

Génération des contraintes à partir de structures rhétoriques

Titre de l'Article: Generating Presentation Constraints from Rhetorical structure

Auteur: Lloyd Rutledge, Brian Bailey, Jacco van Ossenbruggen, Lynda Hardman et Joost Geurts

De CWI, Amsterdam ET Department of Computer Science and Engineering, University of Minnesota, Minneapolis, USA.

Nom du relecteur: Takoua ABDELLATIF

Type d'article: Il s'agit d'un article de recherche, avec développement d'un exemple illustrant l'idée des auteurs.

Année: 2000.

Notes et appréciations:

Appréciation globale (note 4) L'article fait partie des premiers travaux qui transforment les structures rhétoriques en contraintes pour les traduire en présentations finales. On va voir dans la partie « critique » que ce travail a été complété par une thèse de l'un des auteurs Joost Geurts, ainsi que dans d'autres recherches.

Clarté et qualité d'écriture (3) Les concepts de l'article sont bien clarifiés par les exemples donnés. Par contre, on note beaucoup de redondance dans les explications (ex : on rencontre la définition des rhétoriques et des contraintes à plusieurs endroits de l'article) et le rapport entre la partie théorique et l'implémentation n'est pas très claire (il m'a fallu comprendre dans d'autres articles l'architecture du système Berlage avant de comprendre la mise à jour faite par les auteurs).

Application du cours (note 5) Le cours du module MW a expliqué les bases de technologies utilisées dans l'article : XML et SMIL. Il nous a aussi sensibilisé aux difficultés pour présenter un document multimédia, par exemple, les problèmes complexes de synchronisation temporelle de SMIL. Cet article permet d'approfondir la compréhension de ces problèmes, de voir une manière de les aborder et d'avoir des exemples d'implémentations dans des systèmes réels.

Type d'article: Il s'agit d'un article de recherche. On a implémenté un exemple pour illustrer les concepts et on a mis à jour un système (Berlage) pour implémenter ces concepts.

Résumé:

Les structures rhétoriques représentent l'enchaînement conceptuel d'un document et un moyen de décrire sa présentation à un haut niveau d'abstraction. Beaucoup de recherches se sont focalisées sur la génération de présentations à partir de contraintes spatiales, temporelles et/ou de navigation ou alors, directement à partir de rhétoriques. Cet article, propose, plutôt, une méthode pour traduire les structures rhétoriques de l'utilisateur en des contraintes qui seront par la suite transformées en présentation finale.

Générer automatiquement un document multimédia à partir de rhétoriques et contraintes permet d'adapter la présentation à de différents types d'utilisateurs. Ceci permet aussi d'éviter la redondance des données et aider les utilisateurs et les logiciels à trouver plus rapidement l'information à partir de grands documents.

Pour générer les contraintes à partir des structures rhétoriques, les auteurs ont créé deux concepts : le *débordement* (overflow) et la *compensation*. Pour les illustrer, on a développé un guide électronique de programmes de films à la télévision ou EPG (Electronic Program Guide). L'utilisateur envoie ses requêtes à une base de données, spécifiant ce qu'il désire voir (par exemple ses types de films favoris) et comment il veut que ses films soient présentés (par exemple, on spécifie la durée maximale de la présentation). Si on se limite à la structure spatiale, et on veut afficher toute la séquence de films avec leurs titres respectifs à la fois, on risque de dépasser les dimensions de l'écran. On tombe alors dans ce qu'on appelle *débordement*. On est obligé alors d'adopter une autre structure, temporelle cette fois-ci, consistant à placer les films dans des écrans différents et les défilement les uns après les autres. Cette opération s'appelle *compensation*. Si le temps total du défilement est supérieur au temps maximal spécifié par l'utilisateur, on a recours à la navigation en ajoutant par exemple, des barres de menus sur l'écran permettant de choisir le programme qu'on veut au lieu de voir passivement tous les films à la fois. Encore une fois, on a rectifié le *débordement* temporel par le mécanisme de *compensation*. Le *débordement* est donc la violation de contraintes en se limitant à une seule dimension de présentation (spatiale, temporelles ou liens) et la *compensation* est un moyen de corriger ce *débordement* en passant à d'autres formes de présentation afin de respecter les contraintes.

L'exemple ci-dessus nous mène à deux types de contraintes : intra-dimensionnelle et inter-dimensionnelle. Les contraintes intra-dimensionnelles traduisent des relations entre objets multimédia de même structure de présentation. Elle sont résolues indépendamment des autres formes de présentation. Lorsqu'une contrainte intra-dimensionnelle n'est pas satisfaite, on a un effet de *débordement*. Ceci déclenche, comme on a vu une autre série de contraintes intra-dimensionnelles. Le passage d'une structure à une autre constitue les contraintes inter-dimensionnelles qui définissent la stratégie de *compensation*.

La stratégie de *compensation* de l'exemple EPG est en fait linéaire. On passe en effet, en cas de *débordement*, d'une structure de présentation vers une autre et on parcourt à chaque fois, dans l'ordre et une seule fois, les objets multimédia. La *compensation* peut être beaucoup plus complexe, en adoptant par exemple des liens de forme arborescente ou/et en changeant l'ordre des éléments. C'est cette stratégie de *compensation* qui va nous permettre de passer du niveau d'abstraction des rhétoriques vers les contraintes sans avoir forcément de correspondance directe. A partir d'une structure rhétorique, on peut générer plusieurs types de contraintes spatiales, temporelles et/ou de navigation donc des présentations finales différentes.

Considérons la structure rhétorique la plus répondue: les *séquences*. On doit présenter une liste d'éléments dans un ordre précis. La stratégie de compensation ne doit donc pas changer l'ordre des éléments. On peut par exemple, afficher une collection de constructions de Amsterdam selon leur histoire, ou les adresses des rues. Il y a deux types de séquences : strictes et permissives. Dans le cas de *séquences* strictes, on doit maintenir l'ordre des éléments. On peut par exemple, les présenter avec des pages contenant uniquement le bouton « suivant » pour accéder dans l'ordre aux pages suivantes. Dans le cas de séquences *permissives*, on peut présenter les éléments dans des pages avec des barres de menus ou on peut naviguer dans n'importe quel ordre.

On distingue deux types de relations rhétoriques : les relations multi-nucléaires et les relations noyau-satellite. La structure rhétorique multi-nucléaire concerne une liste de

composants équivalents, par exemple, les *séquences*. Inversement, la structure noyau-satellite concerne des éléments de taille ou d'importance différente, par exemple, le *Résumé* d'un texte. Dans le cas de *débordement* spatial, on adopte alors une stratégie de compensation de navigation d'ordre. Par exemple, quand l'écran ne peut contenir la totalité des composants, on présente d'abord le noyau et on met des liens vers les satellites.

Un consortium de chercheurs ont développé un modèle pour la génération automatique de présentations finales à partir de spécifications haut niveau : le SRM-IMMPS (Standard Reference Model for Intelligent Multimedia Presentation Systems). On a utilisé et mis à jour Berlage, un système implémentant les composants de SRM-IMMPS, pour illustrer les concepts de *débordement* et *compensation*. L'article décrit brièvement les composants qui permettent de transformer les formes rhétoriques en contraintes.

Le composant «expert d'application» dans le SRM-IMMPS, fournit les informations sur le domaine en question ou des formes de media particuliers. On définit en fait une ontologie de classes du domaine et on attribue à chaque classe, une série de classes d'attributs (par exemple, longueur, largeur, relation rhétorique ..). La « couche de contrôle » prend en entrée le but de l'utilisateur. Dans l'exemple de l'EPG, le but était la série de films à regarder. Ce but est découpé en sous-buts et le résultat est envoyé à la couche « content layer ». Cette couche permet de choisir les contenus media appropriés qui correspondent aux besoins spécifiés dans les buts. Cette opération n'est pas triviale et repose sur une bonne connaissance du domaine. Une autre couche responsable de la conception de la présentation finale, est la couche « design expert ». Ici, design est équivalent à « stylesheet » en XML. Dans Berlage, spécifier la transformation des structures rhétoriques en contraintes est une question de design. On a rajouté à la couche « design expert », un composant « rhetoric design expert » pour distinguer le design rhétorique des autres formes de design. L'opération de translation des structures rhétoriques en contraintes se fait par contre dans une autre couche, la couche « design layer ». La couche « content layer » fournit comment la présentation finale va se dérouler, c'est à dire la structure rhétorique, et la couche « design expert » fournit le design. A partir de ces données, la couche « design layer » génère une liste de contraintes dans un plan de design sous le format XML, qu'on passe par la suite à la « couche de réalisation ». Cette couche « parse » le document en distinguant les relations temporelles, spatiales et de navigation. Les contraintes spatiales et temporelles sont résolues avec « Cassoway solver ». On génère à partir des contraintes résolues, un document SMIL qu'on peut jouer utilisant « GriNS SMIL player ».

Commentaires/critiques:

Cet article nous a montré qu'il y a énormément d'écart entre les structures abstraites d'un document multimédia et sa présentation finale. Cet écart est tellement grand, qu'on ne peut le franchir en une seule étape mais en plusieurs, ici en passant par les contraintes. On retrouve cette même idée, deux ans après, en Janvier 2002, dans le document de thèse de l'un des auteurs : Joost Geurts. Le document intitulé « Constraints for Multimedia Presentation Generation » présente en fait, un découpage plus fin du problème, en plusieurs étapes permettant de générer un document multimédia à partir de structures sémantiques (dont les structures rhétoriques). Ces étapes sont respectivement : la couche sémantique, « device », contraintes qualitatives, contraintes quantitatives et la présentation finale. Cuypers, un système prototype, implémente ces différentes couches.

Dans les deux documents, on a utilisé les contraintes comme moyen intermédiaire entre les spécifications sémantiques et la présentation finale, même si on ne les classe pas de la même manière. Dans notre document, on a classé les contraintes en contraintes intra-dimensionnelles et contraintes inter-dimensionnelles. Dans la thèse de Joost Geurts, on les classe plutôt en contraintes quantitatives, pour vérifier que la présentation finale satisfasse toutes les contraintes numériques de l'environnement, et contraintes qualitatives, pour faciliter le raisonnement de plus haut niveau.

On note aussi une différence dans la complexité des exemples illustrés dans les deux documents. Dans notre document, on a développé l'exemple des EPG, dans lequel une forme rhétorique très simple est présentée : les *séquences* alors que dans la thèse de Joost Geurts, on a un exemple plus complexe qui traite à la fois plusieurs structures rhétoriques : *exemple*, *élaboration* et *séquence*. Le but est en effet, de construire une présentation autour d'«exemples de Chiaroscuro dans les travaux de Rubrant Van Rijn». On a utilisé les images comme *exemple* du concept, le texte pour l'*élaboration* du concept et pour préserver l'ordre des images dans le temps, les images sont présentées avec une relation de *séquence*. Ces données sont présentées avec un arbre RST (Rhetorical Structure Theory). Dans les deux documents, l'utilisation de l'exemple était très utile au lecteur pour comprendre les concepts théoriques présentés qui sont, en général, difficiles à expliquer s'ils ne sont pas accompagnés de preuves concrètes. Utiliser les exemples, a permis aussi d'expliquer le fonctionnement des systèmes de génération automatique des présentations (Berlage ou Cuypers) et montrer leur utilisation dans des domaines différents.

Les deux systèmes dans les deux articles, reposent sur le même principe et les mêmes technologies. Les choix sémantiques ou rhétoriques sont encodés en XML, ils passent par des couches qui traduisent ces choix en contraintes qualitatives puis discrètes qui vont être résolus par un salvateur de contraintes. La présentation finale est encodé par HTML ou SMIL (ou SVG dans le cas de Cuypers).

La thèse de Joost Geurts confirme aussi, par une étude plus approfondie sur la théorie des contraintes arithmétiques que les contraintes multimédia ne peuvent être traduites d'une manière simple et unique en contraintes discrètes. Cette idée est aussi exprimée dans notre article grâce à l'exemple des EPG. On a vu qu'on peut traduire une structure abstraite et provoquer un *débordement* et qu'on adopte alors la stratégie de *compensation*. Dans la thèse de Geurts, cette possibilité de faire marche-arrière réside dans le découpage de son système en couches ou on peut passer d'une couche à une autre afin de revoir ses choix numériques en cas d'échec. Quelque soit la manière de le présenter, le principe est le même dans les deux documents: trouver un compromis entre le choix sémantique et les ressources.

Dans les deux articles, la méthode adoptée est de décomposer le but voulu de l'utilisateur en d'autres sous-buts et on transforme ainsi le problème en d'autres problèmes de granularité plus fine, que le « constraint solvers » sait résoudre. La distribution des décisions peut par contre, rendre la prise en compte des critères globaux, une opération difficile à réaliser. On a alors moins de liberté pour concevoir la structure globale du document. Il existe d'autres approches qui traitent le sujet dans le sens inverse. Dans l'article «Automatic Construction of Personalized TV News Programs», écrit par Merialdo B., Lee K.T, LuparelloD., Roudaine J., on donne plutôt la priorité au contenu et on cherche après à organiser les éléments. Cette méthode est connue sous le nom de « processus d'optimisation ».

Ces deux approches, celle de l'article et le « processus d'optimisation », sont en fait réutilisées dans l'article « A Generic architecture for automated construction of multimedia presentations », écrit par Frederic Bes, Muriel Jourdan et Farid Khantache de l'Unité de recherche INRIA Rhône Alpes. Malgré les grands avantages que présente ces deux méthodes, on trouve que chaque méthode présente des inconvénients et qu'il est plus efficace de les combiner. On considère en effet, méthode du « processus d'optimisation » est très liée au domaine d'application et qu'elle ne propose pas, par conséquent, une solution plus générale pour générer la présentation de domaines différents. Et on reproche à la méthode de notre article, la difficulté d'exprimer les critères globaux, puisqu'on découpe le problème en de très petites entités. L'article de l'Unité de l'INRIA propose alors une architecture qui combine les avantages de chaque approche pour proposer une solution générique valable pour plusieurs domaines. Le principe consiste à choisir les objets multimédia ainsi que quelques décisions à propos de leur organisation. On utilise les « transformation sheet » dans la construction de la structure globale du document et on ne précise pas dans cette étape les valeurs de formatage (telles que la taille et la durée des objets). C'est le but de l'étape d'optimisation de préciser plus en détail ces valeurs. En choisissant le contenu et l'organisation des objets, on bénéficierait d'une bonne présentation (qui respecte les contraintes) tout en respectant les critères globaux.

Synthèse :

L'article « Generating Presentation Constraints from Rhetorical structure » définit en détail les structures rhétoriques et les différentes classifications des contraintes multimédias. Il propose de traduire les structures rhétoriques en contraintes qu'on transforme par la suite en présentation finale. Cette solution est illustrée d'une manière claire par l'exemple de l'EPG. Ce travail a été complété dans la thèse de l'un des auteurs, Joost Geurts, qui a implémenté un système prototype traitant des cas plus complexes. Malgré les avantages que présente cette solution, on risque en se focalisant sur les critères fins, après découpage du problème initial, de perdre les critères globaux. Il existe alors d'autres solutions, comme celle de l'INRIA, qui reprend la méthode des contraintes de notre article et la combine avec le processus d'optimisation, décrit ci-dessus pour respecter la structure globale du document.