

MADEUS : Un modèle de documents multimédia structurés

Nabil Layaïda et Loay Sabry-Ismail

*Projet Opéra, INRIA Rhône-Alpes
655 avenue de l'Europe
38330 Montbonnot Saint-Martin*

RÉSUMÉ. Les documents électroniques ont jusqu'ici été considérés principalement sous l'angle de leur structure logique (organisation des documents en chapitres, sections, etc.), de leur structure spatiale (présentation graphique et mise en page) et de leur structure sémantique (hypertextes). Avec l'apparition de systèmes informatiques manipulant des données multimédia comme le son et la vidéo, un nouveau type de structure est maintenant considéré, la structure temporelle. L'élaboration d'un modèle de documents permettant d'intégrer ces différentes structures reste encore une tâche difficile, car d'une part les objets multimédia possèdent des propriétés très diverses, et d'autre part, l'adjonction de la structure temporelle aux documents électroniques remet souvent en cause les modèles abstraits habituellement utilisés pour les représenter. Dans cet article, nous exposons cette problématique et proposons des solutions. Ces solutions ont été élaborées à travers notre expérience dans la conception et implémentation d'un éditeur multimédia structuré appelé MADEUS.

ABSTRACT. Electronic documents have been mainly considered through their logical structure (organization of documents in chapters, sections, etc.), their spatial structure (graphical layout) and their semantic structure (hypertext). With the advent of new computer systems handling multimedia data like audio and video, the temporal structure is more and more considered. The design of a document model integrating these various structures remains a difficult task, because, on one hand, the multimedia objects have different properties, and on the other hand, the extension of traditional documents models to the temporal dimension requires heavy changes. In this paper, we expound these problems and propose some solutions. These solutions are based on our experience in designing and implementing a structured multimedia authoring system called MADEUS.

MOTS-CLÉ : Modèles de documents multimédia, structure temporelle, cohérence temporelle, formatage temporel, documents structurés multimédia.

KEY WORDS : Multimedia document model, temporal structure, temporal consistency, temporal formatting, structured multimedia documents.

1. Introduction

Les percées technologiques récentes en matière de multimédia [BUF 94] [BOU 94] [FUR 94] ont permis d'accroître les possibilités d'interaction entre l'homme et la machine. La manipulation dès lors possible de graphiques, de son et d'images animées par des stations de travail ou des ordinateurs personnels a changé la nature d'un grand nombre d'applications.

En particulier, les applications de traitement de documents électroniques, habituellement dédiées à la création et à la présentation de données textuelles et graphiques, trouvent dans le multimédia le moyen d'accroître la richesse de l'information qu'elles manipulent en intégrant, dans ces documents, du son et des images animées. Ces nouveaux types de documents électroniques sont communément appelés *documents multimédia*. De plus, le développement des réseaux de communication rapides permet d'envisager de relier ces documents pour constituer ainsi des réseaux de *documents hypermédia* à l'intérieur desquels on peut naviguer à la manière du *World-Wide Web* [BER 92].

L'intégration de données ayant des dimensions temporelles (durée, vitesse de défilement) comme la vidéo, l'audio et les interactions de l'utilisateur dans les documents électroniques introduit de nouveaux besoins qui doivent être pris en compte dans une nouvelle dimension : la dimension temporelle. Les standards et les modèles employés pour représenter les documents classiques, comme SGML [GOL 90] et ODA [ISO 89] sont devenus en conséquence inadaptés pour représenter de tels types de documents. De ce fait, de nouveaux standards, comme *HyTime* [HYT 92], *MHEG* [PRI 93] et *HyperODA* [ISO 92] émergent pour compléter ceux déjà existants, en particulier pour prendre en compte cette nouvelle dimension.

L'activité de recherche en cours dans les systèmes d'édition multimédia s'articule principalement autour de deux thèmes complémentaires en forte activité. Le premier thème concerne la définition de modèles de documents multimédia et le développement d'outils d'édition et de présentation mettant en œuvre ces modèles. Cette approche, dite descriptive, a pour premier objectif la définition d'un ensemble d'abstractions qui conviennent le mieux pour décrire, créer, manipuler, stocker et échanger des documents multimédia. Les modèles proposés sont généralement mis en valeur par des outils d'édition qui implantent ces abstractions et qui démontrent l'adaptation des fonctions de manipulation du document à l'utilisateur.

Le deuxième domaine s'oriente surtout vers le support à l'exécution des présentations multimédia, l'accès aux données par le biais du réseau, la construction d'interface homme-machine, le support de la concurrence, et la définition de langages de programmation appropriés pour construire ce type d'applications. L'exemple du projet AHM [HAR 93] (*Amsterdam Hypermedia Model*) illustre particulièrement bien la complémentarité entre les deux activités de recherche en multimédia. Le prototype d'éditeur hyper-média appelé *CMIFed* développé dans ce projet, a été réalisé conjointement avec le langage *python* offrant une grande partie des fonctionnalités requises. D'autres plate-formes de développement d'applications

multimédia comme Java [SUN 95] et ScriptX [VAL 94] sont actuellement en plein essor.

Dans le projet Opéra, nous nous intéressons à ces deux aspects qui, ensemble, nous ont conduit à la réalisation d'un premier prototype appelé MADEUS. L'idée de départ consiste à étendre un éditeur structuré existant pour les documents statiques, pour prendre en compte la dimension temporelle.

Dans cette étude, nous mettons en avant les principaux concepts relatifs à la mise en œuvre de modèles abstraits de documents électroniques supportant la structure temporelle. Nous présentons aussi les méthodes qui nous ont permis de réaliser l'intégration de cette structure avec les autres structures du document. Ces méthodes sont basées sur l'expérience que nous avons acquise lors de la réalisation du prototype MADEUS.

La section 2, présente les besoins de l'édition de documents multimédia selon l'approche structurée. La section 3, présente le modèle de document multimédia de MADEUS, tant sur l'aspect logique, spatial, hypermédia et surtout sur l'aspect temporel. La section 4, expose les problèmes et les solutions liés à la construction et à l'intégration de la structure temporelle dans les documents structurés. Un état de l'art en matière de modèles temporels abstraits et d'éditeurs multimédia est présenté dans la section 5. La section 6 termine cet article par des conclusions et quelques perspectives.

2. L'édition de documents multimédia

Un système d'édition de documents multimédia comporte principalement deux types de fonctions :

- *Des fonctions d'édition.* Elles réalisent les opérations de création, de construction et de modification du document par un auteur. L'opération de construction d'un document consiste à y inclure des objets multimédia de base, ensuite à spécifier des relations entre ces différents objets. Ces relations peuvent être liées à leur organisation logique (par exemple une scène composée de trois vidéo et d'un accompagnement musical), leur disposition spatiale sur les différents canaux à travers le temps (écran et/ou le haut parleur), ou encore à leur synchronisation temporelle.

- *Des fonctions de présentation.* Elle consiste à présenter à un utilisateur (le lecteur) le contenu du document une fois que son édition est achevée. Cette phase consiste à fournir au lecteur un ensemble de commandes permettant d'explorer ou de naviguer dans l'espace du document pour découvrir l'information qu'il contient à travers l'espace, le temps et l'interaction avec le document. Les commandes fournies au lecteur du document lui permettent l'arrêt d'une présentation multimédia, son redémarrage, son avance et retour rapide ainsi que la possibilité d'interagir avec certains éléments du document. Ce type d'interaction est aussi basée sur l'activation de liens hypermédia ou de boutons de présentation contenus dans le document, ou de façon générale sur tous les objets du document qui le permettent comme les objets multimédia programmés *applets* [VAL 94] [SUN 95].

Il est à noter que ces deux phases ne sont pas nécessairement dissociées. Le type de processus d'édition visé dans MADEUS est similaire à l'approche dite *WYSIWYG (What You See Is What You Get)* où les deux opérations d'édition et de présentation sont confondues [QUI 86]. À cause de la dimension temporelle des objets multimédia, une telle approche ressemble plus à la mise au point d'un programme puisque le comportement spatio-temporel d'un document n'est restitué que pendant sa phase de présentation (son exécution). Ainsi, le cycle d'édition d'un document est formé par une première phase pendant laquelle l'auteur introduit une ou plusieurs contraintes, puis, d'une seconde phase, dite phase de présentation pendant laquelle, il vérifie que le comportement du document correspond bien à l'effet recherché. C'est l'approche incrémentale de l'édition multimédia.

2.1. Besoins des systèmes d'édition multimédia

Pour construire un système d'édition de documents multimédia adapté à l'utilisateur, l'interface d'édition doit offrir un ensemble d'opérations permettant la création, la modification, le stockage et la présentation des documents. La qualité de ces opérations est fortement dépendante du modèle de document choisi et des mécanismes qui le mettent en œuvre. De façon générale, un système d'édition de documents multimédia doit répondre aux exigences suivantes :

- **Représentation structurée du document** : les aspects liés à l'organisation logique du document doivent être séparés des aspects liés à sa présentation physique. Cette structuration du document permet de garantir l'indépendance de la représentation du document (son format) par rapport à une plate-forme spécifique.

- **Processus d'édition incrémental** : la création d'un document se fait généralement par un enrichissement progressif de son contenu et de sa structure. Chaque opération d'édition effectuée par l'auteur, doit lui permettre de voir immédiatement son effet à la présentation.

- **Vérification de la cohérence** : pour chaque opération d'édition effectuée par l'auteur, le système d'édition doit vérifier sa validité par rapport à l'état courant du document. De plus, cette opération doit être efficace pour répondre à un choix de processus d'édition incrémental.

- **Intégration des différentes dimensions d'un document multimédia** : il est nécessaire de prendre en compte quatre aspects d'un document multimédia qui sont liés à son organisation logique, sa synchronisation temporelle, sa présentation spatiale et à l'interconnexion intra et inter documents (liens hypermédia).

- **Intégration de données multimédia hétérogènes** : Dans un système de document multimédia, il est primordial de réaliser l'intégration de données comme le texte, la vidéo, l'audio et les interactions avec l'utilisateur au sein d'une même structure de document. Cette intégration pose un certain nombre de problèmes liés principalement aux natures temporelles très diverses de ces objets (temporisés, statiques, indéterministes).

- **Résolution automatique des contraintes temporelles** : afin d'offrir à l'utilisateur des opérations de composition temporelle de haut niveau, le système d'édition doit prendre en charge toutes les contraintes numériques induites par ces opérations.

Ces différentes exigences attendues d'un système d'édition multimédia proviennent d'une part du besoin de construire des applications interactives simples d'utilisation, et d'autre part de permettre une représentation du document riche du point de vue sémantique et interchangeable entre différentes applications multimédia.

2.2. Modèles de documents multimédia

La définition d'un modèle abstrait de document multimédia est primordiale pour la conception d'un système d'édition. Le modèle doit permettre de représenter toutes les relations qui peuvent exister entre les objets d'un document multimédia. Ces relations peuvent décrire l'organisation logique du document, sa présentation spatiale, sa synchronisation temporelle ainsi que l'interconnexion entre ses différents éléments (liens). On appellera dans la suite ces différentes relations **les relations multimédia**. L'organisation logique concerne le regroupement des objets du document en entités sémantiquement liées. Par exemple, dans un document utilisé pour une présentation orale (suite de transparents), chaque transparent est généralement formé par un titre et un corps, qui à son tour peut contenir d'autres objets comme des images, de l'audio ou de la vidéo. La présentation spatiale concerne la disposition des objets selon les différents canaux à travers le temps (audio, fenêtre d'écran, etc.). Par exemple, le titre du transparent peut apparaître 2 centimètres plus haut que le corps pendant toute la durée allouée au transparent. La synchronisation temporelle concerne la disposition des objets du document dans le temps. Par exemple, le corps du transparent doit apparaître 2 secondes après le début du titre. La dimension hypermédia concerne la mise en place d'éléments particuliers dans le document qui permettent de naviguer entre des documents différents ou entre des parties différentes d'un même document.

La description de ces différents aspects d'un document peut être réalisée à différents niveaux :

- **générique** : à ce niveau, la description de l'aspect logique, temporel et spatial des documents est basée sur la notion de classes. Une classe de document ou DTD (*Document Type Definition*) [GOL 90][ISO 89] spécifie tous les types d'éléments qui peuvent être utilisés dans un document ainsi que toutes les relations structurales liées à l'organisation logique qui peuvent exister entre ces éléments. Ces informations structurales sont ensuite utilisées pour associer à ces classes des modèles de présentation spatiale, temporelle et hypermédia. C'est le principe des normes de documents hypermédia comme HyperODA[ISO 92] et Hytime[HYT 92].

- **spécifique à un document** : à ce niveau, la description des différentes relations multimédia est spécifique à un seul document. Les relations logiques, spa-

tiales, temporelles et hypermédia sont spécifiées entre les objets contenus dans le document.

- **objet multimédia de base** : à ce niveau, on s'intéresse plus à la description du contenu des objets multimédia de base. Chaque objet est décrit par un ensemble d'attributs qui permettent de définir tous les aspects liés à sa présentation (la durée, la vitesse de défilement, la taille à l'écran, le format, la résolution à l'écran, la langue, etc.).

- **implémentation et protocole d'accès aux objets** : à ce niveau, on décrit la façon de retrouver le contenu d'un objet multimédia en termes d'entités actives du système (processus) et de protocoles d'accès (système de fichier local, HTTP[BER 92], MPEG[LEG 91], JPEG[WAL 91], etc.).

Les modèles de documents multimédia existants couvrent uniquement une partie de ces différents niveaux. Notre approche de l'édition de documents multimédia tente d'intégrer ces niveaux au sein d'un même modèle. On appellera dans la suite ce type de modèles **les modèles de documents structurés multimédia**.

Ce type de modèles a déjà été mis en œuvre pour les documents classiques [QUI 86] [LAM 86]. La nouveauté vient des objets multimédia temporisés qui sont des entités dynamiques dont la présentation s'effectue à travers le temps. Lorsque de tels éléments sont regroupés par des relations temporelles, ils permettent de construire à l'intérieur du document des structures temporelles plus complexes (regroupement de portions de vidéo en scènes, leur synchronisation avec du son et avec les interactions de l'utilisateur etc.). Ceci ajoute à la structure traditionnelle d'un document une dimension structurale supplémentaire : *la dimension temporelle*. Ce dernier aspect place au cœur de la problématique des systèmes d'édition multimédia, la définition d'un modèle abstrait permettant de rendre compte d'une part de sa dynamique, et d'autre part, des autres dimensions (logique, spatiale et hypermédia).

3. Le modèle de document multimédia de MADEUS

Dans cette section, nous présentons le modèle de documents de MADEUS. Ce modèle décrit les différentes structures d'un document en s'appuyant sur une représentation générique des documents. Dans cette représentation des documents, les différents éléments qui le composent sont représentés par leur structure logique, c'est-à-dire par les composants (sections, paragraphes, notes, etc.) et les relations entre ces composants (inclusions, ordre, référence, etc.). Pour chaque type de document, les composants disponibles et leurs relations possibles sont définis par un *schéma de structure* [QUI 86]. L'éditeur est guidé par ce schéma et aide l'utilisateur à produire un document conforme au modèle du schéma.

Dans l'éditeur, la structuration est utilisée pour décharger l'utilisateur des opérations liées à la présentation graphique et à la synchronisation temporelle. A chaque type de document sont associés un ou plusieurs schémas de présentation et de synchronisation qui spécifient les règles de structuration spatiale et temporelle des documents de ce type. C'est en appliquant ces règles que l'éditeur reproduit la

forme dynamique du document à travers laquelle l'utilisateur interagit lors de la présentation (l'exécution).

L'édition structurée permet des opérations puissantes sur les documents multimédia traités, comme le positionnement automatique des objets les uns par rapport aux autres à travers le temps et sur les différents canaux. Le principe de la structuration s'applique aussi bien à l'organisation générale du document qu'à toutes les structures plus spécifiques qu'il contient. Cela permet d'intégrer et de manipuler de façon homogène le document lui-même et les parties qu'il contient.

Au delà de la compacité apparente d'un document multimédia en tant qu'unité sémantique, les relations qui regroupent ses composants déterminent différentes dimensions, orthogonales mais parfois dépendantes, de sa structure globale [LAY 95]. La synchronisation d'un document multimédia est souvent interprétée comme l'ensemble des mécanismes permettant de décrire l'ensemble des relations d'ordre logique, spatial, hypermédia et temporel. Dans le modèle de document de MADEUS, l'originalité vient de la forte intégration de tous ses aspects. Bien que l'accent soit mis sur l'aspect temporel, nous expliquerons systématiquement, dans la suite de cet article, les liens qui existent entre les différentes dimensions.

3.1. Dimension logique

Dans un document multimédia, il est possible de distinguer des parties du document qui sont sémantiquement liées et qui peuvent être regroupées par le biais de relations logiques. Ce regroupement est un objet composite qui constitue une nouvelle entité dont la structure interne, tant sur l'aspect logique, spatial que temporel est spécifiée indépendamment du reste du document (cf. exemple de la Figure. 1).

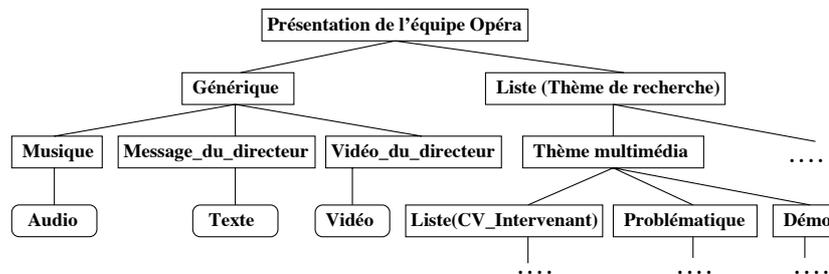


Figure 1. Structure logique d'une présentation multimédia.

De telles entités logiques peuvent être également groupées pour constituer d'autres entités logiques plus complexes. Par exemple, pour construire un document multimédia présentant une équipe de recherche, on peut commencer par une scène de bienvenue de son directeur, ensuite regrouper dans un ensemble de scènes les thèmes de recherche de l'équipe. La présentation de chaque thème étant elle-même

composée par la liste des intervenants décrits par un scénario curriculum vitae et ainsi de suite.

Ce type de construction décrit des relations d'inclusion entre les éléments et définit ainsi une classe de documents. Par exemple, les activités de recherches d'un institut peuvent être décrites de la même façon que dans l'exemple de la Figure. 1. Cette description générique des documents est définie par une grammaire basée sur la notion de constructeurs (agrégats, liste, choix, etc.) pour décrire un schéma de structure. Le schéma de structure correspondant à l'exemple précédant s'écrit de la façon suivante :

```
Présentation = BEGIN
  Introduction = Générique;
  Corps = LIST OF (Thème_de_recherche);
  Fin_Présentation = ... ;
END;
Générique = BEGIN
  Musique = AUDIO;
  Message_du_directeur = Message;
  Vidéo_du_directeur = VIDÉO;
END;
Message = CASE OF (
  TEXTE;
  AUDIO;)
END;
Thème_de_recherche = BEGIN
  Schéma = LIST OF (CV_intervenant);
  Problématique = ... ;
END;
```

3.2. Dimension spatiale

Cette dimension concerne l'affectation dans le temps des composants multimédia d'un document aux ressources physiques (l'espace). Par exemple, un objet vidéo doit disposer d'un espace géométrique sur l'écran délimité par une certaine zone. De la même façon, un objet audio doit disposer d'un canal audio pour toute la durée de sa présentation. Cette opération de mise en correspondance, appelée formatage spatio-temporel, tient compte de différents paramètres du document qui peuvent dépendre aussi bien de sa structure temporelle que de sa structure logique. Cette opération peut être appliquée au niveau générique d'un document multimédia en définissant des règles de présentation spatio-temporelles. Par exemple, dans le document de la Figure. 1, l'auteur peut définir une règle de présentation qui impose de libérer tous les objets affectés à la fenêtre du document après chaque fin d'un objet *thème de recherche* présenté. Il est important de souligner que cette règle est attaché au type *thème de recherche* et non pas à une instance de ce type. Les règles de pré-

sentation spatio-temporelle sont donc génériques, et peuvent ainsi réduire de façon significative le nombre de spécifications requises pour la construction d'un document.

Le formatage spatio-temporel généralise celui du formatage géométrique dans les éditeurs de documents statiques [ROI 94][LAM 86]. Dans ce type de documents, le formatage est restreint au positionnement géométrique (sur papier ou sur écran) d'un document à partir de sa structure logique et de définitions de règles géométriques (taille de la page, des colonnes, etc.) ainsi que des caractéristiques des différents objets du document (taille, fontes, etc.). Dans le cas de documents multimédia, la nouveauté vient de la dimension temporelle des données qui ajoute un aspect dynamique à cette opération. En effet, les espaces géométriques et les canaux sonores requis pour la présentation des objets multimédia sont alloués, libérés, mixés pour le cas de l'audio au fur et à mesure de la présentation du document global [OGA 92]. L'exemple de la Figure. 2 montre la scène correspondant au générique de début de la présentation de l'équipe de recherche. La vidéo du directeur (en bas à gauche de la première fenêtre) ainsi qu'un graphique d'annotation de l'image qui accompagne son message sont présentés au début de la présentation. Ensuite, dès que la vidéo termine, elle disparaît de l'écran en même temps que le graphique d'annotation de l'image.



Figure 2. Exemple de présentation multimédia (scène du générique).

De manière générale, à chaque élément du document on fait correspondre un intervalle temporel qui correspond à la durée de temps pendant laquelle il est actif et occupe les canaux physiques de la machine (l'affectation au canal audio pour le son, la carte de décompression et la fenêtre de présentation pour la vidéo, ou uniquement la boîte associée pour les objets statiques de MADEUS). Toutefois, cette règle n'est pas générale et ne permet pas de décrire toutes les situations. Pour décrire de façon plus complète la dynamique d'un objet multimédia dans l'espace et dans le temps, il faut distinguer quatre instants **spatio-temporels** clés pour chaque objet du document. Ces instants permettent de définir de manière précise les relations qui existent

entre les deux dimensions spatiale et temporelle des différents objets. Ces instants, qui sont notés, (*map*, *start*, *finish*, *unmap*), correspondent respectivement à l'affectation d'un objet à une ressource, à son démarrage temporel, à la fin de sa présentation et à la libération de la ressource qu'il occupe. Ces règles s'appliquent à tous les objets visuels qu'ils soient du texte, des graphiques ou de la vidéo. Par contre, pour les objets sonores, les instants *map* et *start* ainsi que *finish* et *unmap* sont temporellement confondus. De la même façon que pour les objets multimédia de base, les entités logiques possèdent aussi des instants spatio-temporels similaires. L'instant *map* d'une entité logique correspond à un ou plusieurs instants *map* des objets multimédia visuels de base qu'il contient.

3.3. Dimension hypermédia

La structure logique présentée dans la section 3.1 est fondamentalement hiérarchique. Cette représentation n'est pas adaptée pour décrire l'ensemble des relations qui peuvent exister entre les éléments d'un même document ou entre des documents différents. Dans MADEUS, ces liens sont décrits par des objets particuliers du document qui permettent de relier des portions de différents documents indépendamment de leur position dans cette structure. Ces liens dits hypermédia permettent de définir des relations de type sémantique entre des documents ou entre des parties de documents, comme les renvois ou les références. Ils mettent en place un réseau de portions de documents dépassant les structures qui les englobent et constituent ainsi un support pour la navigation dans un grand espace d'informations à la manière des hypertextes ou du *World Wide Web* [BER 92]. Un lien est défini par une ancre de départ (un élément), une ancre d'arrivée (un autre élément) et le lien qui porte une certaine sémantique [NAN 93] (renvoi, référence, annotation, etc.). Au niveau de l'interface utilisateur, un lien nécessite d'être activé explicitement par une interaction de l'utilisateur.

Les liens hypermédia dans les documents statiques, trouvent aussi leur équivalent dans la dimension temporelle. En effet, l'analogie dans le temps des liens statiques est définie par des éléments particuliers (les liens hypermédia) du document qui permettent de relier des documents ou des portions de documents dans l'espace et dans le temps [LIE 94]. Par exemple, une partie d'un objet vidéo délimitée dans l'espace (une zone particulière de la vidéo) et dans le temps (la période de son apparition à l'écran) peut constituer un lien vers un nouveau document. L'activation du lien, comme dans les documents statiques, reste basée sur une interaction explicite de l'utilisateur. La seule différence est que, d'une part l'ancre de départ n'est activable que pendant le laps de temps couvert par sa présentation, et d'autre part l'ancre d'arrivée correspond non seulement à un autre élément ou document mais aussi à une date précise de sa présentation (début, fin, etc.) [CHE 95].

3.4. Dimension temporelle

La caractéristique majeure d'un document multimédia est sa dimension temporelle. Quelle que soit leur granularité, les objets d'un document multimédia sont reliés temporellement de façon à définir un ordre global de présentation. Un modèle de synchronisation temporelle pour les documents multimédia doit principalement représenter ces dépendances.

La Figure. 3 illustre un exemple de composition temporelle. Celle-ci montre à partir de l'exemple de la Figure. 1 la combinaison temporelle des objets multimédia pour constituer l'ordre temporel de la présentation de l'équipe de recherche. Les boîtes de la Figure. 3 représentent des objets multimédia qui consistent en des images fixes (les icônes et le texte) ou des flots continus comme les images animées (vidéo du directeur, des participants et des scènes des thèmes de recherche), ou encore des flots sonores (musique du générique, voix du directeur et des intervenants). La synchronisation de ces différents éléments par le biais de relations temporelles permet de construire une structure temporelle indépendante pour chaque partie du document formant une entité logique (cf. 3.1).

Dans la dimension temporelle, contrairement à la dimension spatiale, on s'intéressera uniquement aux instants *start* et *finish*, car les instants *map* et *unmap* correspondent à d'autres événements d'objets de base *start* et *finish* de la présentation. Ceci explique l'interaction entre le temps et l'espace et nous permet de faire coexister les aspects liés à l'espace et les aspects liés au temps au sein d'une même structure du document.

Quel que soit l'élément auquel il se rapporte, chaque intervalle possède dans la dimension temporelle deux instants extrêmes de début et de fin. Dans notre modèle, on distingue trois types d'intervalles temporels :

- **Les intervalles des éléments logiques du document** : ce sont des intervalles abstraits qui délimitent respectivement l'instant de début du premier élément de base contenu dans cette entité logique, ainsi que l'instant de fin du dernier élément de base de cette entité. Le premier et le dernier élément correspondent dans ce cas à l'ordre temporel et non pas à l'ordre de représentation dans l'arborescence logique du document.

- **Les intervalles de style associés aux éléments** : ces intervalles sont ajoutés pour des besoins de contrôle du style d'une présentation multimédia. La notion de style est relative à chaque média (changement du volume pour le son, changement de couleur du texte au cours de la présentation, défilement de texte, trajectoire d'un objet sur l'écran etc.). Ces intervalles sont reliés aux éléments de la structure logique qu'ils permettent de mieux faire ressortir. Par exemple, deux secondes avant la fin de la présentation de l'élément logique *Démo* de la Figure. 3, on voudrait baisser progressivement le volume sonore jusqu'au niveau zéro, ou encore faire défiler progressivement le contenu de la fenêtre pour laisser place à la scène suivante.

- **Les intervalles d'espacement et de rupture temporels** : ce type d'intervalle permet d'insérer entre les intervalles des éléments logiques du document des inter-

valles d'espace temporel. Un tel espace peut correspondre à un intervalle de durée fixée entre un élément et son successeur (un délai) ou encore à un intervalle de durée indéterminée correspondant à l'activation explicite d'un bouton "suite" par l'utilisateur.

Ces deux derniers types d'intervalles sont importants pour affiner la présentation d'un document multimédia. Ils se distinguent de la première catégorie d'intervalles par le fait qu'ils n'ont pas de représentation dans la structure logique du document.

La description de la structure temporelle d'un document multimédia doit également intégrer des objets logiques ou de base aux propriétés temporelles très diverses. En effet, un document multimédia peut contenir des objets ayant les comportements temporels suivants :

- 1) Des *objets intemporels*, comme les éléments textuels et graphiques.
- 2) Des *objets déterministes*, comme les éléments audio et vidéo de durée connue ainsi que les délais.
- 3) Des *objets indéterministes*, comme les programmes, les flots vidéo de durée indéterminée lorsque, par exemple, on y a accès par le biais du réseau ou qu'ils sont calculés en cours de présentation (images de synthèse, réalité virtuelle, bouton d'interaction avec l'utilisateur).

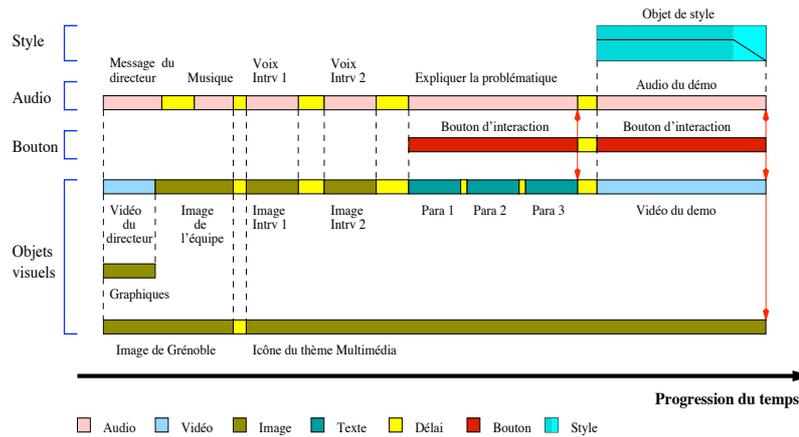


Figure 3. Exemple de composition temporelle.

4. Contraintes temporelles dans les documents multimédia

Le modèle temporel est un aspect central du modèle de documents multimédia. Il permet de définir les méthodes permettant le support de la structure temporelle du document. Il doit donc rendre compte de toutes les dépendances temporelles entre les composants de ce document. La présentation, et donc l'ordonnement temporel du document, découlera de ces dépendances. C'est cette structure qui détermine la synchronisation des composants textuels, sonores et visuels du document. Au ni-

veau d'une application d'édition multimédia, un modèle temporel est mis en œuvre par trois niveaux :

- **Un langage temporel** qui implante les abstractions du modèle et dont le rôle est l'expression des dépendances temporelles entre les différents éléments du document. Le langage définit l'interface de spécification des contraintes temporelles entre l'auteur d'une part, et le document (via l'interface d'édition) d'autre part. La synchronisation à ce niveau est décrite uniquement en fonction de la structure des documents et du comportement temporel que l'on souhaite obtenir avec une plateforme idéale (ressources infinies, et comportement idéal des objets multimédia). Il s'agit donc d'une représentation abstraite indépendante d'une plateforme spécifique et pouvant être implantée sur des systèmes d'exécution différents.

- **Un système de résolution de contraintes temporelles** : ces contraintes sont induites par les relations temporelles spécifiées par l'auteur, par la nature des objets mis en jeu et par la disponibilité des ressources nécessaires à la présentation (cartes de décompression, nombre de canaux audio, etc.). Le rôle de ce module est de définir les règles de calcul des durées entre les objets temporisés, non-temporisés et indéterministes ainsi que la vérification de la cohérence globale du document. Le problème posé à ce niveau concerne la mise en place de méthodes de résolution permettant la prise en compte de ces différentes contraintes.

- **Une machine d'exécution** : qui supporte le langage et qui utilise une forme dite exécutable du document. Cette machine virtuelle, appelée dans MADEUS *médiateur* d'exécution est un module de l'éditeur qui nécessite une expression cohérente de sa structure temporelle du document. Le document est décrit à ce niveau par des règles de transition qui correspondent aux instants (*map*, *start*, *finish* et *unmap*) décrits précédemment. Ces règles de transition sont définies en fonction du type des objets et de leur comportement temporel. Pour les objets visuels, les règles de transition permettent aussi de définir les moments choisis pour leur allocation (*map*) et de leur libération (*unmap* de l'écran) par l'auteur.

La progression du temps au cours de la présentation du document (génération des différents événements spatio-temporels) est portée par deux types d'événements : les événements correspondants aux objets multimédia de base et les événements correspondant aux éléments logiques du document. Le premier type d'événements se fait par une interaction directe entre l'éditeur et l'implantation au niveau support d'exécution des objets de base (entités actives de la présentation du document). Par contre, pour les entités logiques du document, ces instants sont définis en fonction de la structure logique et temporelle du document (instants calculés). Par exemple, la fin de la présentation de l'élément logique scène n'est définie que par rapport à la fin du dernier élément (dans la dimension temporelle) qu'il contient.

4.1. Langages temporels et la représentation du temps

Le rôle d'un modèle temporel étant l'expression des dépendances temporelles entre objets multimédia, l'adoption d'une représentation appropriée du temps est

nécessaire. Le modèle temporel est déterminé par la façon de décrire l'unité temporelle de base, la sémantique associée à cette unité temporelle ainsi que l'expression des relations temporelles entre différentes unités. Il existe deux façons de représenter une unité temporelle de base :

1) Une représentation fondée sur les **instants**. Dans ce cas, un objet multimédia qu'il soit logique ou de base est déterminé par un instant de début et un instant de fin. Les dépendances temporelles entre les différents objets consisteront en des relations entre ces instants.

2) Une représentation fondée sur les **intervalles**. Un objet multimédia est considéré, dans cette représentation, comme une entité temporelle de base. Les dépendances temporelles entre les composants du document multimédia sont exprimées par des relations appliquées aux intervalles.

Au delà de cette différence de représentation, les relations possibles entre deux objets multimédia se réduisent à toutes les combinaisons de positionnement possibles de deux intervalles sur une droite orientée. Le modèle le plus général, proposé par J.F. Allen [ALL 83], dresse la liste exhaustive de toutes ces relations (cf. Figure 4). La liste des combinaisons possibles entre objets multimédia comporte 13 relations consistant en 7 relations de base et de leurs relations inverses, l'égalité étant elle-même son inverse. Ces treize relations se répartissent en deux classes, celle des relations de séquentialité et celle des relations introduisant le parallélisme de présentation. Dans l'avant-dernière colonne du tableau, nous présentons la traduction de chaque relation sous la forme de relations à base d'instant (x^- et x^+ désignent respectivement l'instant de début et de fin d'un objet x).

D'autres constructeurs d'objets sont utilisés dans MADEUS, comme l'opérateur d'itération ($\text{loop}(n,x)$) où n est le nombre d'itérations de la présentation d'un objet x), la projection temporelle d'un objet et les relations mettant en œuvre la causalité entre les événements. Ces constructeurs d'éléments et ces relations ne sont pas présentés ici car ils n'influent pas sur le modèle temporel à base d'intervalles.

Relation	Classe	Contraintes	Valeurs de t
$x \text{ before}(t) y$	Seq	$x^+ < y^-$	$]0, \infty]$
$x \text{ equals } y$	Par	$x^- = y^- \wedge x^+ = y^+$	-
$x \text{ meets } y$	Seq	$x^+ = y^-$	-
$x \text{ overlaps}(t) y$	Par	$x^- < y^- \wedge x^+ > y^- \wedge x^+ < y^+$	$]0, \text{dur}(x)]$
$x \text{ during}(t) y$	Par	$x^- > y^- \wedge x^+ < y^+$	$]0, \text{dur}(y) - \text{dur}(x)]$
$x \text{ starts } y$	Par	$x^- = y^-$	-
$x \text{ finishes } y$	Par	$x^+ = y^+$	-

Figure 4. Les sept relations temporelles.

Dans MADEUS, contrairement à plusieurs systèmes multimédia [BUC 93] [DRA 93], les deux manières de décrire les relations temporelles (à base d'instant ou à base d'intervalles) correspondent à deux niveaux de représentation du

document, plutôt qu'à un choix entre l'un ou l'autre. Celle à base d'intervalles est plus adaptée pour l'interface auteur (représentation symbolique), dans la mesure où les opérations de base de l'édition, comme l'adjonction et la suppression d'une contrainte temporelle, consiste à sélectionner deux objets multimédia ainsi que la relation qui les lie. Par contre, la représentation à base d'instantants convient mieux un support d'exécution (représentation interne) où la synchronisation des entités actives du système est basée sur les événements [BLA 94]. De plus, la correspondance entre les deux niveaux est d'ordre syntaxique et s'exprime de manière directe selon le tableau précédent.

Nous avons vu dans la section 3.4 que chaque élément du document, qu'il soit logique ou de base, est associé à un intervalle abstrait délimitant son instant de début et de fin. L'instant de référence d'un objet quelconque du document est, par défaut, celui de son père selon la structure logique. Ensuite, au fur et à mesure de l'introduction de nouvelles contraintes et de nouveaux objets à ce même niveau de l'arbre logique, l'ordre entre les objets est déterminé de manière implicite. Dans MADEUS, nous utilisons une combinaison de trois types de représentation du temps qui nous permettent d'exprimer plusieurs situations :

- **Une représentation à base d'information temporelle quantitative :** Dans ce cas, la représentation des objets ou des relations temporelles fait référence de manière directe ou indirecte à une mesure du temps physique. Par exemple, pour l'objet musique du document de la Figure. 3, qui est situé au début de la présentation, sa durée et son axe temporel de référence déterminent de manière implicite sa relation temporelle avec les autres objets. Si l'on restreint les relations et les objets multimédia du document à ce type de représentation du temps, l'ordre de positionnement temporel des différents objets document est total.

- **Une représentation à base d'information temporelle qualitative :** Le positionnement des objets reliés par de telles dépendances peut définir un ordre partiel. Cette représentation du temps permet de positionner les objets multimédia les uns par rapport aux autres sans imposer des contraintes précises sur leur durée ou sur les délais associés aux relations (voir Figure. 4). Par exemple, si l'auteur spécifie une relation (*before*) entre deux objets sans définir la durée du délai qui les sépare, le positionnement du deuxième objet par rapport au reste du document n'est pas déterminé de manière précise.

- **Une représentation à base d'information temporelle causale :** Dans ce cas, la représentation des objets ou des relations possède une sémantique de causalité entre les événements temporels [LAM 78]. Par exemple, lors de la synchronisation d'une vidéo avec du texte par une relation d'égalité temporelle, la fin du texte proprement dite est causée par la fin de la vidéo. De la même façon, l'objet bouton de la scène problématique, s'il est activé, cause la fin de toute la scène problématique (sens des flèches de la Figure. 3).

La représentation causale du temps permet de modéliser les interactions de l'utilisateur, car ces interactions peuvent agir sur le comportement temporel du document. Les interactions de l'utilisateur correspondent dans notre modèle aux intervalles où elles sont activables. Ainsi, le choix de l'utilisation de l'intervalle comme

principale abstraction pour représenter tous les éléments d'un document semble la plus appropriée.

4.2. Réseau de contraintes temporelles

Un réseau de contraintes temporelles est une représentation interne du document extraite à partir des relations temporelles introduites par l'auteur. Ce réseau est défini par un graphe acyclique orienté (*DAG*) composé d'un certain nombre de nœuds et d'arcs. Chaque nœud de ce graphe représente un instant temporel où un ou plusieurs intervalles commencent ou se terminent. Chaque arc représente un intervalle dont les extrémités correspondent au début et à la fin de la présentation d'un objet multimédia ou d'un délai. De plus, cet arc est étiqueté par la contrainte de durée associée à l'objet ou au délai qu'il représente.

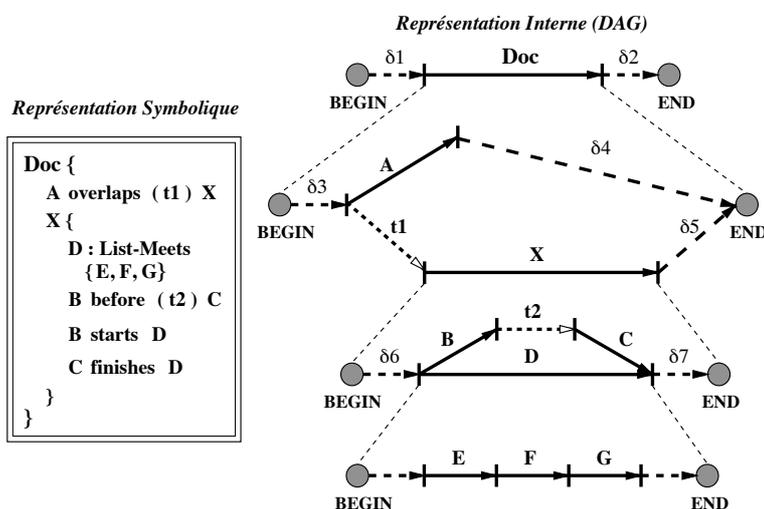


Figure 5. Représentation symbolique et interne (DAG) de MADEUS.

Deux instants abstraits *BEGIN* et *END* sont attachés aux entités logiques du document pour délimiter leur début et leur fin. Ceci permet de définir pour chaque entité logique un sous-réseau de contraintes qui est construit de la même façon que celui du document lui-même. Ainsi, la résolution des contraintes temporelles pour tout un document est réduite à sa résolution pour un seul niveau de la hiérarchie logique.

4.2.1. Contrôlabilité et flexibilité des présentations temporelles

Dans la classification introduite dans la Section 3.4, nous avons fait la distinction entre les objets multimédia de base de type discret, continu et indéterministe. Au niveau de notre modèle de contraintes, tous ces types d'objets

ainsi que les délais peuvent être représentés par deux types d'intervalles. Chacun d'eux possède une borne inférieure et une borne supérieure et diffèrent fondamentalement selon les définitions suivantes :

Définition 1 : *Un intervalle libre* est un intervalle ayant une contrainte numérique sur la durée $[l, u]_l$ (où $l < u$, l et u sont des valeurs positives et u pouvant valoir l'infini), et cette contrainte peut être réduite librement entre l et u (*lower bound et upper bound*). Ce type d'intervalles permet de modéliser le comportement des objets de type continu et discret dont la durée de présentation peut être choisie dans un certain intervalle. Cette durée est bornée pour les objets continus par des limites liées à l'intelligibilité de leur contenu, mais pour les objets discrets comme le texte et les graphiques, elle peut être infinie. Ce type d'intervalles présente un comportement temporel de type élastique [HAM 93].

Définition 2 : *Un intervalle contingent* est un intervalle ayant une contrainte numérique sur la durée $[l, u]_c$, et cette contrainte n'est connue qu'à l'exécution et peut prendre une valeur aléatoire entre l et u , u pouvant valoir l'infini. Ce type d'intervalles permet de modéliser le comportement indéterministes de certains objets multimédia dont la durée de présentation ne peut pas être connue pendant la phase d'édition. Par exemple, la durée de l'intervalle correspondant à un bouton d'interaction avec l'utilisateur ne dépend que de l'instant de cette interaction.

D'après ces deux types d'intervalles, nous pouvons distinguer deux concepts qui caractérisent le scénario temporel d'un document multimédia : la flexibilité et la contrôlabilité. La flexibilité mesure l'aptitude d'un scénario temporel à s'adapter aux spécifications introduites par l'auteur tout en garantissant sa cohérence globale. La flexibilité d'un scénario dépend du nombre de délais libres et des objets multimédia élastiques présents dans le document et de leur position dans la structure temporelle. Dans le cas de MADEUS, cette flexibilité du document est utilisée pour compenser le comportement indéterministe qu'il contient. Cela introduit la notion de contrôlabilité qui mesure l'aptitude d'un scénario à s'adapter dynamiquement (à l'exécution) au comportement indéterministe des objets multimédia.

Le modèle de contraintes de MADEUS se distingue par le fait qu'il permet de décider pendant la phase d'édition si un document est contrôlable ou pas, et nous a permis d'étendre la notion de cohérence temporelle aux documents intégrant un comportement indéterministe [LAY 96].

4.2.2 Structure d'un réseau de contraintes temporelles

Le graphe décrit précédemment est construit et modifié au fur et à mesure de la session d'édition. Cette construction se fait progressivement en insérant des objets multimédia dans le document et par la spécification de nouvelles relations ou de nouvelles contraintes sur les intervalles. Le positionnement des objets dans la structure globale est réalisé en suivant deux types de règles :

- Les règles liées à la position de l'objet dans la structure logique du document. Ces règles permettent de positionner par défaut l'objet dans la structure temporelle du document en utilisant comme référence, les instants de début de son père ou de son frère selon la structure logique. Dans l'exemple de la Figure. 1, le constructeur

liste lié au type *Thème Multimédia* du schéma de structure est associé à une règle de positionnement en séquence de ces éléments *Liste_Before(CV_intervenant_i)* ($1 < i < n$), et l'élément logique *List_CV_intervenants* est positionné par rapport à son élément englobant *Thème Multimédia*, et ainsi de suite.

- Si l'objet n'a pas de position spécifiée par défaut, il est inséré dans la structure temporelle par rapport aux instants début et fin de l'élément englobant de la structure logique par la relation temporelle *contains*.

Les différentes opérations de modification de la structure temporelle du document ont généralement une portée locale. Cette portée est délimitée par les instants de références associés aux objets selon leurs types. Les modifications introduites dans la structure temporelle susceptibles d'avoir lieu après chaque type d'opération d'édition sont détaillées dans [LAY 96]. Ces différentes opérations permettent la construction incrémentale de la structure du graphe par des mises à jour locales. L'utilisation d'un langage d'opérateurs temporels pouvant se transformer syntaxiquement en graphe [GHA 93], nous permet d'avoir une structure temporelle qui est organisée en un ensemble de chaînes temporelles qui sont définies de la façon suivante :

Une chaîne temporelle $[i_1, i_2, \dots, i_n]$ est une séquence d'intervalles contigus tels que chaque intervalle i_j , $1 < j < n$ est relié à un seul intervalle prédécesseur et successeur tel que $begin(i_j) = end(i_{j-1})$ et $end(i_j) = begin(i_{j+1})$, chaque intervalle pouvant être de type libre ou contingent.

4.3. Cohérence temporelle des documents multimédia

Dans MADEUS, la structure temporelle d'un document multimédia est décrite par un ensemble de graphes qui correspondent chacun à un niveau de la hiérarchie logique. L'auteur construit les différentes dépendances temporelles qu'il relie progressivement en spécifiant des relations entre elles. Il se peut toutefois que l'adjonction d'une nouvelle relation introduise dans le document des incohérences en violant certains invariants temporels.

4.3.1 Problème de cohérence temporelle

La modification d'un document multimédia par l'insertion ou la suppression de nouveaux objets ou de nouvelles relations ajoute au niveau du graphe de nouvelles contraintes. Ces contraintes peuvent conduire à une incohérence car les paramètres temporels des objets et des relations (durées des objets et des délais) doivent vérifier certains invariants liés d'une part à la progression du temps (causalité entre les différents événements) et d'autre part à la coïncidence dans le temps des instants de fin des chaînes temporelles concurrentes (aspect quantitatif).

Supposons qu'une partie d'un document soit composée de trois objets a, b et c (voir. Figure. 6-a), et que ces objets soient liés par les relations (a *meets* b) et (b *meets* c). L'adjonction d'une nouvelle spécification c *overlaps* a rend le scénario incohérent. Ce cas correspond à une incohérence de type qualitatif car elle ne dépend

pas des durées des objets mais de la sémantique même des relations mises en jeu. Par contre, un scénario qui peut être cohérent du point de vue qualitatif peut être quantitativement incohérent. Par exemple (voir Figure. 6-b), si l'auteur spécifie que (*a meets b*) et (*a starts c*), l'adjonction de la relation (*c overlaps b*) rend le scénario incohérent pour des durées de *a*, *b* et *c* valant respectivement 30, 20 et 20 secondes. Finalement, dans l'exemple de la Figure. 6-c, on reprend le même scénario précédent mais avec un objet *c* de type indéterministe dont la durée est bornée dans l'intervalle $[20,40]$. Ce cas correspond à une incohérence si, à l'exécution, la durée de l'objet *c* prend une valeur entre 20 et 30, par contre, le scénario sera cohérent si la valeur de cette durée est entre 30 et 40. Du point de vue du document, ce cas est considéré comme une incohérence car cette contrainte ne peut pas être garantie dans tous les cas.

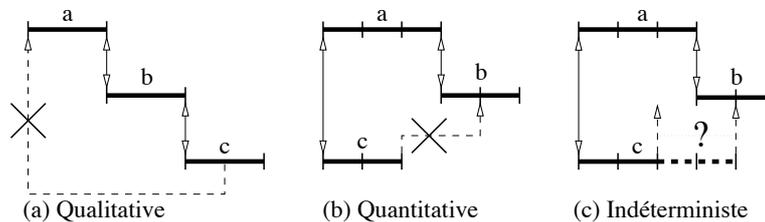


Figure 6. Exemples d'incohérences temporelles.

Dans des scénarios plus complexes, les incohérences ne sont pas aussi faciles à détecter que pour les trois exemples précédents. Dans l'approche structurée des documents multimédia, la structure logique arborescente du document réduit considérablement la taille du problème lié à la gestion de cette cohérence. En effet, les objets de base se trouvant au niveau des feuilles de l'arbre, une gestion ascendante de la cohérence par propagation des contraintes permet d'avoir un mécanisme efficace adapté au processus incrémental de l'édition [VIL 86].

4.3.2. Gestion de la cohérence temporelle dans MADEUS

Si on considère le graphe des contraintes décrit précédemment, les invariants de temps se résument sur ce graphe en deux propriétés suffisantes pour garantir la cohérence globale du document :

- Les circuits ne sont pas autorisés dans le graphe car les instants venant du passé ne peuvent pas rejoindre des instants du futur (mis à part pour les éléments liens qui possèdent une sémantique autorisant des sauts temporels).
- Les cycles du graphe sont toujours composés de deux chaînes temporelles parallèles (propriétés du *DAG*) dont les extrémités temporelles doivent être compatibles dans le temps. Cette compatibilité temporelle est traduite au niveau des chaînes temporelles par le fait que l'intersection de leurs domaines de solutions ne doit pas être vide. Cette propriété nous garantit que la coïncidence temporelle entre les extrémités des deux chaînes sera assurée à l'exécution.

Ces deux propriétés permettent d'identifier les trois types d'incohérences présentées dans la section 4.3.1. Avant de présenter le principe des solutions mises en œuvre dans MADEUS pour détecter ces incohérences, nous adoptons une hypothèse importante par rapport à l'observabilité temporelle d'un document multimédia à l'exécution. Cette hypothèse restreint les instants d'observation du comportement temporel du document aux extrémités de début et de fin des différents intervalles. Cette hypothèse nous permet de refléter le modèle d'exécution à base d'événements utilisé dans l'application MADEUS. Ces instants d'observation de l'état du document sont aussi à la base de notre politique de gestion du comportement indéterministe des objets, car ces instants nous permettent de décider s'il est possible ou non, pour chaque état observable du document, de corriger dynamiquement les paramètres temporels d'une présentation afin de garantir sa cohérence.

Nous présentons d'abord le principe de l'algorithme à travers deux exemples de scénarios. Dans le premier scénario (voir exemple 1 de la Figure. 7), un objet texte **A**, est mis en parallèle avec un objet bouton **B**. Ce scénario se traduit dans le graphe correspondant par deux arcs parallèles partant d'un même nœud de départ vers un même nœud d'arrivée. Ce scénario est consistant car la fin de l'objet **A** représenté par un intervalle infiniment élastique $[0, \infty]_I$, peut être indéfiniment retardé jusqu'à l'occurrence de l'événement fin de l'objet **B** qui est représenté par $[l, u]_C$ (durée déterminée à l'exécution par un clic de l'utilisateur sur le bouton).

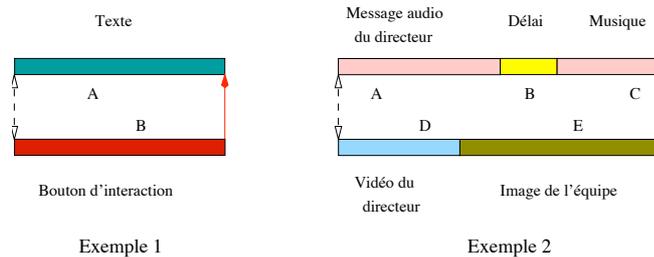


Figure 7. Exemples de gestion de la cohérence pour les objets indéterministes.

Dans le deuxième exemple de la Figure. 7, nous reprenons le scénario de début du générique de présentation de l'équipe de recherche. Ce scénario est composé de deux chaînes temporelles. La première chaîne est composée de trois objets **A**, **B** et **C** correspondant respectivement au message du directeur (audio), à un délai et à un morceau de musique. La deuxième chaîne est composée de deux objets **D** et **E** correspondant à la vidéo du directeur suivi de l'image de l'équipe de recherche. Si l'objet vidéo du directeur est modélisé par un intervalle contingent $[l_A, u_A]_C$ alors il est possible de compenser le montant d'indéterminisme correspondant à $(u_A - l_A)$ unités de temps de deux façons. La première consiste à utiliser le délai **B** modélisé par un intervalle libre $[l_B, u_B]_I$ si le montant de sa flexibilité correspondant à $(u_B - l_B)$ unités de temps est supérieur à l'indéterminisme : $(u_B - l_B) \geq (u_A - l_A)$. L'autre solution consiste à utiliser l'élasticité infinie de l'objet **E**.

Ceci nous mène au principe de l'algorithme de vérification de cohérence qui est basé sur une propagation locale des contraintes temporelles. Cet algorithme, décrit plus en détail dans [LAY 96], est mis en œuvre par plusieurs phases de vérification permettant d'éliminer d'abord les cas triviaux :

- L'algorithme vérifie en premier lieu si la nouvelle contrainte introduit un circuit dans le graphe. Si c'est le cas, une erreur est reportée à l'auteur du document. Cette vérification est réservée aux opérateurs de classe parallèle (voir tableau Figure. 4) qui sont les seuls susceptibles d'introduire ce type d'incohérences.

- Si la nouvelle spécification ne met en jeu que des chaînes de type déterministe, les contraintes liées à cette nouvelle spécification sont d'abord vérifiées par rapport aux contraintes locales de ces chaînes. Ensuite, elles sont vérifiées par rapport aux chaînes concurrentes par propagation progressive des contraintes et ainsi de suite.

- La dernière phase de l'algorithme correspond à l'introduction d'intervalles indéterministes dans le graphe. La méthode utilisée consiste à vérifier, pour chaque instant permettant d'observer la durée d'un intervalle indéterministe, qu'il est possible de le compenser. Cette compensation est basée sur l'utilisation de la flexibilité des objets élastiques dont les événements de début ou de fin se trouvent dans le futur par rapport à cet instant. S'il reste suffisamment de flexibilité pour permettre de faire coïncider les extrémités des chaînes temporelles concurrentes [DUB 93] [GHA 93] [LAY 96], alors la nouvelle contrainte est cohérente.

Cet algorithme permet de garantir que la spécification temporelle du document sera conforme, à l'exécution, aux contraintes spécifiées par l'auteur : c'est la forme dite exécutable d'un document multimédia.

4.4. *Formatage temporel*

La représentation d'un document sous forme de graphe de contraintes, lorsqu'elle est cohérente, permet de produire une description de bas niveau de sa présentation. Cette description du document indique au module de présentation (d'exécution) du système multimédia, de manière précise, les dates des événements qui correspondent au lancement de leurs représentants actifs (entités d'exécution). Pour pouvoir calculer ces dates, il est nécessaire de définir une stratégie permettant de distribuer les contraintes locales aux différentes chaînes définies lors de la vérification de la cohérence. Les différents choix possibles de répartition des contraintes entre les objets multimédia et les délais temporels sont discutés dans [LAM 78]. Dans MADEUS, on se base en priorité sur la modification de l'élasticité des délais dont les durées sont laissées libres par l'auteur. Ceci nous permet d'affecter le moins possible la durée des objets multimédia. La phase permettant la production de cette forme du document s'appelle le *formatage temporel* et est effectuée par un module du système multimédia appelé *formateur temporel*. C'est la dernière étape qui précède l'exécution du document.

La Figure. 8 expose un exemple où les propriétés de contraction d'objets sont exploitées pour satisfaire une contrainte de synchronisation. La dilatation des objets

A et B permet de synchroniser, dans la présentation, l'objet résultant A *meets* B avec un objet rigide C par la relation *equals*. De manière plus générale, le formateur combine ces informations (classes et taux de contraction et dilatation éventuels des objets) pour permettre une synchronisation plus flexible de la présentation. En particulier, il pourra par ce biais, satisfaire des contraintes entre objets qui sont a priori impossibles en vertu de leurs durées initiales de présentation.

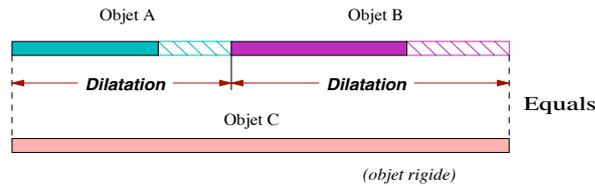


Figure 8. Exemple de formatage temporel.

Il est à noter que les opérations de contraction et de dilatation ne sont pas seulement applicables à des objets de base mais, conformément au regroupement des objets en entités logiques (voir section 3.1), à des portions entières de document. Un processus de formatage qui tient compte de ces aspects est proposé dans le système *Temporal Glue* [HAM 93].

5. Étude de quelques systèmes multimédia

Nous détaillons dans cette section l'état de l'art en matière de modèles temporels de documents multimédia. Comme nous l'avons indiqué précédemment en 4.1, ces modèles se répartissent en deux grandes familles qui diffèrent suivant leur façon de représenter l'information temporelle de base d'un document multimédia.

5.1. Modèles fondés sur les instants

Nous présentons dans cette section les plus courants des modèles temporels fondés sur les instants utilisés par les systèmes multimédia.

5.1.1. Le modèle à axe de temps absolu (time-line)

Ce modèle découle naturellement du choix d'une représentation quantitative du temps [DRA 93][DIR 95] [HYT 92]. Il consiste, en effet, à positionner par rapport à un axe de temps unique les dates des instants de début des objets multimédia. Outre ce positionnement temporel, les objets multimédia sont répartis selon un axe orthogonal sur des positions discrètes représentant les différents canaux physiques disponibles (cf. Figure. 9).

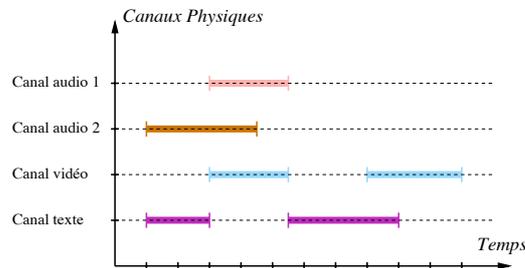


Figure 9. L'axe de temps absolu.

L'avantage de ce modèle est qu'il permet, en mettant à plat la répartition des objets multimédia sur les canaux physiques, d'éviter toute collision en interdisant le recouvrement de segments. Son emploi est, en outre, très intuitif pour l'auteur. Il présente toutefois plusieurs inconvénients :

- Toute insertion d'un nouvel objet impose le repositionnement manuel de tous les objets déjà positionnés qui en sont temporellement dépendants.
- À cause de l'hypothèse du temps absolu adoptée dans ce modèle, il n'est pas possible de tenir compte d'intervalles indéterministes.

5.1.2. Les réseaux de points (temporal point nets)

Ce modèle, mis en œuvre dans le système *Firefly* [BUC 93] considère trois types de relations entre les points : la simultanéité, la précédence et la succession. Le document multimédia est dès lors représenté par un graphe relationnel dont les nœuds sont les instants de début et fin de composants multimédia et dont les arcs joignent les points mis en relation (cf. Figure. 10).

Le système *Firefly* utilise l'algorithme du *simplex* pour effectuer le formatage temporel. La description temporelle du document est traduite sous forme d'équations et d'inéquations linéaires portant sur des variables correspondant aux instants de début et de fin des composants multimédia. Cet algorithme, en minimisant le temps entre les instants de début et de fin de la présentation globale, détermine les dates de tous les événements de début et de fin des objets multimédia du document. Il n'est donc pas possible d'intégrer dans un document des objets ayant des durées indéterministes.

L'un des principaux avantages de ce modèle est sa puissance expressive et la finesse de la synchronisation qu'il permet. Il est possible, en effet, de représenter dans ce modèle pratiquement tous les scénarios possibles de documents multimédia quelle que soit la complexité des dépendances temporelles qu'il contiennent. Cette représentation sous forme de graphe est utilisé directement comme interface utilisateur. Cette représentation étant basée sur les instants et n'étant pas structurée, elle a l'inconvénient d'être très peu lisible lorsque la taille du document augmente et l'opération de formatage devient de plus en plus coûteuse.

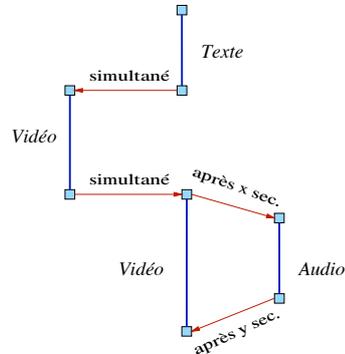


Figure 10. Les réseaux de points.

5.2. Modèles fondés sur les intervalles

Les modèles fondés sur les intervalles ne considèrent plus les instants de début et fin d'un composant multimédia, mais représentent celui-ci de manière plus compacte comme un intervalle de temps abstrait correspondant à l'exécution de sa présentation. Au lieu des graphes complexes et peu lisibles rencontrés dans les modèles fondés sur les instants, nous rencontrons dans ce cas des expressions plus structurées des contraintes temporelles. En particulier, certains modèles supportent, en plus de la structure temporelle, la structure logique du document.

5.2.1. Les réseaux de Petri temporisés (temporal Petri-nets)

Dans ce modèle, la présentation d'un document multimédia est vue comme l'évolution d'un système de transitions. Dans cet esprit, les réseaux de Petri peuvent être utilisés pour modéliser formellement le comportement temporel d'un document multimédia [LIT 93]. L'étape atomique comprise entre deux changements d'état du réseau de Petri correspond à la présentation atomique d'un certain nombre d'objets multimédia de base. Le document multimédia est dès lors modélisé de la manière suivante :

Une place du réseau de Petri représente un objet multimédia de base. Lorsqu'une telle place contient un jeton, la présentation de cet objet est exécutée pendant l'étape atomique inter-transitions. Les places sont aussi utilisées pour modéliser des portions de temps de durée fixée, vides de toute exécution de présentation, qui sont utilisées pour affiner la synchronisation. Dans ce cas, l'étape de transition comporte simplement la décrémentation d'un compteur qui mesure l'écoulement de cette durée. Les transitions du réseau de Petri, quant à elles, représentent des instants de début ou de fin d'objets multimédia de base.

La représentation des relations temporelles entre objets multimédia de base est supportée par une organisation particulière de la topologie du réseau (cf. Figure. 11).

L'utilisation des réseaux de Petri comme modèle de documents multimédia convient bien, d'après notre expérience, à la réalisation de l'ordonnancement (séquenceur) de l'exécution. Mais leur utilisation ne permet pas de résoudre les problèmes de cohérence, de gestion de l'indéterminisme et de formatage dont le principe est basé sur la durée des objets et non pas sur la propagation des événements de présentation (les transitions). C'est pour pallier cet inconvénient que certains modèles comme [IIN 94][LIT 93][SEN 96] préfèrent représenter le document en utilisant le modèle d'Allen mais en extraient une description sous forme de réseaux de Petri pour supporter l'exécution.

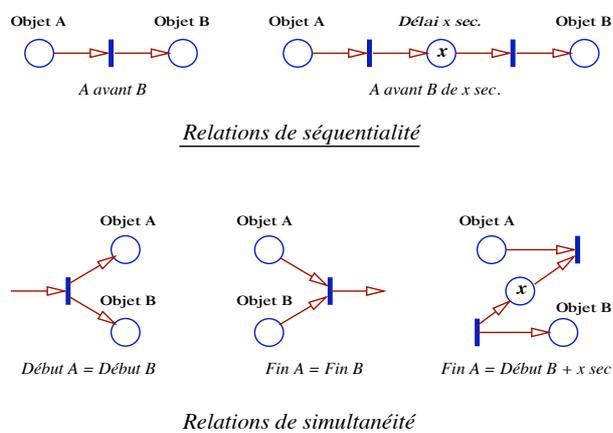


Figure 11. Les réseaux de Petri temporisés.

5.2.2. Modèle de Wahl-Rothermel

Le modèle de Wahl-Rothermel [WAH 93] est une adaptation du modèle général d'Allen [ALL 83]. Il propose un jeu d'opérateurs de composition intégrant la notion de délai dans les relations de simultanéité et de séquentialité (cf. Figure. 12).

L'élaboration de ces opérateurs a été motivée par le souci de faciliter la spécification de la structure temporelle du document par un auteur en réduisant le jeu d'opérateurs qu'il manipule. En effet, un même opérateur peut remplacer plusieurs de ceux proposés par Allen ou plusieurs relations entre instants. Par exemple l'opérateur *overlaps*, présenté dans la Figure. 11, permet de décrire respectivement les opérateurs *starts*, *finishes* et *overlaps* d'Allen pour ($\delta_1 = 0$), ($\delta_2 = 0$) et ($\delta_1 > 0$ et $\delta_2 > 0$ et $\delta_3 > 0$). Il est à noter que ce modèle n'est proposé que pour les intervalles déterministes, ce qui, sous cette hypothèse, entraîne que l'auteur ne doit spécifier, pour

certains opérateurs d'arité 2 ou 3, qu'un sous ensemble des paramètres, les autres étant déductibles par un calcul simple.

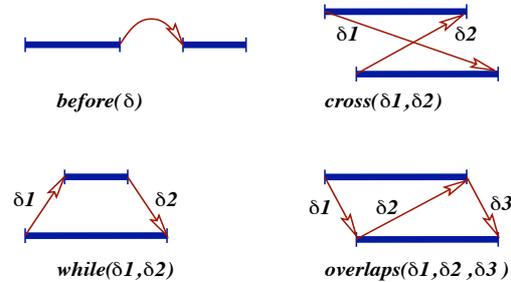


Figure 12. Quelques opérateurs du modèle Wahl-Rothermel.

6. Conclusions et perspectives

Nous avons exposé dans cet article, la nature des problèmes liés à l'intégration de données sensibles au temps dans les documents multimédia structurés. Ces problèmes ont été posés lors de la réalisation du prototype MADEUS. Parmi ces problèmes, la mise en œuvre de modèles temporels abstraits occupe une place centrale. Des efforts de modélisation actuellement entrepris pour proposer de tels modèles, ressortent plusieurs points dont la prise en compte conditionne la construction d'éditeurs multimédia :

- Le modèle temporel abstrait d'un document s'insère dans une structure globale considérant des dimensions de diverses sortes (temporelle, spatiale, logique et hypermédia). La description temporelle d'un document est à cheval sur plusieurs dimensions et ne doit pas être considérée séparément. En particulier, les structures temporelles peuvent tirer profit de la structuration logique d'un document, et les structures spatiales et temporelles d'un document sont fortement corrélées.
- Le compromis nécessaire entre l'utilisation de modèles à base de temps qualitatif, qui fournissent une description de plus haut niveau que ceux à base de temps quantitatif, et la prise en compte des incohérences qu'ils introduisent mettent en avant l'importance de la recherche de méthodes de vérification de la cohérence. Celles-ci, puisqu'elles influent sur le processus d'édition d'un document, se doivent d'être efficaces, et la mise en œuvre de modèles temporels abstraits doit en tenir compte.
- Fournir un formatage temporel permettant une souplesse dans le processus d'édition s'impose comme une condition nécessaire à la conception d'éditeurs faciles d'utilisation.

La complexité de conception d'un système d'édition multimédia apparaît clairement dans la diversité des modèles actuellement proposés, chacun essayant d'en réaliser une partie. Certains problèmes liés à la vérification de la cohérence et au formatage temporel ne sont pas encore résolus de manière adaptée. En particulier,

les méthodes de vérification de la cohérence et de formatage temporel s'adaptent peu à une édition interactive. Cela est dû à des lacunes qui se situent à deux niveaux :

- Les algorithmes de vérification de la cohérence existants nécessitent, pour chaque introduction d'une nouvelle spécification, une exécution portant sur la totalité du document. Ceci rend la vitesse de la phase de vérification tributaire de la complexité du graphe de contraintes représentant la structure temporelle du document. Elle devient de ce fait, de plus en plus lente, au fur et à mesure que le contenu du document croît.

- Lorsqu'un auteur spécifie une contrainte qui s'avère incohérente, il serait souhaitable que l'éditeur lui indique la source de cette incohérence et lui propose des solutions alternatives. Celles-ci consisteraient, par exemple, à modifier certaines des contraintes existantes pour que la nouvelle devienne possible.

Par ailleurs, le souci de concevoir des éditeurs hypermédia interactifs met en évidence d'autres problèmes comme la conception d'interfaces adaptées qui permettent la manipulation du document suivant ses différentes dimensions structurales. En effet, la nature multidimensionnelle d'un document multimédia rend l'opération d'édition particulièrement complexe. Il est à noter toutefois, que l'expérience acquise jusque là en matière d'éditeurs structurés multimédia est relativement faible. Les quelques systèmes existants [HAR 93] [HAM 93] sont encore au stade de prototype et ne couvrent pas encore tous les problèmes qui sont posés.

7. Bibliographie

- [ALL 83] ALLEN J.F., "Maintaining Knowledge about Temporal Intervals", *Communications of the ACM*, Vol. 26(N° 11), p. 832-843, Novembre 1983.
- [BER 92] BERNERS-LEE T., CAILLIAU R., GROFF J., POLLERMANN B., "World-Wide Web : The Information Universe", *Electronic Networking: Research, Applications and Policy*, Vol. 1(N°), Printemps 1992.
- [BLA 94] BLAIR G.S., PAPHOMAS M. et al. "Supporting Real-time Multimedia Behaviour in Open Distributed Systems: An Approach Based on Synchronous Languages", *Proceedings of the Second ACM International Conference on Multimedia*, p. 299-306. San Francisco, CA, 15-20 Octobre 1994.
- [BOU 94] BOURSIER P., TAUFOUR P.A., *La technologie Multimédia*, Hermès, Paris, Mars 1994.
- [BUC 93] BUCHANAN M.C., ZELLWEGER P.T., "Automatic Temporal Layout Mechanisms", *Proceedings of the First ACM International Conference on Multimedia*, p. 341-350, Anaheim, Californie, Août 1993.
- [BUF 94] BUFORD J.F.K., *Multimedia Systems*, Addison-Wesley, New York, 1994.
- [CHE 95] CHEN Z., TAN S., CAMPBELL R. et LI Y., "Real Time Video and Audio in the World Wide Web", *Proceedings of the Fourth International Conference on the World Wide Web*, Boston, MA, USA, décembre 1995.
- [DIR 95] DIRECTOR 4.0, "User's Guide", Macromedia. 1995.

- [DRA 93] DRAPEAU G.D., "Synchronization in the MAestro Multimedia Authoring Environment", *Proceedings of the First ACM Conference on Multimedia*, p. 331-340, Anaheim, California, Août 1993.
- [DUB 93] DUBOIS D., FARGIER H., PRADE H., "The Use of Fuzzy Constraints in Job-Shop Scheduling", *In Proceedings of IJCAI 93 Workshop on Knowledge-Based Planning, Scheduling and Control*, Chambéry (FRANCE), 1993.
- [FUR 94] FURTH B., "Multimedia Systems: An Overview", *IEEE Multimedia*, Vol. 1(No. 1), p. 47-59, Printemps 1994.
- [GHA 95] GHALLAB M., VIDAL T., "Focusing on a Sub-Graph for Managing Efficiently Numerical Temporal Constraints", *Proceedings of FLAIRS*, Melbourne Beach (FL), 1995.
- [GOL 90] GOLDFARB C. F., *The SGML Handbook*, Oxford University Press, Oxford, 1990.
- [HAM 93] HAMAKAWA R., REIKIMOTO J., "Object Composition and Playback Models for Handling Multimedia Data", *Proceedings of the First ACM International Conference on Multimedia*, p. 273-281, Anaheim, California, Août 1993.
- [HAR 93] HARDMAN L., VAN ROSSUM G., BULTERMAN D.C.A., "Structured Multimedia Authoring", *Proceedings of the First ACM International Conference on Multimedia*, p. 283-289, Anaheim, California, Août 1993.
- [HYT 92] HyTime Information Technology, "Hypermedia/Time-based Structuring Language (HyTime)", *ISO/IEC 10743*, Novembre 1992.
- [IIN 94] IINO M., DAY Y.F., GHAFOR A., "An Object-Oriented Model for Spatio-Temporal Synchronization of Multimedia Information", *Proceedings of the International Conference on Multimedia Computing and Systems*, p. 110-119, Boston, Massachusetts, 14-19 Mai 1994.
- [ISO 89] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, *International Standard ISO/IEC 8613: Information Processing - Text and Office Systems - Office Document Architecture (ODA) and Interchange Format*, International Organization for Standardization, Geneva/New York, 1989.
- [ISO 92] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, *International Standard ISO/IEC 8613: HyperODA - extensions for temporal relationships, extensions for non-linear structures*, International Organization for Standardization, 1992.
- [KIM 95] KIM M. et SONG J., "Multimedia Documents with Elastic Time", *Proceedings of the third ACM International Multimedia Conference and Exhibition*, p. 143-154, San Francisco, CA, USA, Novembre 1995.
- [LAM 78] LAMPORT L., "Time, Clocks and the Ordering of Events in a Distributed System", *Communications of the ACM*, Vol. 21(N° 7), p. 558-565, Juillet 1978.
- [LAM 86] LAMPORT L., *LATEX: a document preparation system*, Addison-Wesley, New York, 1986.
- [LAY 95] LAYAÏDA N., VIONDURY J.Y., "Multimedia Authoring: a 3D Interactive Visualization Interface based on a Structured Document Model", *Sixth International Conference on Human-Computer Interaction*, Yokohama, Japan, Juillet 1995.
- [LAY 96] LAYAÏDA N., SABRY-ISMAIL L., "Maintaining Temporal Consistency of Multimedia Documents Using Constraint Networks", *Proceedings of the International Conference on Multimedia Computing and Networking 1996*, p. 124-135, SPIE 2667, San José, CA, USA, Février 1996.

- [LEG 91] LEGALL D., "MPEG : A Video Compression Standard for Multimedia Applications", *Communications of the ACM*, Vol. 34(No. 4), p. 45-68, Avril 1991.
- [LIE 94] LIESTØL G., "Aesthetic and Rhetorical Aspects of Linking Video in Hypermedia", *Proceedings of the ACM European Conference on Hypermedia Technology*, p. 217-222, Edinburgh, Scotland, Septembre 1994.
- [LIT 93] LITTLE T.D.C., GHAFOR A., "Synchronization and Storage Models for Multimedia Objects", *IEEE Journal on Selected Areas in Communications*, Vol. 8(N° 3), p. 413-426, Avril 1990.
- [LIT 93] LITTLE T.D.C., GHAFOR A., "Interval-Based Conceptual Models for Time-Dependent Multimedia Data", *IEEE Transactions on Knowledge and Data Engineering (Special Issue : Multimedia Information Systems)*, Vol. 5(N° 4), p. 551-563, Août 1993.
- [NAN 93] NANARD J., NANARD M., "Should Anchors be typed Too? An Experiment with MacWeb", *Fifth ACM Conference on Hypertext Proceedings*, pp. 51-62, ACM Press, Seattle, Washington, Novembre 1993.
- [OGA 92] OGAWA R., TANAKA E., TAGUCHI D., HARADA K., "Design Strategies for Scenario-based Hypermedia: Description of its Structure, Dynamics, and Style", *Proceedings of the ACM Conference on Hypertext*, p. 71-80, Milano, Italy, Novembre 92.
- [PRI 93] PRICE R., "MHEG: An Introduction to the Future International Standard for Hypermedia Object Interchange", *Proceedings of the First ACM Conference on Multimedia*, p. 121-128, Anaheim, California, Août 1993.
- [QUI 86] QUINT V., VATTON I., "Grif: An Interactive System for Structured Document Manipulation", *Text Processing and Document Manipulation, Proceedings of the International Conference*, p. 200-213, Cambridge, 1986.
- [ROI 94] ROISIN C., VATTON I., "Merging Logical and Physical Structures in Documents", *Proceedings of the Fifth International Conference on Electronic Publishing, Document Manipulation and Typography*, p. 327-337, Wiley Publishers, Darmstadt, Allemagne, Avril 1994.
- [SEN 96] SENAC P., DIAZ M., LEGER A., de SAQUI-SANNES P., "Modeling Logical and Temporal Synchronization in Hypermedia Systems", *IEEE Journal On Selected Areas In Communications*, Vol 14(N° 1), p. 1-29, Janvier 1996.
- [SUN 95] SUN MICROSYSTEMS, *The Java Language Specification*, Sun Microsystems, 1995.
- [VAL 94] VALDES R., "What's UP at Kaleida ?", *Dr Dobb's Developer Update*, Vol. 1(N° 7), Septembre 1994.
- [WAH 94] WAHL T., ROTHERMEL K., "Representing Time in Multimedia Systems", *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, p. 538-543, Boston, Massachusetts, 14-19 Mai 1994.
- [WAL 91] WALLACE G., "The JPEG Still Picture Compression Standard", *Communications of the ACM*, Vol. 34(N° 4), p. 30-44, avril 1991.
- [VIL 86] VILAIN M., KAUTZ H. A., "Constraint Propagation Algorithms for Temporal Reasoning", *Proceedings of AAAI*, pp. 377-382, Philadelphia, Août 1986.