Indices visuels et métriques combinatoires pour la visualisation de donnee hierarchiques - Article long

Guy Melancon, Ivan Herman

CWI – Centrum voor Wiskunde en Informatica
P.O. Box 94079
1090 GB AMsterdam, Nederland
{Guy.Melancon,Ivan.Herman}@cwi.nl
téléphone: (31) 20 592 41 80
télécopie : (31) 20 592 41 99

Maylis Delest

LaBRI, Université Bordeaux I
351, Cours de la Libération
33405 Talence Cedex, France
maysli@labri.u-bordeaux.fr
téléphone: (33) 5 56 84 60 82
télécopie : (33) 5 56 84 66 69

RESUME
La visualisation de grand corpus de données nécessite de définir des outils de navigation pour l'utilisateur. Nous présentons une technique qui automatise en partie la définition d'indices visuels pour des structures hiérarchiques, en particulier les arbres. Notre technique repose sur le calcul d'une métrique pour chaque sommet et/ou arc de la structure. Nous donnons des exemples de métriques qui sont particulièrement bien adaptées aux structures hiérarchiques et qui permettent de rendre compte de leur complexité. Ces valeurs numériques peuvent alors être utilisées pour définir des attributs graphiques des éléments d'une représentation. Nos résultats ne sont pas tributaires d'un représentation particulière et peuvent être *a priori* être appliquée à n'importe quelle représentation des arbres ou de graphes plus généraux.

MOTS CLES Visualisation, navigation, indices visuels, graphes, métriques, attributs graphiques.

NOM ET VERSION DU LOGICIEL UTILISE POUR ECRIRE LA SOUMISSION :
ex :Word 2.0c Microsoft pour Windows

Microsoft Word 97

FORME DE PARTICIPATION :

x	Cocher la forme de participation
	Article court
x	Article long
	Ateliers
	Expériences industrielles
	Forum
	Rencontres doctorales
	Table ronde
	Tutoriaux

THEMES DE LA SOUMISSION :

x	Cocher les thèmes choisis (vous pouvez en cocher plusieurs)
	- Fondements théoriques des IHM
	- Modèles de l'utilisateur, modèles cognitifs
	- Analyses et modèles de tâches
	- Formalismes et IHM
	- Modèles de développement des systèmes interactifs
	- Modèles distribués de coopération homme-machine
	- Modélisation des organisations humaines
	- Dispositifs, techniques d'interaction

	- Morphologie de l'interaction
	- Nouvelles techniques d'interaction et de présentation
	- Nouvelles interfaces pour la supervision
	- Interfaces haptiques, interaction au stylo
	- Interaction multimodale, gestuelle ou en 3 dimensions
	- Intégration de la vision dans l'interface
	- Interaction vocale
	- Interfaces écologiques
	- Coopération entre modalités/média
	- Analyse, conception et évaluation
	- Architecture des systèmes interactifs
	- Processus de conception et ergonomie
	- Méthodes de conception et d'évaluation
	- Evaluation ergonomique et validation
	- Méthodes formelles dans la conception ou l'évaluation
	- Assistance à l'utilisateur
	- Ergonomie du web
	- Systèmes, outils
	- Outils de prototypage
	- Outils de construction
	- Environnements de développement
	- Programmation visuelle
	- Programmation par démonstration
	- Communication médiatisée par l'ordinateur
	- Interaction asynchrone : téléenseignement, télémédecine, etc.
	- Interaction synchrone
	- Conception coopérative
	- Édition coopérative
	- Interaction et coopération
x	- Visualisation et représentation de l'information
x	- Visualisation de grands corpus de données
	- Recherche dans des documents complexes
x	- Aide interactive
	- Stratégies multi-écrans
	- Interaction dans les mondes 3D
x	- Interaction et informatique graphique
	- "Visualisation" pour les handicapés
	- Nouveaux champs d'application
	- Réalité virtuelle
	- Réalité augmentée
	- Collecticiels
	- Interfaces intelligentes
	- Interfaces pour les systèmes critiques
	- Informatique mobile et informatique disséminée
	- Multimédia, hypermédia, et interaction sur le WEB
	- IHM et publics adaptés
	- Interaction pour les personnes âgées et handicapées
	- Interfaces adaptées aux besoins du grand public
	- IHM et Société
	- Applications et expériences
	- Commerce électronique
	- Jeux et loisirs
	- Santé
	- Banque
	- Transport
	- Bâtiment
	- Enseignement de l'IHM

Indices visuels et métriques combinatoires pour la visualisation de données hiérarchiques

Guy Melancon, Ivan Herman

CWI – Centrum voor Wiskunde en Informatica
P.O. Box 94079
1090 GB AMsterdam, Nederland
{Guy.Melancon,Ivan.Herman}@cwi.nl
téléphone: (31) 20 592 41 80
télécopie : (31) 20 592 41 99

Maylis Delest

LaBRI, Université Bordeaux I
351, Cours de la Libération
33405 Talence Cedex, France
maysli@labri.u-bordeaux.fr
téléphone: (33) 5 56 84 60 82
télécopie : (33) 5 56 84 66 69

RESUME

La visualisation de grand corpus de données nécessite de définir des outils de navigation pour l'utilisateur. Nous présentons une technique qui automatise en partie la définition d'indices visuels pour des structures hiérarchiques, en particulier les arbres. Notre technique repose sur le calcul de valeurs métriques pour chaque sommet et/ou arc de la structure. Nous donnons des exemples de métriques particulièrement bien adaptées aux structures hiérarchiques qui permettent de rendre compte de leur complexité. Ces valeurs numériques peuvent alors être utilisées pour définir des attributs graphiques des éléments d'une représentation. Nos résultats ne sont pas tributaires d'un représentation particulière et peuvent *a priori* être appliqués à n'importe quelle représentation des arbres ou de graphes plus généraux.

MOTS CLES Visualisation, navigation, indices visuels, graphes, métriques, attributs graphiques.

INTRODUCTION

Les outils informatiques actuels tendent à systématiser l'exploration et l'analyse de données abstraites à l'aide d'outils de visualisation. Dans beaucoup de cas les données viennent équipées de relations et on peut alors se ramener au problème de la visualisation et de l'interaction avec les graphes. Le problème de visualisation de données intrinsèquement hiérarchiques pose à lui seul un défi de taille, même si la structure sous-jacente est tout simplement un arbre. Il existe de nombreux systèmes de visualisation de graphes qui, en particulier, implémentent des algorithmes de positionnement des sommets d'un arbre (cf, par exemple, [1] et les éditions précédentes). L'algorithme de Reingold et Tilford [2,3] calcule une représentation « classique » d'un arbre et offre en outre des qualités au niveau des performances et des aspects esthétiques du dessin qu'il produit. D'autres algorithmes ont été proposés récemment pour surmonter les problèmes causés par les grands corpus de données. L'utilisation de la troisième dimension, avec par exemple les *Cones Trees* [4,5] permet d'échapper au problème d'espace en disposant les sommets d'ancêtre commun sur la frontière d'un disque à la base d'un cône. L'utilisation de la 3D introduit toutefois des difficultés de perception qui lui sont propres (cf, par exemple, [6]). Le système Cheops [7] propose une disposition pyramidale dans le plan où les sommets de l'arbre sont représentés par des triangles qui se chevauchent et permet ainsi de stocker dans un espace relativement réduit une quantité exponentielle de sommets. L'utilisation de la géometrie hyperbolique [8,9], en 2D ou 3D, permet aussi de vaincre les problèmes d'espace caractéristiques aux arbres ou plus généralement aux graphes.

Cependant, le problème d'espace n'est pas le seul aspect sur lequel se sont penchés les travaux récents. Plusieurs auteurs ont proposé des modèles qui s'éloignent du modèle plus classique pour dessiner les graphes. Les *Tree Maps* [10] propose une représentation bisimensionnelle d'un arbre juxtaposant des zones rectangulaires par bissection répétées du plan. MAPA™ [11] propose une vue en perspective sur un ensemble de fiches qui se déploient dynamiquement. Les « Perspective Walls » [12] sont encore une autre alternative pour visualiser les arborescences. On trouve sur le web quantité d'applications qui proposent des alternatives, tablant sur l'utilisation d'une métaphore ou d'icônes pour représenter les sommets de la structure, ou offrant parfois une vue dynamique de l'arbre évoluant au gré des interactions avec l'utilisateur[1].

Comme le suggère Schneiderman [13], lors d'une exploration l'utilisateur développe d'abord une image de la globalité de la structure, pour ensuite appliquer une séquence de filtres ou transformations et arriver finalement aux détails portés par la structure. Tous s'accordent à dire qu'un algorithme de positionnement ne suffit pas pour offrir à un utilisateur une représentation de ses données. Les algorithmes de positionnement plus classique recherchent une solution qui optimise certaines esthétiques (même longueur pour toutes les arêtes, dispositions des noeuds par niveaux, respect d'un espace

[1] Voir la page `http://www.cwi.nl/InfoVisu` qui contient, en outre, une liste de liens sur d'autres sites.

minimal entre les noeuds, respect des symétries dans la structure, etc; cf, par exemple, [14]). Mais ces algorithmes ne permettent pas nécessairement de donner du relief à la structure. De même, une métaphore et l'utilisation d'icônes peuvent dans une certaine mesure aider l'utilisateur à naviguer, mais ne lui suffisent pas pour développer une image mentale de la structure dans sa globalité. De ce fait, il doit lui-même développer ses repères pour guider son exploration.

Les résultats que nous présentons ici proposent d'utiliser des outils combinatoires pour mesurer la structure et enrichir graphiquement sa présentation. Les indices visuels que nous proposons aident l'utilisateur, non seulement à développer une image globale d'un arbre, mais aussi à mieux se repérer par rapport à l'arbre tout entier au cours de l'application de filtres (zoom, distorsions et/ou déplacement latéral, etc). Notre méthodologie n'est pas tributaire d'une représentation visuelle particulière des arbres et peut, *a priori*, être appliquée à toute représentation.

La suite de l'article est organisé comme suit. Nous rappelons le calcul de valeurs métriques des sommets d'un arbre à travers un exemple. Nous montrons ensuite comment ces valeurs peuvent être utilisées pour définir des attributs graphiques des noeuds et arcs d'une représentation de l'arbre. Puis, nous décrivons brièvement comment ces valeurs peuvent permettre de filtrer les éléments de la structure pour n'en afficher qu'un sous-ensemble et alléger la représentation. La discussion se tourne ensuite vers les choix possibles de métriques et la dépendance de ces choix par rapport au contexte de l'application. Dans certains cas, les données hiérarchiques sont plus adéquatement représentées par des graphes acycliques orientés. C'est le cas, par exemple, des graphes décrivant les scènes de réalité virtuelle ou encore les structures d'héritage des bases d'objets. Les algorithmes de visualisation d'arbres ne s'étendent pas toujours à ces structures et les algorithmes existant (cf [14,15,16]) ne sont envisa-geables que pour des structures relativement petites. La dernière section indique comment une adaptation simple de l'algorithme de Reingold et Tilford [2,3] combinée à une généralisation appropriée des métriques permet d'étendre notre technique aux graphes acycliques orientés.

ANALYSE COMBINATOIRE DE L'ARBRE

Notre technique repose sur une analyse automatique de la structure. Une visite récursive de l'arbre permet d'affecter à chacun de ses éléments une valeur numérique, appellée *métrique*. La métrique la plus naturelle, pour un arbre, est certainement la distance d'un noeud à la racine de l'arbre ; la valeur attribuée a un sommet est souvent appellée sa *profondeur* dans l'arbre.

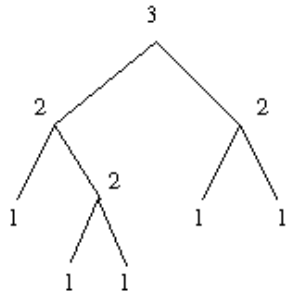
La combinatoire [17] regorge d'exemples de métriques, qui sont souvent utilisées pour caractériser certaines propriétés des arbres. Nous n'allons pas nous pencher sur cet aspect ici, mais plutôt chercher à comprendre leur intérêt dans un contexte de visualisation. De ce point de vue, comme nous allons le voir, des métriques de nature heuristique peuvent être définies.

Nous supposerons donc qu'un positionnement des éléments de l'arbre nous est donné. Le problème que nous abordons n'est pas celui de calculer une représentation, mais bien de définir des attributs génériques qui puissent enrichir visuellement toute représentation.

Nombre de Horton-Strahler des arbres

L'utilisation des arbres comme outils de modélisation ou d'études dépasse largement le domaine de l'informatique. C'est dans des articles sur l'étude morphologique de bassins hydrologiques qu'apparaît la première fois la définition d'un nombre quantifiant les noeuds d'un arbre pour en décrire la « forme » [18,19]. Depuis, de nombreux travaux en mathématiques combinatoires en ont determiné les propriétés. Ces nombres, appelés nombre de Horton-Strahler des arbres (binaires), ont aussi été utilisés pour produire des images réalistes d'arbres [20].

Pour un arbre binaire, la valeur d'un sommet est calculée comme suit: les feuilles reçoivent une valeur 1. Un sommet interne reçoit la valeur: s+1 si ses deux fils ont même valeur s, et $s = \max(s_1, s_2)$ si ses deux fils ont des valeurs distinctes s_1 et s_2 (cf l'exemple ci-contre).



On peut facilement généraliser cette métrique aux arbres d'arité variable. On calcule alors la valeur d'un sommet en fonction des valeurs de ses fils (cf [21] pour une définition possible).

METRIQUES ET ATTRIBUTS GRAPHIQUES

Nous allons utiliser la métrique de Horton-Strahler pour illustrer nos idées. Nous décrirons plus loin d'autres métriques pour lesquels le même procédé peut être repris. On peut définir, à partir des valeurs données par la métrique de Horton-Strahler, des attributs graphiques pour une représentation d'un arbre. L'arbre dessiné au haut de la figure 1 illustre la représentation d'un arbre obtenu de l'algorithme de Reingold et Tilford [2,3]. La métrique peut être utilisée pour définir une épaisseur et une couleur pour chacun des arcs. La variation des valeurs entre sommets adjacents permet aussi d'obtenir un effet de dégradé le long d'un arc. L'arbre dessiné au

bas de la figure 1 est obtenu de l'arbre plus haut en appliquant ce procédé[2]. D'autres artifices graphiques sont aussi possibles ; on peut aussi faire varier la couleur en fonction des valeurs de la métrique, par exemple.

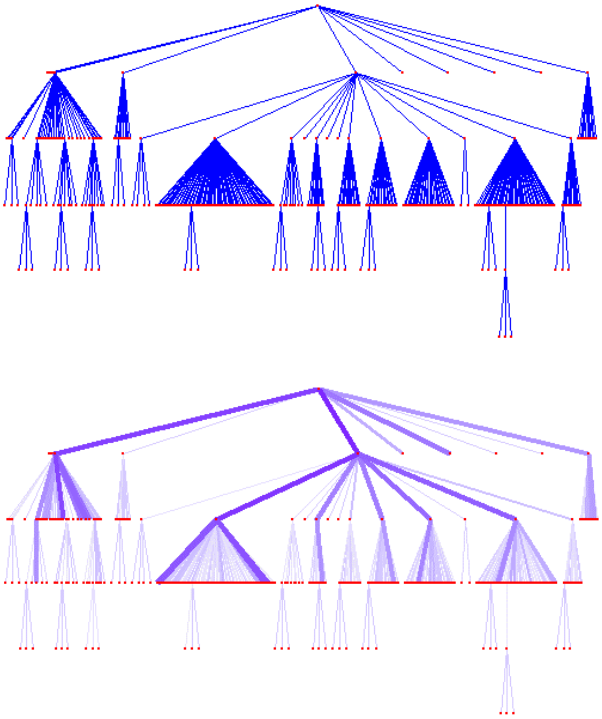


Figure 1. Haut : vue « classique » d'un arbre (répertoire UNIX).

Bas : vue en « relief » (épaisseur des arcs en fonction de la métrique Horton-Strahler)

NAVIGATION

L'étude d'exemples montre que le relief donné au dessin de l'arbre à l'aide d'une métrique permet à l'utilisateur de développer une image globale de la structure et d'en identifier les composantes principales. Par exemple, avec les nombres de Horton-Strahler, si l'arbre code les ramifications d'un bassin hydrologique, les arêtes de plus grande épaisseur correspondront aux cours d'eau majeurs du réseau.

Cet enrichissement du dessin permet aussi de guider la navigation. Une vue locale (conséquence d'effets de zoom, de recadrages et/ou de distorsions) donne, grâce au relief, des indications sur la direction à suivre pour retrouver la racine de la hiérarchie, ou encore les endroits voisins où se trouvent les ensembles d'informations les plus denses.

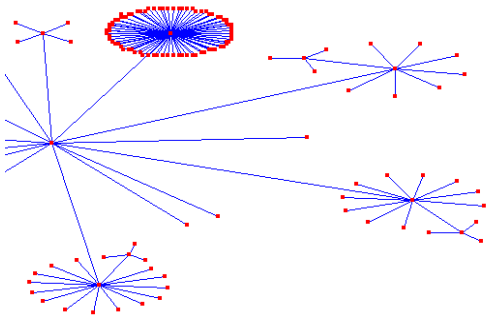


Figure 2. Vue « plate » d'une partie d'un arbre.

Par exemple, il est impossible de déterminer, dans la figure 3 si le sommet qui est l'ancêtre commun des sous-arbres qui se déploient sur la droite est effec-tivement la racine de l'arbre tout entier. Il pourrait très bien être la racine d'une sous-structure; et il est dans ce cas difficile de prédire dans quelle direction se trouve la racine et/ou les autres parties importantes de la structure (au sens de la métrique utilisée). L'utilisation des indices visuels, dans la figure 4, donne une réponse claire à ces questions. La racine est nécessairement vers la gauche ; mais l'épaisseur des traits montrent qu'il faut suivre celui qui monte vers le nord-ouest pour trouver la racine.

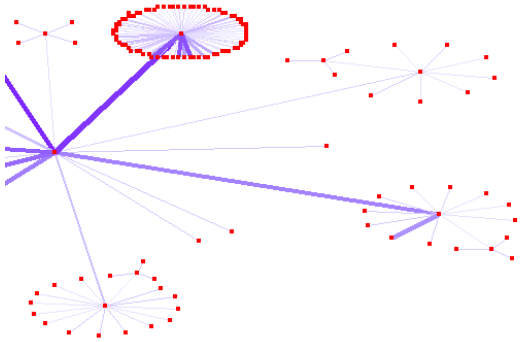


Figure 3. Vue « enrichie » d'un arbre.

ADAPTATION AU CONTEXTE ET AUTRES METRIQUES

La métrique définie par les nombres de Horton-Strahler n'est pas universelle. Ses qualités tiennent au fait qu'elle permet de quantifier l'importance d'un sommet dans le réseau sous-jacent à l'arbre étudié. Elle n'a de valeur que dans les cas où la mesure qu'elle fournit est pertinente par rapport au contexte. Ce qu'il faut en retenir tient plutôt dans le schéma de calcul qui est suivi pour attribuer une

[2] Les images proposées dans cet article ont été produites à l'aide du logiciel *Latour* développé par notre équipe et disponible sur notre site : `http://www.cwi.nl/InfoVisu`. La version électronique de cet article y est accessible et contient des images couleurs.

valeur aux sommets de l'arbre. Dans ce sens, la généralisation proposée dans [21] laisse une place à un poids qui peut être attaché à un sommet de l'arbre pour en relever l'importance et ainsi affecter le calcul. Par exemple, le poids attribué aux feuilles d'une arborescence UNIX pourrait être la taille du fichier correspondant. Le lecteur attentif aura remarqué que l'image au bas de la figure 1 utilise nécessairement la version « avec poids » des nombres de Horton-Strahler. En effet, certains arcs émergeant des feuilles sont déjà d'une épaisseur relativement grande.

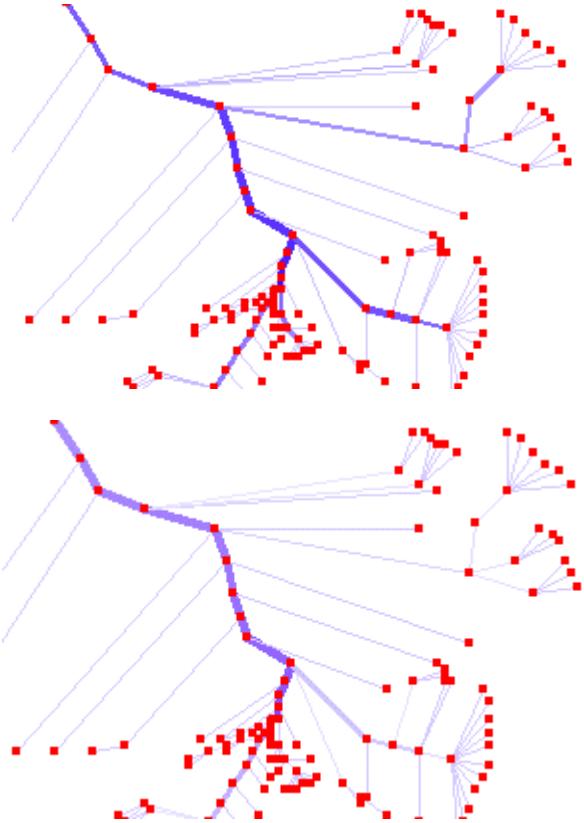


Figure 4. Enrichissement d'un arbre avec deux metriques differentes : Horton-Strahler (haut) et largeur (bas).

La métrique qui calcule la *largeur* d'un arbre, et qui compte le nombre de feuilles qu'il possède, suit aussi un calcul ascendant. Dans ce cas, la valeur d'un sommet est égal à la somme des valeurs de ses successeurs. On peut ici aussi altérer cette valeur par un poids. Le contexte peut lui ausi définir une métrique pour les éléments de la structure. On trouve dans un travail de Furnas [22] une métrique qui définit le *niveau d'intérêt* de l'utilisateur pour un élément dans la hierarchie, qui peut être une portion d'un texte (paragraphe, section ou chapitre) ou encore un bloc de code dans un programme. Mais dans tous les cas, la métrique utilisée doit permettre d'associer aux sommets des poids – ou des valeurs arbitraires - qui dépendent de l'application.

Le choix de la métrique lors de la visualisation de la structure dépend clairement de l'application et doit être laissé à l'utilisateur, qui seul peut déterminer si le relief qu'elle donne à la structure l'aide dans son analyse. La figure 4 illustre le même arbre enrichi par deux métriques différentes. L'algorithme de positionnement est cette fois radial [23]. Le premier arbre est enrichie avec la métrique Horton-Strahler ; le second enrichissement repose sur la métrique « largeur » de l'arbre (nombres de feuilles). On peut facilement constater les différences des indices visuels proposés : la métrique de Horton-Strahler donnent un relief qui descend beaucoup plus profondément dans les sous-arbres représentés ici. La valeur relative de chacune de ces métriques ne peut être déterminée que lors de leur utilisation dans un contexte particulier.

VUES SCHEMATIQUES

La visualisation de grands corpus de données nécessitent l'utilisation de méthode originale pour éviter les problèmes d'espace. Cette difficulté n'échappe pas aux structures arborescentes et a déjà reçu une attention de la part de nombreux chercheurs (cf, par exemple, [7,8,24]). Nous avons déjà mentionner les efforts faits dans la recherche de representation alternative pour vaincre le problème d'espace. Une autre approche consiste à filtrer la structure pour en extraire les composantes les plus importantes et n'afficher que celles-là. Certains algorithmes de « clustering » des graphes (cf [25,14]) peuvent être utiliser pour filtrer les données afin de pouvoir en présenter une vue non détaillée qui en minimise la complexité.

Dans le même ordre d'idée, on peut utiliser les valeurs calculées pour chacun des sommets d'un arbre à l'aide d'une métrique pour filtrer et diminuer l'ensemble des informations à visualiser. En effet, les métriques étudiées jusqu'à maintenant ont toutes en commun une propriété: les noeuds associés aux valeurs les plus élevées sont *de facto* les plus importants dans la structure. Une idée simple est alors de choisir une valeur socle pour décider si un sommet doit être affiché ou non. Dans le cas des arbres, l'affichage des noeuds ayant une valeur qui excède la valeur socle donne ce que nous appelons un « squelette » de l'arbre. L'utilisation d'un curseur permet dans ce cas de modifier interactivement la valeur socle pour faire varier le niveau de détail visuel de la structure. Cette idée est aussi présente dans le travail de Furnas [22] ; c'est ce qu'il désigne par « Fish-Eye Views » [of a graph].

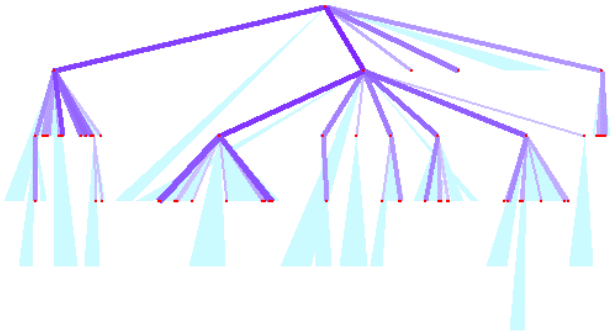


Figure 5. Vue « schématique » d"un arbre (cf Figure 1).

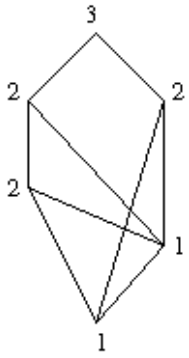
L'ensemble des sommets qui sont alors écartés appartiennent tous à des sous-arbres attachés aux feuilles du squelette. Une alternative est de remplacer chacun de ces sous-arbres par un triangle qui indique, par exemple, la surface qu'il occupe sur l'écran, ou le nombre de sommets qu'il contient (cf figure 5). On pourrait aussi en dessiner le contour; ce choix dépend parfois de l'application et du contenu des données.

GRAPHES ACYCLIQUES ORIENTES

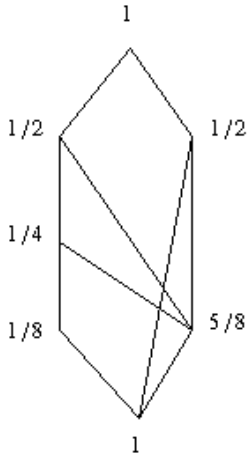
Les graphes acycliques orientés (souvent appelés DAGs) diffèrent des arbres en ce qu'un sommet peut avoir plusieurs ancêtres. Ils apparaissent naturellement dans les bases d'objets qui admettent l'héritage multiple. Ou encore pour décrire les scènes de réalité virtuelle; les ancêtres commun d'un sommet partagent alors la définition de cet objet dans la scène.

Metriques pour les DAGs

L'absence de cycle dans ces structures permet d'étendre la définition de métriques comme les nombres de Horton-Strahler. Il existe toujours dans les graphes acycliques orientés des sommets sans prédécesseurs ; ce sont les sommets de niveau 0, traditionellement placés au haut du diagramme. Le niveau d'un sommet est alors égal à la longueur maximale d'un chemin le reliant à un sommet de niveau 0. On peut alors définir la métrique de Horton-Strahler pour un DAG de la manière suivante : les sommets de niveau maxi-mum (au bas du diagramme) reçoivent la valeur 1 (tout comme les feuilles d'un arbre). La valeur d'un sommet interne est alors calculée avec la même formule, appliquée à l'ensem-ble de ses successeurs. La figure de gauche en donne un exemple.



D'autres métriques peuvent aussi être définies pour les DAGs. On peut, pas exemple, imaginer que les arcs sont des tuyaux dans lesquels on laisse couler de l'eau à partir des sommets du haut. On peut alors calculer, de manière similaire aux lois de Kirchoff en électricité, le flux qui passe en chaque sommet du graphe. Plus précisement, on convient de donner aux sommets de niveau 0 une valeur 1. Puis on calcule la valeur de la métrique en procédant par niveau. Les sommets de niveau 1 reçoivent de chacun de leurs ancêtres une contribution égale à la métrique de l'ancêtre divisée par son nombre de successeurs ; tout comme si l'eau passant par un sommet était équidistribué parmi ses successeurs. La figure ci-contre en donne un exemple.



On peut aussi combiner les métriques pour en définir de nouvelles ; on peut par exemple imaginer que le flux passant dans un sommet est réparti parmi les successeurs selon leur valeurs de Horton-Strahler, comme si les arcs représentaient des tuyaux de diamètres variables.

Attributs graphiques et indices visuels pour les DAGs

Cette mesure peut ensuite être utilisée pour définir des attributs graphiques des éléments du graphes. Comme pour les arbres, la métrique peut être utilisée pour définir une épaisseur pour les arcs et leur attribuer une couleur.

On peut aussi donner une vue schématique du graphe. La figure 6 montre un exemple de vue schématique pour un DAG. La métrique utilisée ici est celle du flux descendant décrite précédemment. Le positionnement des sommets du graphes se fait en adaptant l'algorithme de Reingold et Tilford [2, 3] aux DAGs comme suit. On calcule un arbre recouvrant du DAG en sélectionnant au départ les sommets de niveau 0. Puis on choisit itérativement parmi les sommets qui sont connectés aux sommets déjà sélectionnés, ceux qui se trouvent à distance minimale. Reste ensuite à prendre soin que l'ordonnée d'un sommet du DAG soit égale à son niveau dans le DAG (qui peut être différent de sa profondeur dans l'arbre recouvrant).

Dans le cas des arbres, les sommets écartés par la valeur socle avaient été remplacés par des triangles. Il est

cependant plus difficile, dans le cas des DAGs, de définir une forme géométrique qui rendent compte des parties cachées du graphe. En effet, d'une part, la structure de DAGs donne lieu à des ensembles d'arcs plus complexes. D'autre part, le « squelette » d'un arbre est relativement simple ; dès qu'un sommet est sous la valeur socle il en va de même pour ses descendants. Cela tient au fait que les métriques que nous avons définies pour les arbres sont toujours croissantes en direction de la racine. Pour un DAG, la métrique du flux descendant ne possède pas cette propriété.

Nous avons donc conçu un artifice différent. On a cette fois fait ressortir les sommets sélectionnés et les arcs qui les connectent en leur attribuant une couleur et une intensité en fonction du flux qui les traversent. Les autres sommets et arcs ont été dessiner en fond et dessinés en grisés. La représentation du DAG permet alors de saisir la globalité de la structure et de situer les éléments essentiels de la structure dans leur contexte.

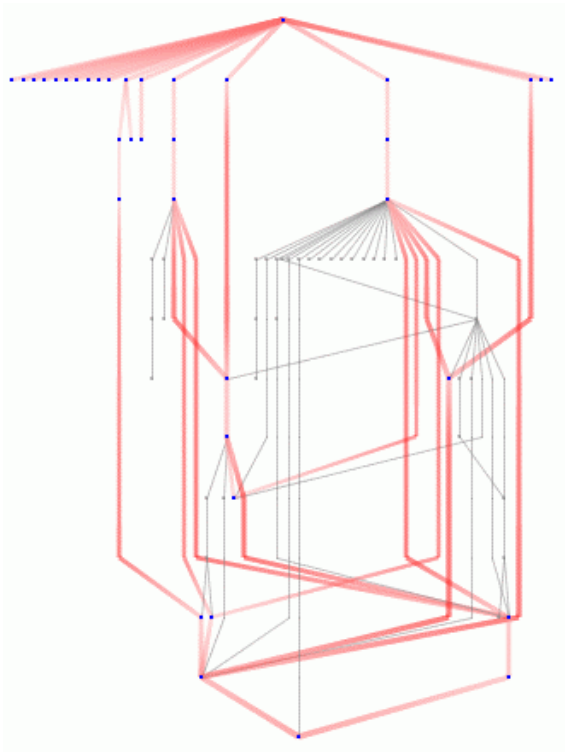
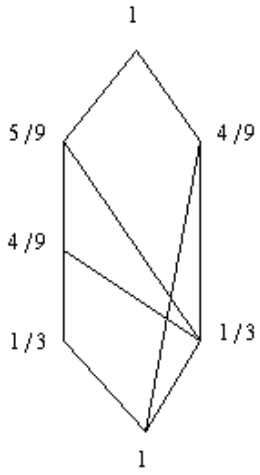


Figure 6. Vue schématique d'un DAG (métrique du flux descendant)

Cette technique s'est avéré utile pour analyser des DAGs qui apparaissent naturellement lors de déploiement de réseaux de Pétri. Les sommets de flux maximaux correspondent alors aux états bloquants du réseaux. On a pu aussi analyser le contenu de site web ; dans les exemples analysés, les sommets mis en exergue corrrespondaient aux pages les plus susceptibles d'être visiter lors de l'exploration du site.

Autres métriques

Comme on l'a mentionné, l'absence de cycle dans le DAG permet d'étendre de manière naturelle les métriques définies pour les arbres. En effet, il est alors possible de traverser le graphe récursivement ou par niveau. Mais les DAGs possède une symétrie verticale ; en d'autres mots il peuvent être traversé en partant du bas. Cet ordre de traversée n'a pas de sens pour un arbre, mais peut être pertinente dans le cas d'un DAG. Cela permet alors de « dualiser » toutes les métriques qui reposent sur un parcours descendant du DAG. On peut par exemple définir une métrique du flux montant dans le DAG. On obtient alors des valeurs différentes, comme le montre la figure ci-contre.



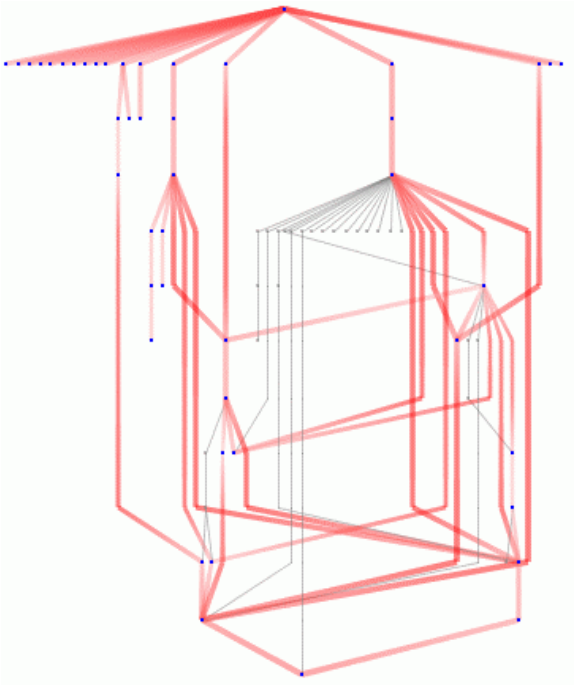


Figure 7. Vue schématique d'un DAG (moyenne des flux montant et descendant).

Une combinaison qui semble alors naturelle est de calculer le flux moyen en tout sommet. La figure 7 illustre le même graphe qu'à la figure 6, cette fois enrichie par la métrique du flux moyen. On peut noter non seulement une différence au niveau de l'épaisseur et de la couleur des traits dans le dessin, mais aussi la présence de certains nouveaux sommets dans le « squelette ». Cela peut s'expliquer par la symétrie relativement forte dans l'exemple. Ces noeuds, alors qu'ils sont moins importants dans le flux descendant, retrouvent leur importance dans le flux montant.

CONCLUSIONS ET PERSPECTIVES

Les résultats présentés ici ont tous été implémentés dans le cadre du développement d'un système de visualisation d'arbres - *Latour*. Ce système permet aussi, dans une certaine mesure, de manipuler les graphes acycliques orientés et est actuellement utilisé et/ou testé pour visualiser des structures hiérarchiques dans les domaines suivants:

Structures de donnés internes de programmes (dans notre cas, celles d'un compilateur paramétrables produit par une entreprise hollandaise ACE b.v., à Amsterdam).

Structures de données décrivant des scènes de réalités virtuelles.

Déploiements de grands réseaux de Pétri.

Suivis d'exécution d'applications massivement parralèles basées sur un langage de coordination.

Structures de site web.

Le défi qui se pose de manière la plus évidente est l'extension des métriques aux graphes plus généraux. Les métriques de flux utilisés pour les graphes acycli-ques orientés s'apparentent aux méthodes de la théorie des chaînes de Markov et semblent offrir des pistes à explorer. Incidemment, la manipulation de grands graphes nécessitent souvent d'appliquer une technique de « clustering ». Nos méthodes peuvent alors être étendues pour définir des outils de navigation construits autour de la manipulation d'un graphe décrivant la structure des « clusters ».

REMERCIEMENTS

Nos remerciements vont à B. de Ruiter qui a contribué de manière importante au développement du logiciel *Latour* et nous a ainsi permis de concentrer nos efforts sur l'expérimentation avec les métriques. Les remarques de plusieurs collègues - J. P. Domenger, M.S. Marshall, D. Duke, et les commentaires des gens de ACE, ont aussi guidé notre travail d'expérimentation.

BIBLIOGRAPHIE

1. Whitesides S. ed., *Graph Drawing* '98, 6th *International Symposium*, Montreal, Canada, vol. 1547 of Lectures Notes in Computer Science, Springer Verlag, 1998.
2. E.M. Reingold and J.S. Tilford: "Tidier Drawing of Trees". In*: IEEE Transactions on Software Engineering*, SE-7(2), pp. 223-228, March 1981.
3. Walker J.Q., A Node-Positioning Algorithm for General Trees}. In: *Software - Practice and Experience* 20(7), pp.685 -- 705, 1990.
4. Robertson G. G., Mackinlay J. D., Stuart S. K. "Cone Trees : Animated 3D Visualizations of Hierarchical Information". In : *Proceedings of CHI '91, ACM Conference on Human Factors in Computing Systems*, New York, pp. 189-194, 1991.
5. J. Carrière, R. Kazman: "Interacting with huge hierarchies : beyond cone trees"; in: *IEEE Information Visualisation* '95, IEEE CS Press, 1995, 74-81.
6. Wiss U., Carr D., Jonsson H., Evaluating Three-Dimensional Information Visualization Designs: a Case Study of Three Designs. In: IV '98 *Conference Proceedings on Information Visualization*, pp. 137-144. IEEE Computer Society Press, 1998.
7. Beaudoin L., Parent M.-A., Vroomen L. C., Cheops: A Compact Explorer For Complex Hierarchies. In: *Visualization 1996: proceedings of the seventh annual IEEE conference*, Yagel R., Nielson G. M. eds, I.E.E.E. Computer Society Press, 1996.
8. T. Munzner, "H3: Laying out Large Directed Graphs in 3D Hyperbolic Space". In: *Proceedings of the 1997 IEEE Symposium on Information Visualization*, IEEE CS Press, Los Alamitos, 1997.
9. T. Munzner, "Exploring Large Graphs in 3D Hyperbolic Space". In: *IEEE Computer Graphics & Applications*, July/August 1998.
10. Schneiderman B., Tree visualization with tree maps: a 2D space-filling approach. In: *ACM Transactions on Graphics* 11 (1), pp. 92-99, 1992.
11. Durand D., Kahn P., MAPA, In: *Ninth ACM Conference on Hypertext and Hypermedia* (*Hypertext*'98), Pittsburgh, USA, June 20-24, 1998.

12. Mackinlay J.D., Robertson G.G., Card S.K., The perspective wall: detail and context smoothly integrated, *Proceedings of CHI '91 Human Factors in Computing Systems*, pp. 173-179, ACM Press, 1991.

13. Schneiderman B., The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization. In: *Proceedings of the 1996 IEEE Visual Languages*, pp. 336-343, 1996.

14. G. di Battista, P. Eades, R. Tamassia and I.G. Tollis: *Graph Drawing: Algorithms for the Visualisation of Graphs*, Prentice Hall, 1999.

15. Gansner E.R., North S.C., Vo K.P., DAG - A Program that Draws Directed Graphs, *Software - Practice and Experience* 18 (11), pp. 1047-1062, 1988.

16. P. Eades and K. Sugiyama: "How to draw a Directed Graph". In: *Journal of Information Processing*, 13(4), pp. 424-437, 1990.

17. Knuth D. E. *The Art of Programming*. Addison-Wesley, Reading (MA), 1973.

18. Horton R. E., Erosioned development of steams and their drainage basins, hydrophysocal approach to quantitative morphology, *Bulletin of the Geologic Society of America* 56, pp. 275-370, 1945.

19. Strahler A. N., Hypsometric (area-altitude) analysis of erosional topology, *Bulletin of the Geologic Society of America* 63, pp. 1117-1142, 1952.

20. Viennot X.G., Eyrolles G., Janey N., Arquès D., Combinatorial analysis of ramified patterns and computer imagery of trees. In: *Computer Graphics* (*SIGGRAPH* '89) (23), pp. 31-40, 1989.

21. I. Herman, M. Delest and G. Melançon, "Tree visualisation and navigation clues for information visualisation". In: *Computer Graphics Forum*, 17(2), pp. 153-165, 1998.

22. Furnas, G. W., Generalized fisheye views. *Human Factors in Computing Systems CHI '86 Conference Proceedings*, Boston, April 13-17, 1986, 16-23.

23. Eades P. Drawing Free Trees. In : *Bulletin of the Institue for Combinatorics and its Applications* 5, pp. 10-36, 1992.

24. M.C. Chuah, "Dynamic Aggregation with Circular Visual Designs". In: *Proceedings of the IEEE Symposium on Information Visualization* (*InfoVis*'98), G. Wills and J. Dill (eds.), IEEE CS Press, Los Alamitos, pp. 35-43, 1998.

25. Jain A. K., Dubes R. C., *Algorithms for clustering data*, Prentice-Hall, Englewood Cliffs (NJ), 1988.