Cartes conceptuelles pour l'ingénierie de cartes de connaissances adaptatives

Michel Crampes[1], Jean Villerd[1], Filip Velickovski[2], Nicolas Mille[3] et Sylvie Ranwez[1]

[1] CR LGI2P, Ecole des Mines d'Alès, Site EERIE, Parc sc. G. Besse, F – 30 035 Nîmes cedex 1 {prénom.nom}@ema.fr
[2]School of Computer Science and Engineering, UNSW Sydney NSW 2052, Australia, filipv@cse.unsw.edu.au
[3]NETIA Inc., Halle industrielle de Farjou, F – 34 270 Claret, n.mille@netia.net

Résumé

Nous présentons une méthode et un environnement pour l'ingénierie de cartes de connaissances adaptatives. L'approche adoptée consiste à concevoir et intégrer différentes cartes conceptuelles qui modélisent la sémantique du domaine et la pragmatique imposée par une application donnée. Cette approche est appelée "Domaine–Vue–Contrôles", en référence au modèle MVC, bien connu pour la conception des interfaces Orientées Objet.

La méthode DVC s'appuie sur un environnement de visualisation d'information interactif qui prend en compte les différentes étapes du modèle. Elle est présentée en suivant l'exemple de la conception d'une carte musicale. Elle permet d'identifier clairement la sémantique de la carte et favorise les interactions d'un opérateur humain dans un contexte particulier auquel il convient de s'adapter. L'ensemble du processus peut être automatisé.

Mots clés : Ingénierie des cartes de connaissances ; cartes auto-organisées ; navigation ; interaction.

1 Introduction

Le traitement d'informations toujours plus nombreuses et hétérogènes nécessite la mise en place de modalités d'organisation et de manipulation de ces informations. Les solutions proposées sont souvent basées sur une modélisation de la connaissance d'un domaine. Cependant plus ce modèle est approfondi, plus difficile est sa perception et son utilisation.

La carte de connaissances est un support efficace pour la communication d'informations complexes et la manipulation interactive de concepts d'un domaine particulier. Toutefois les attentes en termes de simplification et de clarification de la connaissance se heurtent à la difficile tâche de construire une telle carte. Cette difficulté augmente si l'on souhaite personnaliser la carte en fonction des contextes des utilisateurs. Cet article présente une méthode et un environnement dynamique pour la structuration et la visualisation de cartes adaptatives. L'approche consiste à concevoir et fusionner (*to map* en anglais que l'on pourrait plus rigoureusement traduire par "mettre en correspondance") différentes cartes de concepts modélisant la sémantique du domaine et la pragmatique de l'application. Elle repose sur un modèle que nous avons baptisé Domaine–Vue–Contrôles (DVC) car il s'inspire du principe Modèle–Vue–Contrôleur (MVC) utilisé en programmation par objets pour la conception d'interfaces. Elle s'en écarte cependant par certains aspects analysés dans les sections suivantes, par exemple le caractère distribué du contrôle. La méthode est mise en œuvre dans un environnement de visualisation dynamique, appelé MolAge, qui intègre les différents éléments du modèle DVC. Afin d'illustrer sa pertinence, nous commencerons par décrire un exemple concret d'application : le projet MBox. Nous montrons comment, en utilisant comme support une carte de connaissances interactive (ici un paysage musical), les tâches de navigation, d'indexation et de composition de *playlists* sont facilitées. La suite de l'article présente le modèle DVC à partir de ses trois composantes (Domaine, Vue, Contrôles) et la carte finale, résultat de la fusion de plusieurs cartes conceptuelles. Finalement nous montrons qu'il est possible d'automatiser le processus DVC pour concevoir dynamiquement des cartes de connaissances.

2 Une carte musicale bâtie sur des critères artistiques

Le terrain d'expérimentation de nos travaux est l'environnement MBox, conçu et développé pour assister les DJ dans la programmation musicale pour les radios ou les chaînes commerciales. Son but est de fournir un environnement interactif pour créer des *playlists* musicales à partir de vastes ensembles de titres musicaux, en interagissant avec une carte musicale. Celle-ci offre quatre fonctionnalités : la navigation dans une collection de titres, l'indexation d'un nouveau titre, la création de *playlists* d'apprentissage et la composition automatique d'une *playlist*. Nous nous concentrerons ici sur la navigation et sur l'approche d'Ingénierie de la Connaissance (IC) qui sous-tendent toutes ces fonctionnalités. Les autres aspects ont été détaillés dans l'article [5].

2.1 Contexte de l'application MBox

L'idée centrale est de permettre aux utilisateurs d'interagir avec les titres musicaux de la même façon qu'un DJ manipule ses pochettes de disques lorsqu'il prépare manuellement une *playlist*. Si nous imaginons ce DJ disposant un ensemble de titres dans un espace bidimensionnel selon des critères de similarité, le résultat serait un paysage musical dans lequel la proximité géographique représente la proximité artistique des titres. Nous avons développé un environnement orienté agents appelé MolAge, qui permet de visualiser un ensemble d'éléments. Dans notre cas ces éléments sont des titres musicaux et le paysage que nous proposons est le résultat de la projection sur un plan de l'indexation des titres musicaux. L'indexation consiste à valuer un ensemble d'attributs associés à chaque titre. Un DJ professionnel, partenaire dans ce projet, a identifié 23 descripteurs artistiques tels que goût, attitude, intelligence ou rythme. Chaque descripteur peut être valué entre 0 et 100. L'affectation manuelle de ces nombreuses valeurs ne peut pas être envisagée.

La volonté première de notre paradigme visuel est donc d'offrir une méthode rapide et efficace pour indexer plusieurs milliers de titres selon plus de 20 descripteurs sémantiques, tout en conservant le contexte habituel de l'expert : un paysage musical familier composé de quelques pochettes de référence disposées sur une table virtuelle. L'indexation se fait alors par un simple glisser-déposer d'un nouveau titre sur le paysage. Les valeurs de ses descripteurs sont obtenues par propagation des valeurs des descripteurs de ses voisins au moyen de l'algorithme K Nearest Neighbours (KNN). Sur FIG. 1, le titre *U2 : Sunday bloody Sunday* est indexé avec ses quatre voisins représentés par le spectre sémantique de leurs descripteurs.

Une fois ce paysage mis en place, il est possible d'interagir avec la carte pour naviguer dans l'ensemble de titres, en sélectionner certains et construire des *playlists* de référence. Celles-ci serviront de base d'apprentissage pour la composition automatique de nouvelles *playlists* comme exposé dans [5].

2.2 L'environnement MolAge

Dans l'environnement MolAge, les éléments visualisés sont des atomes animés par différentes forces. La structure d'un paysage résulte de la projection d'un nuage de points, représentant des données multivaluées, d'un espace m-dimensionnel sur un espace k-dimensionnel avec k<m. Ici les points sont des titres musicaux, l'espace m-dimensionnel est composé des 23 descripteurs (m=23) et k=2. L'approche utilisée pour la projection est le MultiDimensional Scaling (MDS) [15][2][13][16] qui minimise une fonction de *stress* entre n points décrits dans un espace m-dimensionnel [1][2]. Différents algorithmes peuvent être utilisés pour minimiser la fonction de stress. Nous utilisons l'algorithme *Spring Embedding* [18][6][7] car il ajoute un fort degré d'interactivité. Il considère chaque point comme étant un nœud d'un graphe où tous les sommets seraient reliés deux à deux par des ressorts virtuels. Ces ressorts les contraignent à maintenir une certaine distance entre eux. Des améliorations ont été apportées à cet algorithme pour pallier les problèmes qui surgissent quand le nombre de nœuds devient trop important [3]. Les distances entre les nœuds sont celles de leur espace d'origine. Un ressort est installé entre les n(n-1)/2 paires de nœuds, avec une longueur égale à cette distance. FIG. 1 montre le résultat obtenu en appliquant cet algorithme sur 46 objets musicaux représentés par leurs pochettes et décrits dans un espace à 23 dimensions. En cliquant sur un objet on obtient à gauche la liste de ses descripteurs valués. Ce paysage peut être vu comme un espace sémantique bi-dimensionnel. Il est simple et sémantiquement expressif du point de vue des experts DJ, mais il pèche par l'absence d'un dispositif d'assistance à la navigation.

2.3 Premières évaluations et limites

Une évaluation par des DJ professionnels a été effectuée. Il leur était demandé de charger le paysage de référence (FIG. 1), de tester l'indexation et la composition, et de se prononcer sur la pertinence des résultats obtenus et la qualité de l'environnement dans son ensemble.

Cette évaluation a montré la rapidité et la fiabilité des processus d'indexation et de composition dès le premier essai, ce qui montre une prise en main immédiate de l'environnement par l'utilisateur. Les experts ont en outre confirmé la pertinence des descripteurs utilisés et l'expressivité du paysage musical. Ils ont toutefois pointé la difficulté de parcourir un paysage enrichi et de retrouver un objet particulier du fait du trop grand nombre de titres visibles simultanément. Pourtant plusieurs outils d'assistance sémantique facilitant la visualisation (lentilles *fish-eye* ou *panning zoom*) ont été intégrés [5]. Néanmoins les experts attendent des dispositifs de recherche sémantiques plus élaborés. L'un voudrait que les titres d'un même compositeur ou interprète soient reliés, l'autre que l'on puisse retrouver les titres appartenant à une période musicale donnée, un troisième que l'on puisse mettre en évidence les croisements entre les artistes et les œuvres et faire ainsi apparaître, par exemple, que la chanson *With a little help from my friends* a été composée et enregistrée par les *Beatles* avant d'être interprétée par *Joe Cocker*.

Il ressort donc un besoin de personnalisation de la carte en fonction du contexte d'utilisation. Elle devrait être différente si elle s'adresse à un amateur de musique classique, de jazz ou de rock car leurs centres d'intéret sont différents.

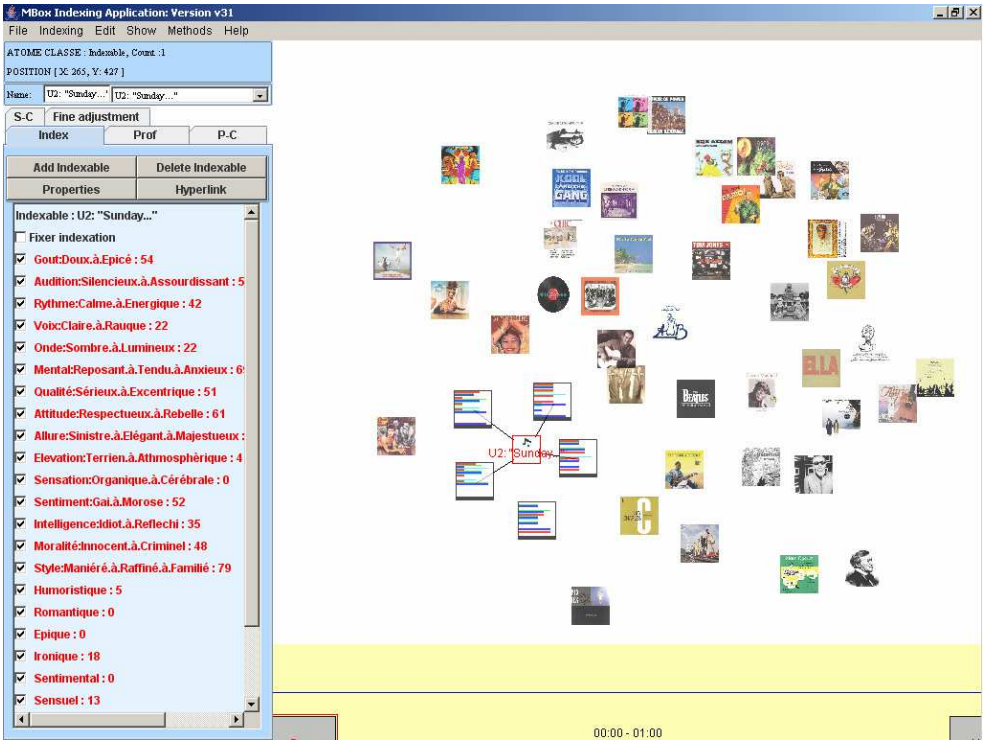

FIG. 1 – *Indexation sur une carte musicale de 46 titres générée par MDS (lentille spectrale sur les voisins)*

Ainsi, si l'on prend en compte l'aspect pragmatique (l'usage dans un contexte précis), nous ne devons pas concevoir une unique carte de connaissances pour la musique, mais générer différentes cartes de connaissances qui s'adaptent aux contextes de culture et de travail de l'utilisateur. Un paysage et un dispositif de navigation adaptés doivent être proposés pour augmenter l'utilité et la pertinence de notre environnement. Cela implique une approche totalement nouvelle et de nouveaux outils.

Nous avons ainsi défini l'approche Domaine–Vue–Contrôles et une méthode pour la mettre en œuvre dans MolAge. Le résultat est une nouvelle carte de connaissances musicale (voir FIG. 4) qui sera présentée plus loin. Le paysage est très différent de celui de FIG. 1 et prend en compte la plupart des fonctionnalités attendues. Il conserve toutefois l'idée originale d'une visualisation selon des critères artistiques. Nous présentons à présent l'approche DVC qui nous a permis de concevoir ce nouveau paysage et nous a ouvert la perspective d'une génération automatique de cartes de connaissances adaptatives.

3 L'approche DVC

La carte musicale représentée dans FIG. 1 est le résultat d'une projection automatique en une seule fois à partir de données renseignées à la main. Nous avons besoin d'une méthode plus formelle qui accompagne le concepteur de cartes de connaissances depuis la sémantique et la pragmatique vers la visualisation adaptée (personnalisée).

L'approche traditionnelle Modèle-Vue-Contrôleur (MVC) en conception orientée objet [10] implique que chacun des trois modèles soit conçu séparément, après quoi ils sont fusionnés. Le *Contrôleur* prend en charge les interactions de l'utilisateur, le *Modèle* est la partie du code qui transforme l'état interne des données en fonction des évènements déclenchés par le *Contrôleur*, la *Vue* gère les différents objets graphiques de l'interface représentant les données. Cette approche, selon laquelle les principales dimensions d'une application sont séparées, conduit à une conception adaptative et permet de spécifier différentes fonctionnalités (vues et contrôles) à partir d'un unique modèle qui constitue le noyau de l'application.

Nous détaillons maintenant la méthode DVC qui s'inspire de ce modèle.

3.1 La carte de concepts Domaine

La première étape de la méthode DVC consiste à identifier les concepts du domaine considéré qui sont pertinents pour l'utilisateur. Dans FIG. 2-a nous présentons la carte de concepts du domaine de la musique. Chaque concept du domaine est considéré comme un atome représenté par un cercle vert étiqueté. Différentes liaisons ont été dessinées entre les concepts. Nous utilisons des liaisons de types et de formes différents pour montrer leur signification spécifique. Trois principaux types de liaisons sont représentés

ici. La liaison **classe-superclasse** est une ligne terminée par un petit segment au niveau de la superclasse, l'équivalent du triangle en UML. La liaison **agrégation** est terminée par un point. Les autres sont des liens d'association. Leur sémantique peut être explicitée par une étiquette, mais elle est souvent évidente en fonction des noms de leurs extrémités.

La carte de concepts Domaine est dessinée grâce à l'environnement MolAge. On pourrait nous demander pourquoi cette carte n'est pas simplement construite en UML via un environnement tel que Rational Rose. Nous verrons à la fin de cet article qu'utiliser MolAge avec ses propres symboles est une meilleure solution dans le cadre de l'approche DVC.

La carte de concepts musicale rejoint les attentes des utilisateurs. En particulier, sept concepts constituent des points d'entrée pour la navigation et l'indexation : *Work*, *Albums*, *Compositor*, *Interpreter, Event* (naissance ou décès d'une personne), *Person* et *Descriptor*. *Descriptor* et *Album* étaient les seuls concepts pris en compte dans la carte de FIG. 1.

Cette carte de concepts pourrait être enrichie par les instruments, les influences entre les œuvres, les droits, etc. En particulier il y manque les interprétations qui seront intégrées dans une prochaine version. La carte pourrait devenir une ontologie complète, cependant nous n'avons pas besoin pour le moment d'une modélisation du domaine plus sophistiquée si l'on s'en tient aux attentes des utilisateurs. Nous avons néanmoins plus de dimensions que dans notre ancien modèle, ce qui pourrait accroître la complexité de la carte finale. Or nous allons justement les utiliser pour mieux organiser et clarifier le paysage.

3.2 La carte des Vues et Contrôles

Les Vues et les Contrôles sont également conçus dans l'environnement MolAge. **FIG. 2**-b représente la carte des principaux concepts de MolAge devant être utilisés afin de créer une carte de connaissances pour un domaine particulier. Indépendante du domaine, elle pourrait aussi devenir une ontologie des dispositifs de représentation. Un nouveau type de liaison est introduit, la simple flèche. Sa sémantique diffère de celle d'UML : ici elle signifie qu'un Contrôle est appliqué à un autre objet MolAge.

Les Vues sont associées à des atomes, des types d'atomes, ou des liaisons. Le type des atomes peut être prédéfini ou créé à l'exécution. Un atome peut changer de type dynamiquement (transtypage). Les liaisons constituent des relations topologiques entre atomes. Elles sont explicitement spécifiées dans ce modèle. *Alphabetic*, *Time* et *Spatial* sont des *structuralAtomTypes* qui peuvent être considérés comme des briques permettant de structurer la carte finale selon une certaine dimension, comme nous le verrons par la suite sur l'exemple de la carte musicale.

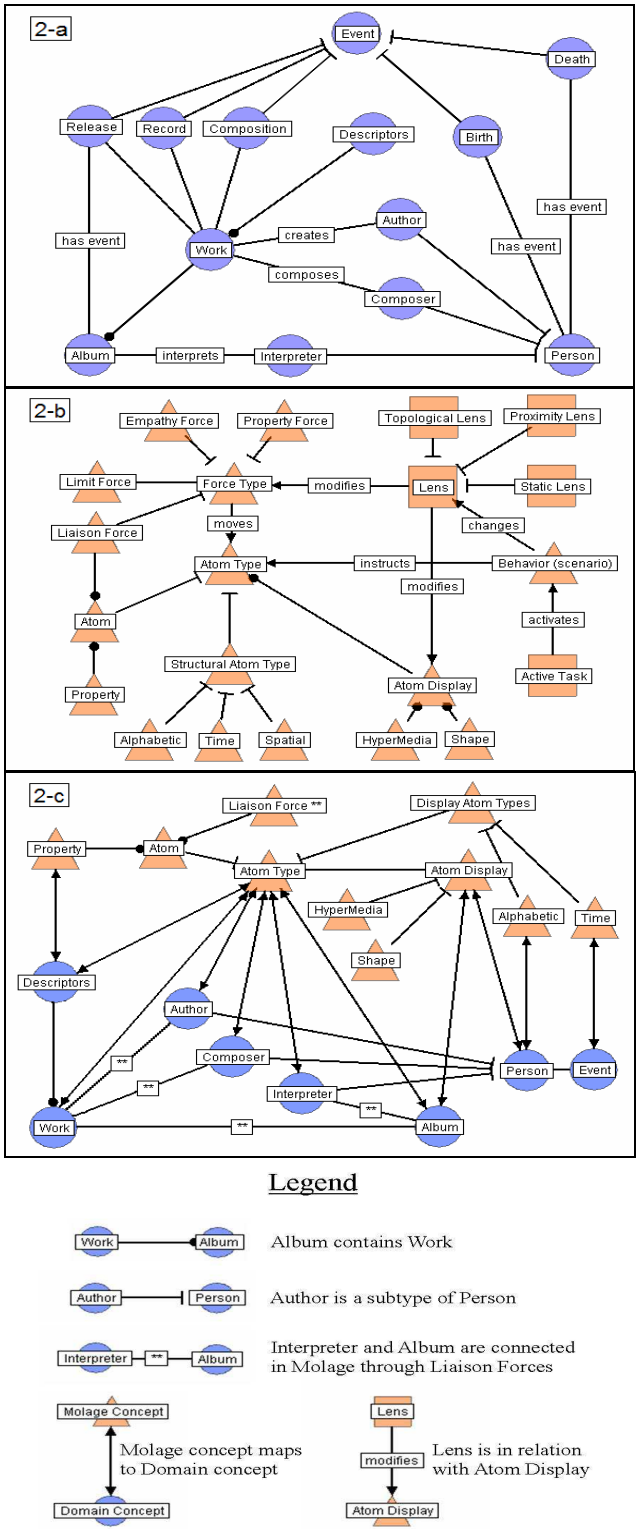

FIG. 2 – *Cartes conceptuelles pour l'approche DVC*
a – Carte conceptuelle du domaine,
b – Carte conceptuelle de MolAge
c – Fusion des deux cartes

Les propriétés sont des couples attribut-valeur associés aux atomes pour décrire des caractéristiques particulières. Elles jouent le même rôle que dans l'ancienne application mais, selon le nouveau modèle, elles peuvent être déclarées en tant qu'atomes et être typées. Enfin *atomDisplay* est un ensemble d'instances de vues possibles qui peuvent être associées à chaque atome ou à un type d'atome. Il comprend des images, des formes, des hyperliens, etc.

Les Contrôles dépendent de deux types de contraintes : les contraintes d'affichage et les contraintes de comportement. Un contrôle peut être appliqué à un atome particulier, à un type d'atomes ou à tous les atomes de la carte.

Les forces et les lentilles, deux sortes de contrôle d'affichage, peuvent être appliquées pour réaliser les trois fonctionnalités suivantes.

1. **Structurer avec les forces**. Les forces s'appliquent entre les atomes d'un même type ou entre types d'atomes. Elles constituent l'élément dynamique organisant la carte et sont utilisées pour séparer ou regrouper des atomes (force empathie), pour empêcher les chevauchements (force limite), pour maintenir les distances sémantiques (force propriété), et pour conserver une longueur donnée sur les liens topologiques (force liaison). Ces forces imitent les forces moléculaires et, de même que pour les molécules, leur combinaison produit différents effets structurants sur la carte.
   Elles peuvent utiliser différents algorithmes *push-pull* entre les atomes, parmi lesquels le célèbre algorithme des ressorts pour les forces de propriété et de liaison. L'emploi de ces forces permet à la carte de réagir aux manipulations de manière flexible tout en conservant une structure définie, ou de se transformer dynamiquement pour adopter une nouvelle structure plus pertinente pour l'usage considéré.
2. **Visualiser avec les lentilles**. Les lentilles sont un moyen d'ajuster les différents affichages relatifs aux atomes ou aux types d'atomes. Elles sont également définies en fonction de l'usage en cours. Les lentilles de **proximité** s'appliquent aux objets situés à une certaine distance paramétrable du curseur de la souris. Les lentilles **topologiques** sélectionnent différentes entités et propagent le mode d'affichage sélectionné dans le sous-graphe. FIG. 4 et FIG. 5 illustrent l'application de lentilles topologiques.
3. **Interagir**. L'utilisateur interagit avec la souris. Il peut sélectionner un atome en particulier, le déplacer et modifier son profil (forces, affichage, position, liaisons, etc.). Il peut faire de même avec un type d'atomes. Il peut encore déclencher ou arrêter les forces, donner une intensité à une force, créer des atomes, des liaisons, éditer et activer des lentilles, etc.

L'utilisation de tous ces contrôles peut se révéler complexe et engendrer de l'instabilité. Mais il apparaît que certains patrons de séquences de contrôles sont particulièrement efficaces et fiables pour structurer la carte selon une sémantique et l'organiser selon les besoins de l'utilisateur (pragmatique). Nous avons identifié les patrons les plus efficaces et allons les réaliser sous forme de scénarios dans un futur proche.

3.3 Fusion des cartes de concepts

La fusion de la carte de concepts musicale dans la carte de concepts MolAge est représentée dans **FIG. 2**-c.

Concernant les concepts, la fusion consiste à créer une liaison entre les atomes concepts du domaine (par exemple *Interpreter*) et les atomes concepts de MolAge; La plupart de ces liaisons sont dirigées vers l'atome *AtomeType* puisque l'on peut supposer qu'un type d'atomes MolAge spécifique sera créé pour chaque concept du domaine devant être visualisé dans la carte finale.

Concernant les liaisons, lorsque deux concepts sont liés dans la carte du domaine, leur liaison doit implicitement être fusionnée avec un type de liaison de MolAge.

La fusion est une spécification topologique de la carte finale. Les concepts du domaine sont à présent des types d'atomes, et les liens sémantiques entre ces concepts des liaisons entre types d'atomes dans MolAge.

3.4 Structuration spatiale

Étant donné que la fusion des deux cartes de concepts a été réalisée dans le même environnement que celui de destination de la carte finale, il est possible de spécifier la visualisation finale et les perspectives spatiales attendues. Le principe est de positionner par glisser-déposer les atomes du domaine sur des zones particulières de la carte et de spécifier des contrôles sur ceux-ci. La carte de spécification qui résulte de ce processus est représentée dans FIG. 3.

Sur cette figure, chaque concept de la carte du domaine est à présent considéré comme la spécification d'un type d'atome dans MolAge. L'emplacement d'un concept définit les coordonnées (x,y) initiales de tous les atomes de ce type dans la carte de connaissances finale avant d'appliquer les forces. Une combinaison particulière de contrôles (forces, lentilles et contraintes spatiales) est également spécifiée et appliquée à chaque concept. Dans notre exemple, le concept *Album* est ainsi structuré selon le temps par une force dont l'unique degré de liberté est l'axe x. À l'opposé, la force de propriété est appliquée à ce même *Album* sur y. Le résultat attendu est que les instances du concept *Album* se disposent selon le temps en x et selon une projection MDS unidimensionnelle en y dans la carte finale.

Différentes perspectives spatiales peuvent être ainsi spécifiées à partir de différentes cartes de spécification. Il est

alors possible d'observer et de manipuler la carte musicale finale différemment selon les usages et les centres d'intérêts de l'utilisateur.

3.5 Déploiement de la carte finale

La carte finale est construite d'après la carte de spécification (FIG. 3). L'utilisateur peut charger ou créer des ensembles d'atomes correspondant aux différents concepts du domaine. Toutes les instances d'un même type sont positionnées au même endroit, i.e. aux coordonnées du concept parent dans la carte de spécification. Les forces et les contraintes spécifiées sont ensuite appliquées sur les atomes. Ceux-ci sont libérés dans le paysage et se répartissent selon les spécifications. Le résultat de cette étape finale sur notre exemple de carte musicale est illustré dans FIG. 4.

Nous décrivons dans un premier temps comment l'on obtient ce résultat, ensuite comment nous gérons la complexité due au nombre d'atomes, notamment les albums et les œuvres.

Au bas de l'écran (FIG. 4), les compositeurs et les auteurs sont répartis automatiquement selon une ligne de temps avec deux atomes aux extrémités. Le processus suivant est utilisé pour construire une telle ligne de temps. Deux atomes d'un type fonctionnel spécifique *TemporalIndicator* sont créés. Ils ne sont munis que d'une seule propriété, *date,* qui prend les valeurs extrêmes de la ligne de temps, ici 1650 pour l'un, 2005 pour l'autre ; et sont fixés aux extrémités de la future ligne de temps. Chaque compositeur possède la propriété date valuée par sa date de naissance. Les valeurs de date sont normalisées entre 0 et 100. La force de propriété est alors appliquée sur le type compositeur/auteur (seuls les atomes de type compositeur/auteur s'attirent ou se repoussent entre eux) en ne prenant en compte que la propriété *date*. Les compositeurs/auteurs étant fixés en y, ils s'organisent selon la ligne de temps virtuelle qu'ils créent par leurs interactions. S'ils se chevauchent, ils peuvent être séparés grâce à une lentille de proximité qui applique la force de limite sur leur type lorsque l'on approche la souris.

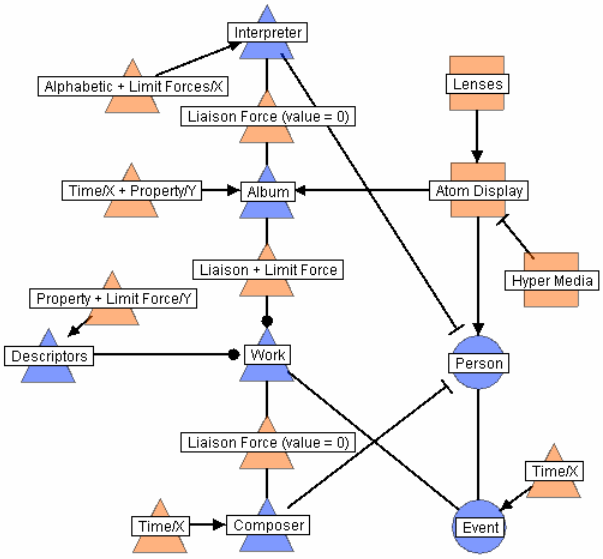

FIG. 3 – *Spécification de la carte musicale finale*

Un processus très similaire est mis en œuvre pour construire la ligne alphabétique en haut de l'écran en introduisant une propriété *alphabet,* valuée selon le rang de la première lettre du nom des interprètes.

Déployer les œuvres et les albums est plus problématique car on peut suivre différentes stratégies. Tous doivent être déployés au centre de l'écran selon deux dimensions, le temps et les propriétés artistiques. Une propriété temporelle est définie sur les albums dont la valeur correspond à leur date de composition. Les albums sont libérés selon x et se disposent ainsi le long de la ligne de temps. Ils sont ensuite fixés en x et libérés en y pour se déployer en fonction des valeurs de leurs propriétés artistiques. On obtient ainsi une projection bidimensionnelle des albums selon deux dimensions indépendantes, le temps en x et les propriétés en y. Sur cette dernière dimension, le résultat est une projection MDS unidimensionnelle à comparer avec la projection 2D de la première carte. Nous y reviendrons. Puisqu'un album contient plusieurs œuvres et qu'une œuvre peut être présente dans différents albums, nous avons *n x m* liaisons entre les albums et les œuvres. Les œuvres peuvent être déployées de la même façon que les albums si elles disposent de propriétés artistiques, mais on pourrait également utiliser la force de liaisons pour qu'elles se disposent autour de leurs albums. Enfin, les descripteurs artistiques sont disposés sur la gauche, conformément aux spécifica-

tions, selon la dimension y comme le montre FIG. 4. Pour cela chaque atome descripteur reçoit une valeur 100 pour la propriété qu'il représente. La force de propriété est appliquée depuis les albums vers les descripteurs toujours selon la carte de spécification. Ceux-ci sont ensuite libérés en y et ils se placent à la position du centre de gravité des albums pour la propriété qu'ils représentent. La force limite est enfin appliquée pour séparer au mieux les descripteurs.

3.6 Interaction et montée en charge

À ce stade, deux questions surviennent : cette nouvelle carte est-elle mieux adaptée aux usages de l'utilisateur et permet-elle de gérer des centaines voire des milliers d'objets musicaux (compositeurs, albums, œuvres, etc.) ?

Notre application a pour but l'indexation, la navigation et la recherche d'objets musicaux puis la composition de *playlists*. Nous ne nous intéresserons pas ici à l'indexation et à la composition à partir de cette nouvelle carte (ces thèmes sont traités dans un autre article en préparation). Nous nous concentrons donc sur les progrès relatifs à la navigation et à l'interaction.

FIG. 4 montre une nouvelle carte musicale selon une approche orientée albums, l'équivalent de la carte musicale de FIG. 1, à ceci près que la projection MDS est seulement unidimensionnelle selon y, l'autre dimension x représentant la date de sortie pour les albums de pop music ou la date de composition pour les albums de musique classique. La carte a été simplifiée pour plus de clarté et recentrée sur les *Beatles*. Cette figure démontre que la nouvelle carte permet différentes stratégies de navigation correspondant à différentes attentes de l'utilisateur.

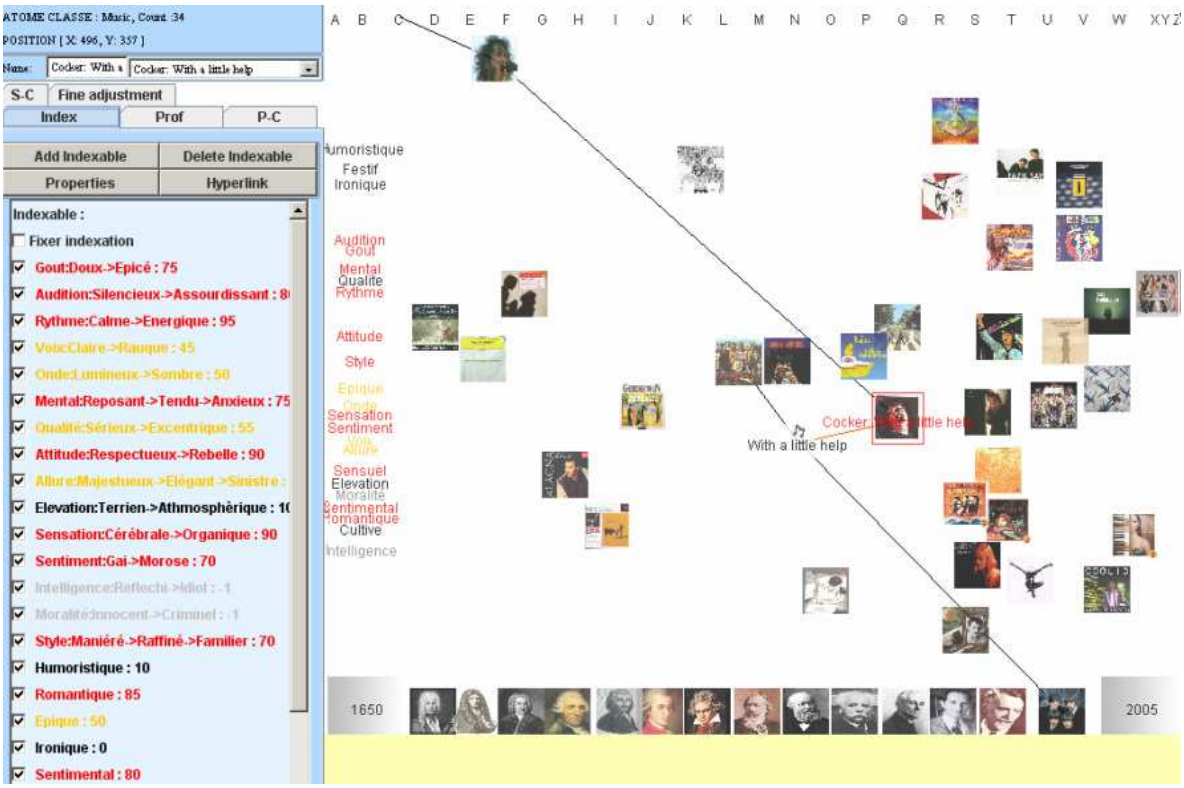

FIG. 4 – *La nouvelle carte résultant de la mise en œuvre de l'approche DVC (à comparer avec* **FIG. 1***)*

En cliquant sur une œuvre, par exemple *With a little help*, des liaisons apparaissent vers les différents albums contenant cette chanson, vers le compositeur et l'interprète. Sur la figure, deux albums apparaissent, l'un interprété par les *Beatles* (*Sgt Pepper*), l'autre par *Joe Cocker*. Les compositeurs en bas sont les *Beatles*. Une autre information est visible en cliquant sur un album ou sur une œuvre. Les descripteurs des propriétés dont la valeur dépasse 70 deviennent rouges, jaunes entre 40 et 70, les autres restant noirs ou devenant gris si la propriété n'est pas valuée. Les valeurs artistiques sont ainsi synthétisées et la liste des descripteurs dans le cadre de gauche de FIG. 1 devient inutile. Il a été conservé dans FIG. 4 pour montrer la différence entre les deux figures. La nouvelle carte fournit toute l'information artistique en un seul coup d'oeil. En effet, la liste de descripteurs de FIG. 1 ne présentait pas d'organisation particulière. À présent la liste se trouve à l'intérieur de la carte et est organisée selon la projection artistique des albums et des œuvres. Il y a plus d'informations pour une complexité moindre. Cette façon d'interagir est celle atten-

due par les DJs s'intéressant à la pop music. Ils construisent des *playlists* reposant sur des variations *d'humeurs* (en utilisant les descripteurs à gauche), et naviguent ensuite à travers les œuvres et les albums au centre de la carte. Il en va autrement d'une de nos évaluatrices, amatrice de rock et de jazz. Son point d'entrée dans la carte musicale réside plutôt dans les interprètes et le temps (le jazz des années 60 en l'occurrence). Elle peut parcourir les lettres en haut de l'écran, identifier un interprète d'après les photos qui apparaissent alors, et cliquer sur l'un d'entre eux. Un exemple très simplifié de cette interaction est montré dans FIG. 5. Un clic sur la lettre B révèle plusieurs interprètes comme les *Beatles* et *Count Basie*. En sélectionnant ensuite les *Beatles*, la carte se clarifie, ne laissant visible que l'information pertinente : les albums correspondants apparaissent au bout des liaisons, au centre de l'écran. On peut à présent naviguer à travers les albums, les œuvres, ou les compositeurs. En cliquant sur un album, on fait apparaître son contenu. Par exemple, dans FIG. 5 on voit le résultat d'un clic sur l'album *Sgt Pepper*. Il est intéressant de noter comment les œuvres se répartissent de façon à ne pas couvrir un autre album des *Beatles* ou à se chevaucher l'une l'autre ; ceci est possible grâce à l'application d'une combinaison de forces (liaison, empathie, limite). La lentille que nous avons utilisée dans cet exemple rend visibles toutes les liaisons actives pour mieux observer la topologie. Pour simplifier la visualisation, nous pourrions appliquer une autre lentille topologique cachant les liaisons entre les albums et les œuvres. Ces lentilles peuvent être prédéfinies ou créées à l'exécution par l'utilisateur au moyen d'une fenêtre spéciale.

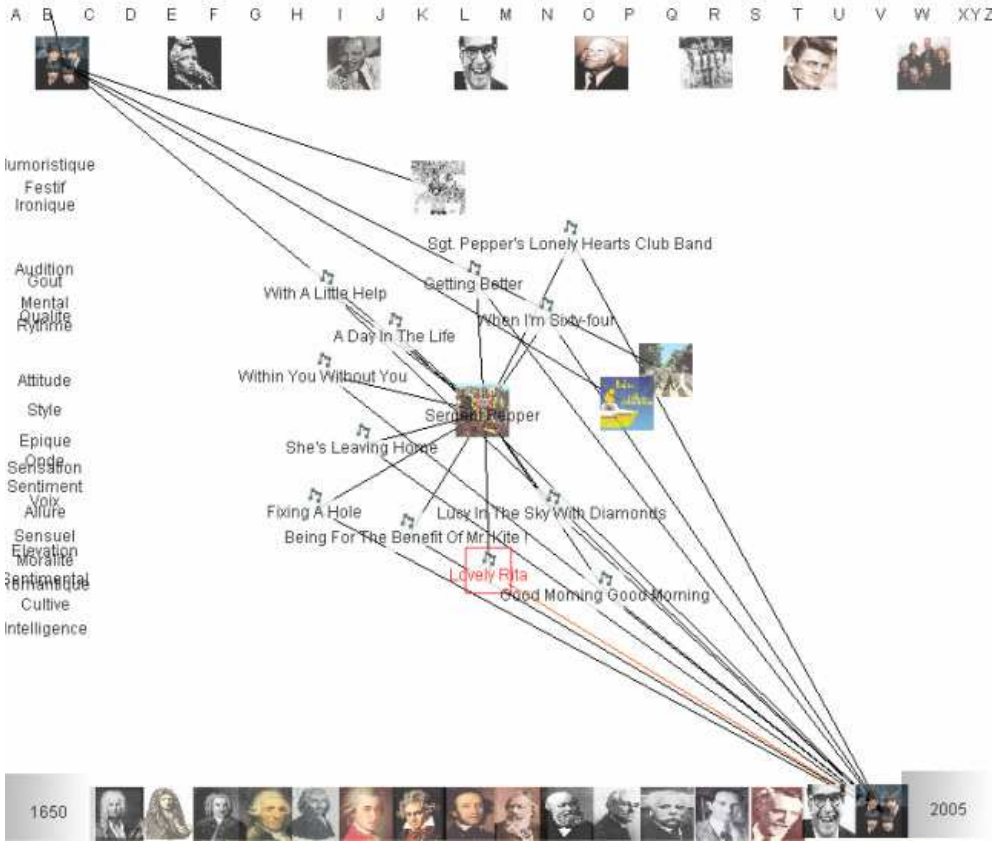

FIG. 5 – *Navigation via les Beatles*

Une autre évaluatrice, amatrice de musique classique souhaitait une navigation orientée temps et compositeurs. Elle peut trouver ces deux entrées en bas de l'écran avec les dates de composition (non représenté sur FIG. 5) et avec les compositeurs déployés selon la ligne de temps. Elle peut alors naviguer à travers les albums, les œuvres et même les interprètes.

4 Personnalisation de la carte

Dans l'exemple de FIG. 5 nous présentons une carte simplifiée unique pour trois mélomanes de styles différents : pop, jazz et classique. En pratique cela n'est pas souhaitable. Une raison évidente est que le paysage doit être cons-

truit à partir d'albums et d'œuvres totalement différents. Mais il existe d'autres raisons plus fondamentales.

Les amateurs de musique classique s'intéressent plutôt aux œuvres et aux compositeurs. Dès lors, la carte devrait être organisée avec les œuvres comme points de repère structurants et les compositeurs comme jalons sur la ligne de temps. Les amateurs de pop music et de jazz sont eux plus intéressés par les albums et les titres, un titre étant une œuvre interprétée dans un album particulier. La carte devrait être organisée selon les titres. Si la ligne de temps pour la carte est la date de composition des œuvres interprétées, ou la date de sortie des albums, la carte sera ramassée sur la droite. Notre exemple synthétise cela. Nous avons mélangé des albums pop avec des œuvres et des interprètes classiques. De fait, les titres récents se regroupent sur la droite puis, appliquant la force de limite pour éviter les chevauchements, repoussent des titres comme *Sgt Pepper* vers la gauche. Ainsi il serait mieux pour les titres pop d'avoir une ligne de temps différente (débutant en 1930 par exemple) ou de transformer celle-ci en ligne alphabétique comme celle des interprètes.

À travers cet exemple, il est clair que la carte de connaissances musicale doit présenter différentes perspectives qui rejoignent la culture et les attentes de l'utilisateur. Elles peuvent être spécifiées au cours de la fusion entre les cartes conceptuelles du domaine et de MolAge (**FIG. 2**), puis au cours de la mise en espace (FIG. 3).

Il est d'ores et déjà possible de changer de lentille pour passer d'une visualisation à une autre. Il devrait pouvoir être également possible de changer de perspective à l'exécution mais cela nécessite de nouveaux développements que nous sommes en train d'introduire dans une nouvelle version de MolAge. Cette partie du projet n'est pas achevée, mais nous la présentons en tant que travail en cours ou, plus généralement, comme une spécification d'un environnement d'ingénierie de cartes de connaissances (KME) semi-automatisé.

5 Vers la génération de cartes

Les cartes de connaissances et l'ingénierie de cartes de concepts sont sans doute les aspects les plus prometteurs de notre approche DVC. Par cette méthode, nous pouvons facilement spécifier différentes stratégies de visualisation (**FIG. 2**-c) mais, pour le moment, le déploiement final du modèle conceptuel est exécuté manuellement par le concepteur de cartes en utilisant le cadre sur la gauche dans FIG. 3. Notre prochain objectif est de générer les cartes automatiquement avec différentes perspectives pour assurer l'adaptabilité, à partir d'une carte de concepts du domaine et des besoins de l'utilisateur. En outre, s'il est possible de générer automatiquement une visualisation particulière, la personnalisation pourrait être accomplie en temps réel selon les intentions de l'utilisateur. Il apparaît que l'approche DVC et la technique de MolAge offrent de multiples possibilités d'automatisation. Deux caractéristiques de MolAge sont particulièrement intéressantes. Premièrement, toutes les visualisations peuvent être sauvées en XML grâce à un schéma spécifique reprenant tous les concepts de MolAge, y compris les types d'atomes, les forces et les lentilles. Deuxièmement, MolAge permet de spécifier des *scénarios* et de les sauvegarder selon un autre schéma.

Grâce à ces deux supports technologiques, notre KME devrait être en grande partie automatisé. En effet il convient d'observer que :

1. le modèle MolAge de **FIG. 2**-b est donné une fois pour toutes ;
2. le modèle du domaine de **FIG. 2**-a est spécifique au domaine et doit être réalisé manuellement, mais on peut imaginer l'import et la traduction automatique d'une ontologie OWL ;
3. le déploiement de la carte de connaissances (FIG. 4) à partir de la carte de spécification peut être entièrement accompli dans MolAge.

Ces automatisations possibles montrent l'intérêt de concevoir, éventuellement d'importer, tous ces modèles dans le même environnement mais le déploiement automatique nécessite toutefois un progrès important : les scénarios.

Nous avons vu en effet que le déploiement nécessite une coordination des différentes étapes. Pour les décrire et les contrôles, nous avons conçu un gestionnaire de scénarios. Un scénario spécifie un enchaînement particulier d'étapes soumises à des conditions de déclenchement et d'arrêt. Ce gestionnaire, en partie opérationnel, a été testé avec succès sur plusieurs cas.

6 État de l'art et conclusion

Utiliser des cartes de concepts pour l'ingénierie n'est pas nouveau [9][4]. En particulier, UML peut être considéré comme un meta-modèle de cartes de concepts pour le génie logiciel à objets. À partir d'un modèle de classes, plusieurs outils permettent de générer automatiquement le squelette du code Java, C++ ou Ada correspondant. L'inverse, i.e. la génération de cartes de concepts ou de cartes de connaissances à partir de spécifications sous forme de cartes de concepts, n'est pas si fréquent.

La génération automatique de cartes de concepts, et plus généralement de cartes de connaissances, est aussi ancienne que le dessin de graphe ou la visualisation d'informations, et beaucoup de solutions ont été déjà proposées dans ces domaines. Elles prennent en compte les critères esthétiques [18][11], la montée en charge [19][14][3], la navigation dans les graphes denses [20], la visualisation de relations complexes entre des entités structurées [17][8][12], la recherche de similarités entre des objets ou des données par projection [1][3]. Cependant dans tous ces

cas, la carte est générée en fonction de la topologie considérée ou des vecteurs d'attributs associés aux entités de la carte. À notre connaissance, aucune approche guidée par la sémantique et la pragmatique n'a été proposée.

Notre contribution est triple. Premièrement ni la topologie de la carte ni les données ne sont prédéfinies. Elles doivent être spécifiées en fonction du contexte de l'utilisateur. Les spécifications sont produites grâce à la fusion d'au moins deux cartes de concepts, celle du domaine et celle de l'environnement de visualisation. Différentes perspectives peuvent être spécifiées pour un unique modèle du domaine, ce qui permet de considérer des cartes de connaissances adaptées aux besoins de l'utilisateur.

Deuxièmement, ces spécifications, la conception des cartes de connaissances et leur déploiement, sont basés sur une méthode appelée Domaine–Vue–Contrôles qui consiste en une approche de conception de cartes guidée par la sémantique et la pragmatique. Cette méthode est intégrée à un environnement qui fournit les moyens de déployer la carte de connaissances à partir du modèle de concepts.

Enfin, nous montrons comment le processus DVC peut être en grande partie automatisé. Cet aspect est non seulement intéressant pour accroître la productivité et la fiabilité dans le domaine de l'ingénierie des cartes de connaissances, mais aussi pour la réorganisation de cartes en temps réel en fonction des changements survenant dans le contexte de l'utilisateur, assurant ainsi une meilleure personnalisation.

Références

[1] Basalaj, W. Proximity Visualisalisation of Abstract Data. *Technical Report UCAM-CL-TR-509*, University of Cambridge, 2001.

[2] Buja, A.,and Swayne, D.F Visualization Methodology for Multidimensional Scaling. In *Journal of Classification* 19(1):7-43, 2002.

[3] Chalmers, M. A Linear Iteration Time Layout Algorithm for Visualising High–Dimensional Data. *Proc IEEE Visualization'96*, San Francisco, pp. 127-132, 1996.

[4] Cornwell P.J. Concept Maps in the Mechanical Engineering Curriculum, *ASEE Annual Conference*, Washington D.C., 1996.

[5] Crampes M., Ranwez S., Velickovski F., Mooney C., Mille N. An integrated visual approach for music indexing and dynamic playlist composition, *Thirteenth Annual Multimedia Computing and Networking (MMCN'06),* San Jose, California, January 18-19, 2006.

[6] Eades, P. A heuristic for graph drawing. In *proc. of the 13th Manitoba Conference on Numerical Mathematics and Computing*. Vol. 24. Utilitas Mathematica, University of Manitoba, Winnipeg, Canada, Sept-Oct, 1984.

[7] Eades, P., Lin, X., Spring algorithms and symmetry. In *Theoretical Computer Science* 240(2):379-405, 2000.

[8] C. Fluit, M. Sabou and F. van HarmelenOntology-based *Information Visualization*, in Visualising the semantic web, Springer-Verlag, 2002;

[9] Gaines B. and Shaw M. WebMap: Concept mapping on the Web. *The World Wide Web Journal*, 1(1):171-183, 1996.

[10] Goldberg A. Smalltalk-80: The Interactive Programming Environment. Addison-Wesley Series in Computer Science, Boston MA, USA, 1984; 516pp.

[11] I. Herman, G. Melancon, and M. S. Marshall. Graph visualization and navigation in information visualization: A survey. IEEE *Transactions on Visualization and Computer Graphics*, 6(1):24-43, 2000.

[12] Le Grand B. Visualisation of the Semantic Web: Topic Maps Visualisation, *Information Visualisation Conference (IV2002)*, London.

[13] Morrison, A., Ross, G., Chalmers, M. Fast Multidimensional Scaling through Sampling, Springs and Interpolation. In *Information Visualization* 2(1):68-77, March 2003.

[14] Quang Vinh Nguyen Q. V. Huang M. L. EncCon: an approach to constructing interactive visualization of large hierarchical data, *Information Visualization*, 2005, 4(1):1-21.

[15] Platt, J.C. FastMap, MetricMap, and Landmark MDS are all Nyström Algorithms. *10th International Workshop on Artificial Intelligence and Statistics*, 261-268, 2005.

[16] Rodden, K., Basalaj, W., Sinclair, D., Wood, K., Does Organisation by Similarity Assist Image Browsing. In *Proc. of Human Factors in Computing Systems* (CHI 2001) ACM Press, pp. 190-197, April 2001.

[17] Shneiderman B (1996) The eyes have it: a task by data type taxonomy for information visualizations. *Proceedings of the 1996 IEEE symposium on visual languages (VL'96)*, pp. 336-343.

[18] Tollis, I., Battista, G., Eades, P., Tamassia, R., *Graph drawing : Algorithms for the visualisation of graphs*. Prentice Hall, 1999.

[19] Ware Ware, C. and Bobrow, R., Supporting visual queries on medium-sized node-link diagrams. *Information Visualization*, 2005, 4(1):49-58

[20] N. Wong M. Sheelagh T. Carpendale, Greenberg, S. EdgeLens: An Interactive Method for Managing Edge Congestion in Graphs. *IEEE Symposium on Information Visualization, InfoVis03.* IEEE Press. October 2003.