un outil de stockage, d'interrogation et d'analyse de données musicales. Cette plateforme est architecturée autour de cinq services permettant respectivement de stocker, interroger, rechercher par le contenu, afficher et annoter des documents musicaux symboliques. Les documents sont stockés et interrogés selon le modèle décrit dans la thèse, et l'annotation peut être soit textuelle libre, soit textuelle dont les termes sont choisis dans une ontologie du domaine, soit symbolique (variante musicale d'une portion de partition).

Ce chapitre présente une contribution importante car il s'agit de la mise en oeuvre effective des contributions théoriques de la thèse. Toutefois, afin de mieux illustrer et mettre en valeur cette partie, il aurait été utile, à titre d'exemple, de présenter une ou plusieurs applications utilisant la plateforme.

En conclusion de ce rapport, les contributions de la thèse sont *importantes* et *significatives*. Ces contributions concernent en effet un domaine de recherche original (les données musicales) et peuvent être appliquées à d'autres types d'applications utilisant les séquences temporelles. De plus, les concepts introduits dans la thèse ont donné lieu à l'implémentation effective d'une plateforme.

Le mémoire est rédigé de manière claire et illustré de nombreux exemples qui en rendent la lecture facile et agréable. Enfin, les résultats ont été publiés dans les actes de conférences internationales de renom. Pour toutes ces raisons, je donne un avis *très favorable* à ce que la thèse soit soutenue.

Fait à Cergy-Pontoise, le 2 novembre 2011.



Pr. Dominique Laurent
ETIS – Université de Cergy-Pontoise