

Génération intégrée de textes et de graphiques statistiques  
Massimo Fasciano  
Guy Lapalme  
Département d'informatique et de recherche opérationnelle  
Université de Montréal  
C.P. 6128, Succ. Centre-Ville  
Montréal, Québec, Canada, H3C 3J7  
e-mail: {fasciano,lapalme}@iro.umontreal.ca}

## Résumé

Les graphiques et le texte sont deux médias très différents. Heureusement, lorsque leur intégration est bien effectuée, ils se complètent à merveille: une image permet de montrer alors qu'un texte permet de décrire. Dans le cadre de notre recherche, nous étudions l'interaction entre le texte d'un rapport statistique et ses figures. Nous nous intéressons aussi à l'influence de l'intention du rédacteur(ses buts) sur un rapport. Notre but final est la réalisation d'un système capable de générer automatiquement des rapports statistiques contenant du texte et des graphiques à partir de données brutes.

### 1 Introduction

#### 1.1 Qu'est-ce qu'un rapport?

Le but du rapport est de produire une synthèse organisée à partir de données ou expériences concrètes. Bien qu'en principe un rapport doive être objectif, cette synthèse est quelquefois biaisée. Le rapport a un but bien précis, il cherche à faire passer le point de vue de la personne ou organisation qui le rédige. Pour arriver à ce but, il faut souvent “tordre” un peu les faits, c'est-à-dire mettre en évidence ceux qui supportent nos conclusions et en éclipser d'autres.

Le niveau d'organisation d'un rapport est plus élevé que celui d'un texte narratif. En effet, le rapport doit être structuré de façon à ce que son analyse ait plus de poids. Pour un texte narratif, la structure est plus libre et suit l'événement qui est raconté. Les rapports n'offrent toutefois pas une structure aussi stricte que les textes procéduraux: tous les manuels d'instructions sont structurés de façon similaire, de même pour les recettes de cuisine.

On peut classer les rapports selon le niveau d'organisation de l'information qu'ils présentent. En effet, un rapport peut être construit à partir de données brutes comme des tableaux de nombres ou d'une information beaucoup plus structurée telle un autre texte. Ce dernier type de rapport peut être un résumé d'un texte plus long ou une critique de celui-ci. Ces exemples montrent deux extrêmes, mais il existe d'autres rapports qui partent de données plus ou moins organisées.

Cette classification est très importante car elle indique le degré de liberté dont on dispose pour produire un rapport. Il est clair qu'il y a beaucoup plus de choix en partant de nombres que d'un texte organisé. Un résumé suivra l'ordre et la structure du texte original; on se limite alors au choix de l'information essentielle.

Les rapports qui nous intéressent, les rapports statistiques, ont une particularité intéressante: les données sont nombreuses et peu évocatrices lorsque présentées de façon brute. Sans une analyse statistique préalable pour faire ressortir les points importants ainsi qu'une organisation et présentation très efficaces, on risque de perdre le lecteur.

#### 1.2 Intégration de graphiques au texte

Les graphiques et le texte sont deux médias très différents. Heureusement, lorsque leur intégration est bien effectuée, ils peuvent se compléter à merveille: une image permet de montrer alors qu'un texte permet de décrire. Ainsi, un rapport qui présente des données peut se servir d'images pour les montrer sous leur forme brute et du texte pour en faire l'analyse.

Évidemment, la frontière montrer/analyser n'est pas absolue et il est tout à fait possible de se servir d'outils graphiques pour présenter les résultats d'une analyse, mais sans un accompagnement textuel, l'analyse risque de ne pas atteindre son plein potentiel.

Inversement, le texte peut aussi servir à “montrer” par une description très détaillée. Cependant, on dit qu'une image vaut mille mots, et dans une telle situation, il est beaucoup plus efficace d'utiliser la puissance visuelle des graphiques plutôt que d'alourdir le rapport avec de longues descriptions. Ainsi, il est préférable de montrer une courbe plutôt que de lister la valeur des points qui la composent. D'un seul coup d'oeil, on voit les tendances importantes sur la courbe.

Dans le cas des rapports statistiques, on dispose d'outils graphiques comme les courbes, les tartes, les barres/colonnes qui sont adaptés à des types particuliers de messages. Par exemple, pour montrer la décomposition d'un ensemble de données, on peut les présenter sous forme de tarte, où l'angle d'un secteur indique la proportion occupée par l'élément associé. C'est efficace car le lecteur décode d'un simple coup d'oeil une information qui ne serait pas aussi évidente si les données étaient sous une forme brute. On peut ensuite associer un paragraphe de texte au graphique pour faire passer notre message de façon plus précise. Par exemple, la phrase “La part de marché de notre compagnie dépasse le total des parts de nos concurrents” donnerait au lecteur une interprétation des données qu'il n'aurait peut-être pas eue en ne regardant que le graphique.

### 1.3 Un exemple

L'information à présenter dans un rapport statistique est souvent donnée sous une forme peu parlante. Par exemple, le tableau 1 contraste l'évolution de 2 compagnies (Xyz inc. et Pqr inc.) entre 1985 et 1994.

année	chiffre d'affaires	
	Xyz inc.	Pqr inc.
1985	50	75
1986	60	79
1987	62	76
1988	64	71
1989	69	72
1990	76	70
1991	83	72
1992	89	83
1993	93	98
1994	102	120

Tableau 1: Évolution des profits de Xyz. et Pqr.

Il est très difficile d'avoir une vue d'ensemble sur les données lorsqu'elles sont présentées sous cette forme. On a beaucoup de difficulté à visualiser l'évolution du chiffre d'affaires de chaque

compagnie et il est encore plus laborieux de les comparer. Cependant, on peut lire avec précision les valeurs individuelles.

Pour donner un sens à cette information, on dispose de 2 outils: le texte et les graphiques. La figure 1 montre comment on pourrait l'analyser en utilisant une description textuelle.

Le chiffre d'affaires de Xyz inc. a augmenté de 52 millions entre 1985 et 1994. Celui de Pqr inc. a augmenté de 45 millions. L'augmentation a été graduelle pour Xyz, alors que celle de Pqr a eu lieu surtout entre 1991 et 1994 (48 millions) après quelques fluctuations les années précédentes. Pqr a commencé et fini en tête mais a été légèrement dépassée entre 1990 et 1992 par son compétiteur.

Figure 1: Description textuelle pour les données du tableau 1

Cette description rend les données du tableau beaucoup plus utiles mais elle ne résoud pas tous les problèmes. En effet, le texte est un peu long, ce qui risque de perdre le lecteur. Celui-ci n'en dégagera donc pas les points essentiels. Il faut essayer de le simplifier sans affaiblir le message. De plus, certaines affirmations du texte sont assez difficiles à vérifier à l'aide d'un simple coup d'oeil au tableau.

L'ajout d'un graphique au texte permet de le simplifier en transférant certains éléments quantitatifs vers le graphique. Par exemple, il devient inutile<sup>1</sup> de quantifier les augmentations ou d'indiquer si elles ont été graduelles. De plus, tout ce qui est dit dans le texte se vérifie d'un simple coup d'oeil et attire beaucoup plus l'attention du lecteur car l'information n'est pas noyée dans un texte trop long.

La figure 2 présente un cas où le texte montre les points importants de l'analyse, alors que le graphique s'occupe de donner une vue d'ensemble.

Dans d'autres situations, on peut ajouter un texte à un ensemble de graphiques pour établir un lien entre ceux-ci. Cette situation est assez courante lorsqu'on cherche à présenter plusieurs "vues" sur les mêmes données afin de mieux exposer toutes les tendances. À la figure 3, on commence par présenter l'évolution de 2 données pour ensuite montrer qu'elles sont fortement corrélées. Le texte évite donc de "parachuter" une série de graphiques sans description. Il sert à décrire l'utilité de chaque graphique plutôt qu'à compléter ceux-ci par une analyse.

---

<sup>1</sup>à moins de vouloir fortement insister sur ces faits.

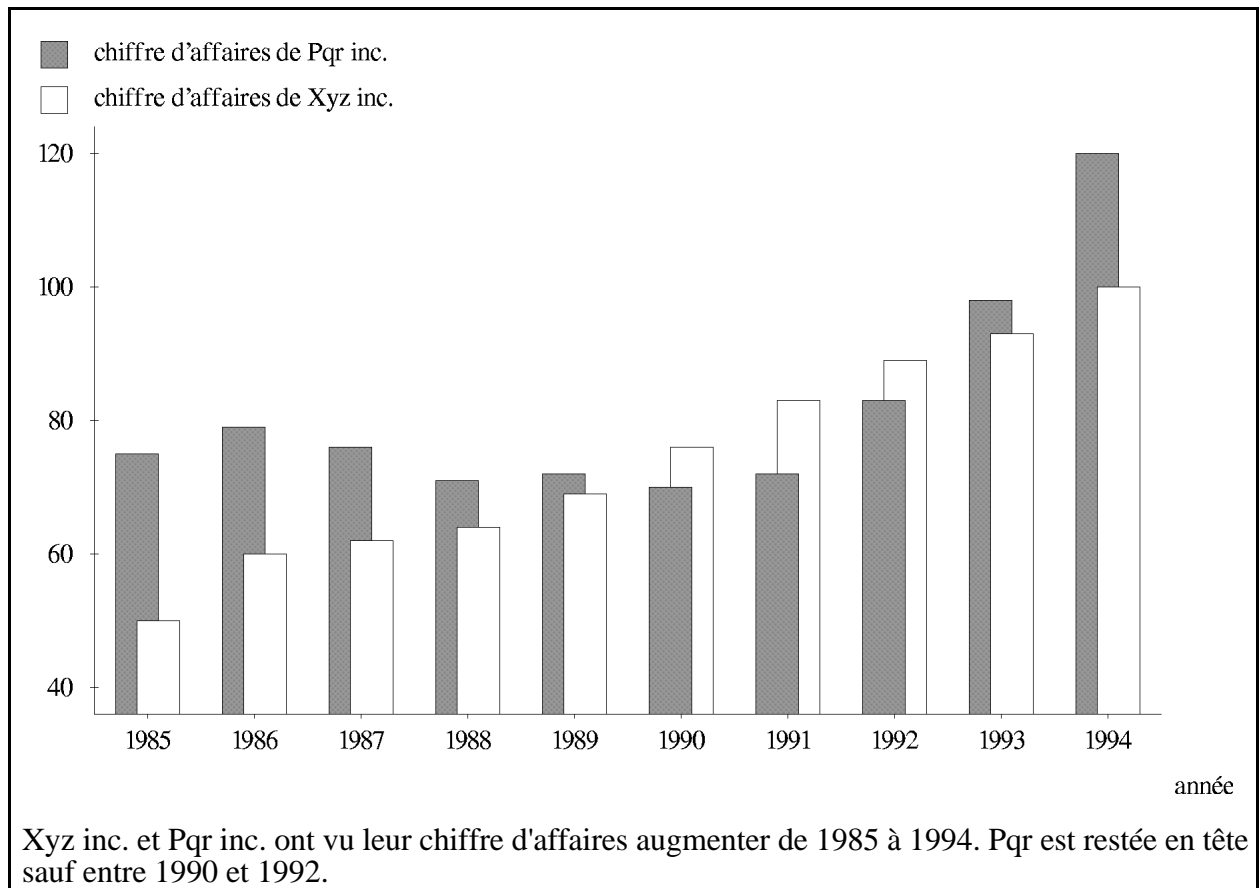
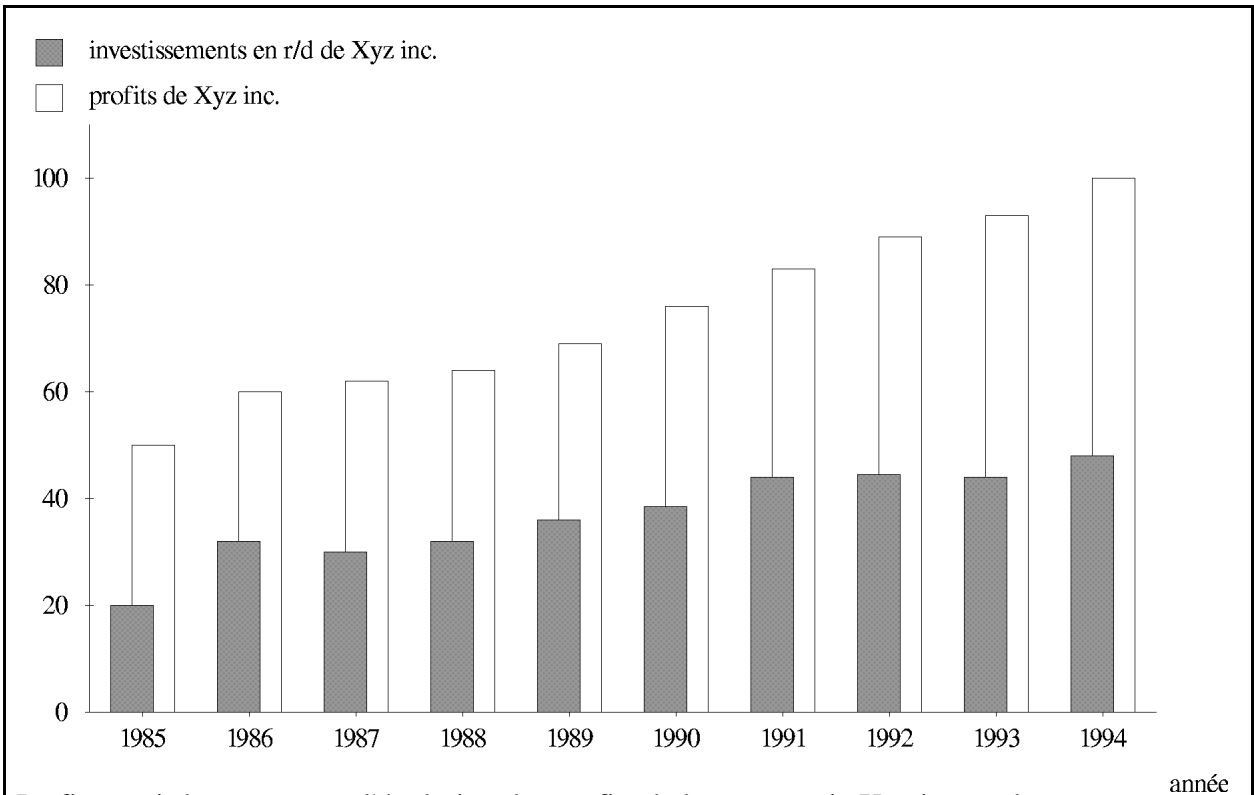


Figure 2: Combinaison texte/graphique pour les données de la figure 1



La figure ci-dessus montre l'évolution des profits de la compagnie Xyz inc. et de ses investissements en recherche et développement entre 1985 et 1994. Tous deux ont augmenté graduellement pendant la période en question (50 et 28 millions respectivement). Il semblerait que l'augmentation des profits soit liée à l'augmentation des investissements en r/d. Cette tendance est confirmée par la figure ci-dessous.

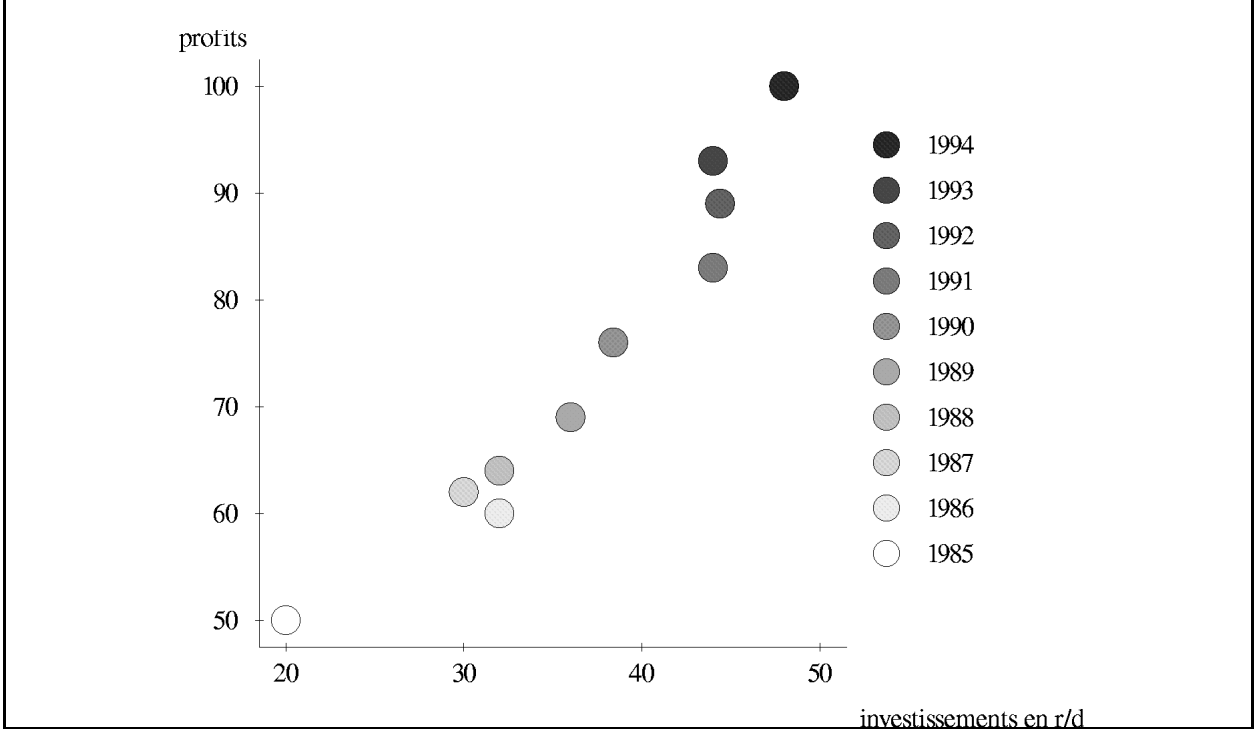


Figure 3: Combinaison d'un texte et de plusieurs graphiques

## 2 Caractéristiques des graphiques

Parmi les études sur le choix de graphiques statistiques, nous nous sommes attardés sur celle de MacKinlay (1986) qui a l'avantage de fournir un algorithme basé sur une extension de la classification de Bertin (1983) (variables nominales, ordinales et quantitatives) et qui définit un ensemble de méthodes graphiques pour l'expression de chaque type de variable. Ces méthodes sont ordonnées et l'algorithme fait l'allocation des méthodes aux variables en évitant de générer des solutions impossibles (par exemple, 3 coordonnées spatiales dans un graphique à 2 dimensions).

Nous avons aussi intégré à notre modèle des résultats plus théoriques comme ceux de Tufte (1983, 1990) et de Zelazny (1989).

On retient de ces travaux une classification des propriétés d'un graphique tant au niveau de ses composantes élémentaires (Bertin, MacKinlay) qu'au niveau global (Zelazny). Les propriétés des composantes élémentaires jouent surtout sur l'efficacité d'un graphique pour exprimer un type donné de variable. Par exemple, les positions spatiales sont plus utiles que les couleurs pour exprimer des variables continues. Les caractéristiques globales d'un graphique sont surtout utiles pour déterminer son rôle dans la transmission du message. Ainsi, des graphiques dont les éléments de base sont similaires peuvent avoir des rôles un peu différents (les barres sont adéquates pour comparer des valeurs, alors qu'on utilise toujours les colonnes ou les courbes pour des données temporelles).

Notre classification des propriétés intègre ces résultats et les applique à un ensemble beaucoup plus général de graphiques. Au niveau des propriétés élémentaires, nous retenons l'efficacité (facilité et précision de lecture) pour chaque variable ainsi que la possibilité de comparer soit les valeurs d'une même variable entre elles ou bien les valeurs de plusieurs variables.

Au niveau de propriétés globales on retrouve la taille des graphiques utilisés pour présenter les données ainsi que des facteurs liés au message à présenter. Connaître la taille d'un ensemble de graphiques est très utile pour déterminer s'ils seront utiles dans un contexte donné. Par exemple, pour une courte introduction à un ensemble de données on voudrait éviter de remplir plusieurs pages de figures avant même de commencer le texte.

Pour ce qui est du message<sup>1</sup>, il est important de savoir si le graphique transmet bien les informations suivantes:

- évolution des valeurs
- corrélation entre les variables
- répartition des valeurs
- décomposition des valeurs

Parfois, la sélection d'un graphique peut se faire à l'aide d'un critère global. On peut voir à la figure 4 qu'une tarte permet de montrer la décomposition d'un total (ici un pourcentage) beaucoup mieux que des barres. Bien qu'elle soit moins utile pour la sélection, on peut noter aussi une différence au niveau de l'efficacité des 2 graphiques: les barres permettent une lecture et une comparaison des valeurs beaucoup plus précise que la tarte (par exemple, Genco et Xyz inc.).

---

<sup>1</sup>Bien que la comparaison soit considérée comme un message dans notre système, elle est utile à un niveau plus élémentaire dans la sélection des graphiques.

Dans d'autres situations, on ne peut pas compter sur le rôle du graphique. C'est le cas à la figure 5 où on retrouve des petites différences liées à la lecture et à la comparaison de valeurs mais aucune au niveau du message. Ainsi, on remarque que la figure du haut permet une lecture précise des valeurs à l'aide de son échelle linéaire alors que celle du bas utilise des surfaces, méthode qui ne permet que des lectures relatives. Par contre, celle du bas est plus orthogonale dans ses comparaisons. En effet, il est aussi facile de comparer selon les mois que selon les provinces, alors que la figure du haut encourage plutôt la comparaison entre les mois pour une province donnée (localité des colonnes). Évidemment, la comparaison orthogonale de la figure inférieure n'est possible que lorsque les valeurs sont très différentes car l'oeil est peu sensible aux différences de surface.

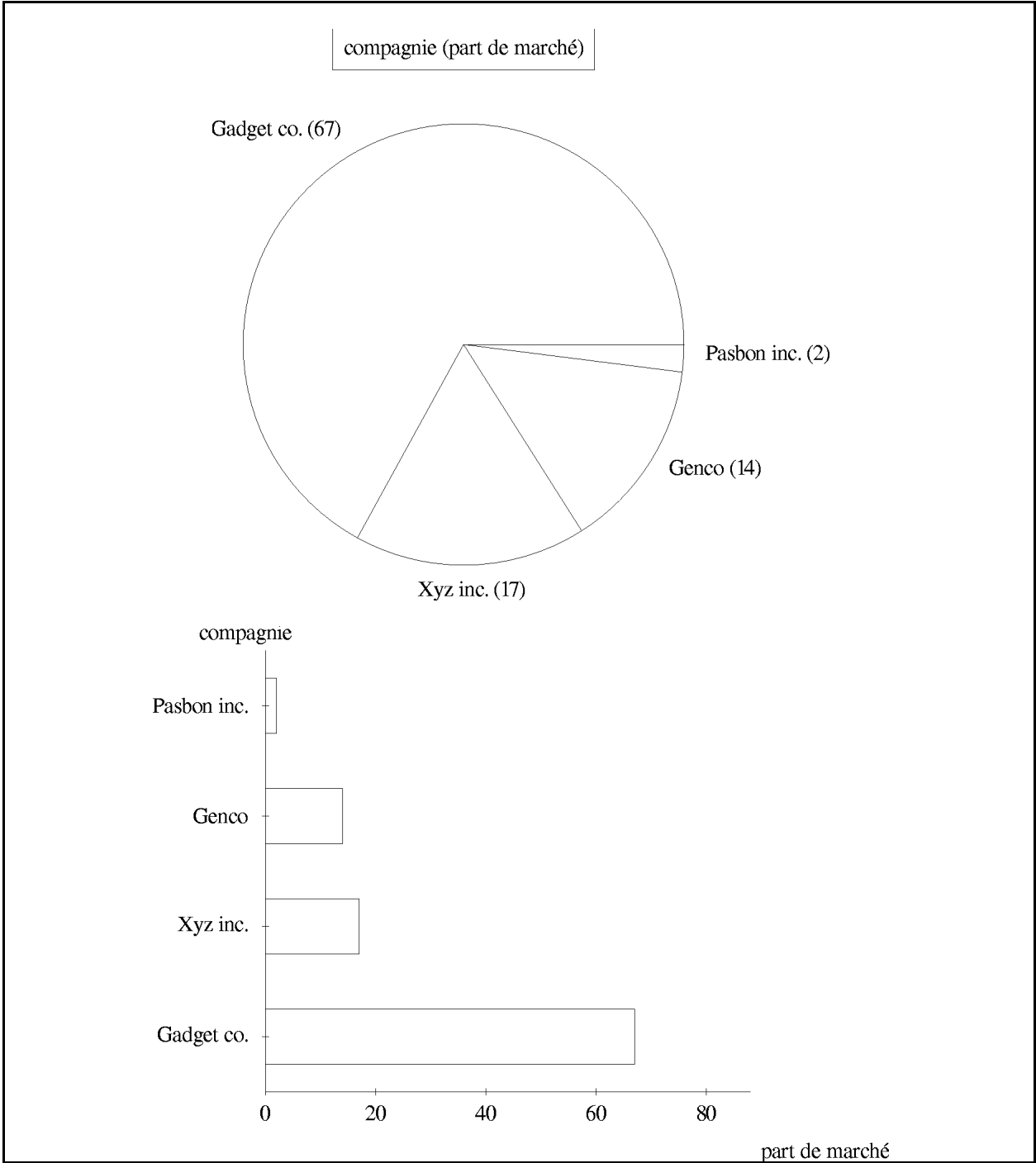


Figure 4: Comparaison de l'efficacité d'une tarte et d'un graphique en barres



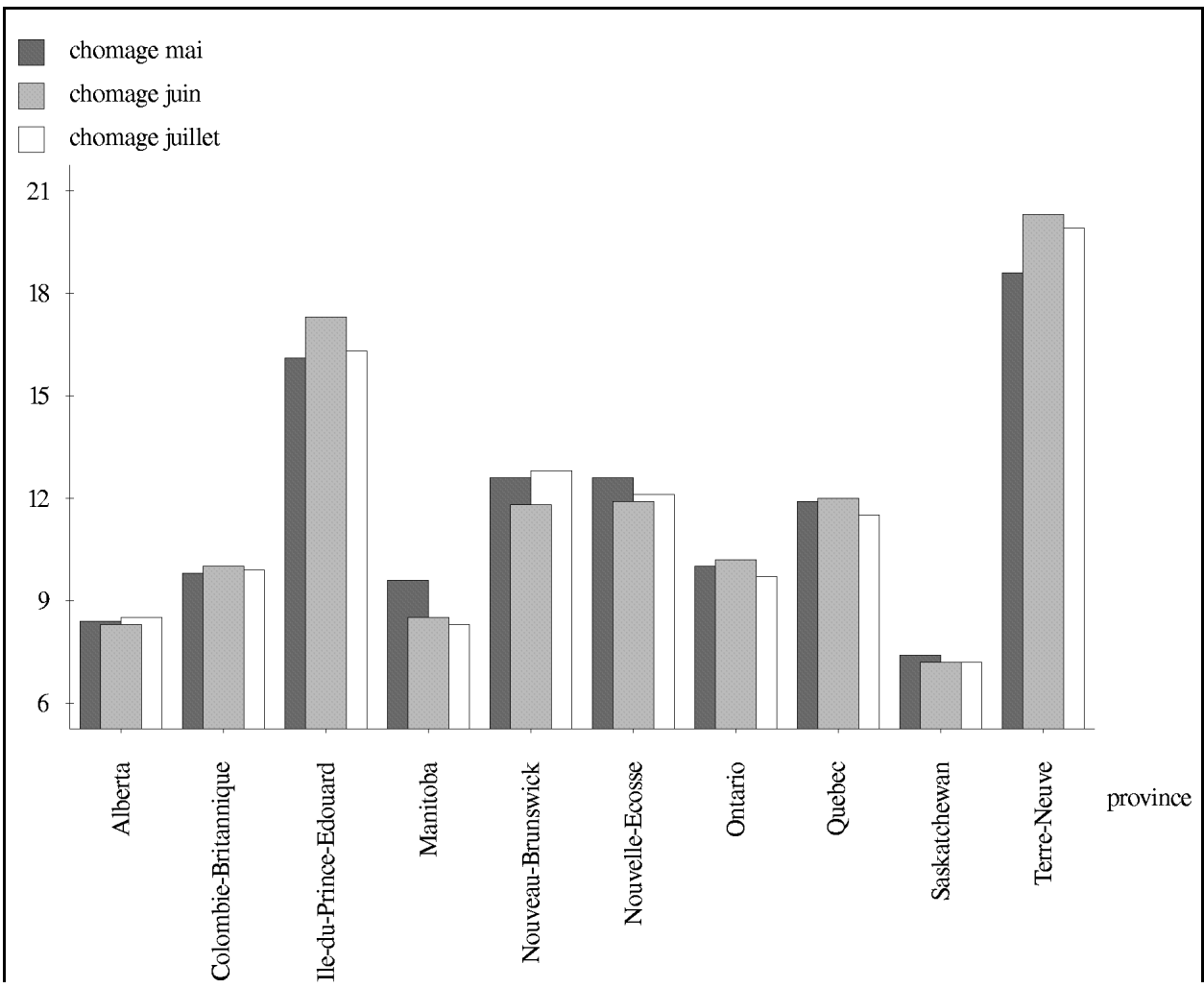




Figure 5: Comparaison de l'efficacité des colonnes et des surfaces (graphique à 3 variables)

### 3 Intention du rédacteur

Alors que la recherche porte d'habitude sur le modèle du lecteur (Paris, 1991), nous étudions plutôt celui du rédacteur du rapport. Les rapports sont parfois et même souvent biaisés par les intentions et les buts du rédacteur. On doit donc tenir compte des rapports objectifs mais aussi des rapports subjectifs dans lesquels on essaie de faire “mentir” un peu les données. L'intention du rédacteur affecte également la façon dont les graphiques et le texte sont combinés. Cet aspect du problème a été très peu étudié dans le domaine.

#### 3.1 Classification

L'intention du rédacteur affecte directement le message à présenter dans un rapport. Une très bonne étude sur la correspondance entre le message et le graphique a été faite par Gene Zelazny (1989). Dans cet ouvrage, il identifie 5 messages<sup>1</sup> de base (la décomposition, la position, l'évolution, la répartition, la corrélation) et 5 graphiques (voir figure 6) très utilisés dans les rapports et il explique le lien entre ces deux ensembles.

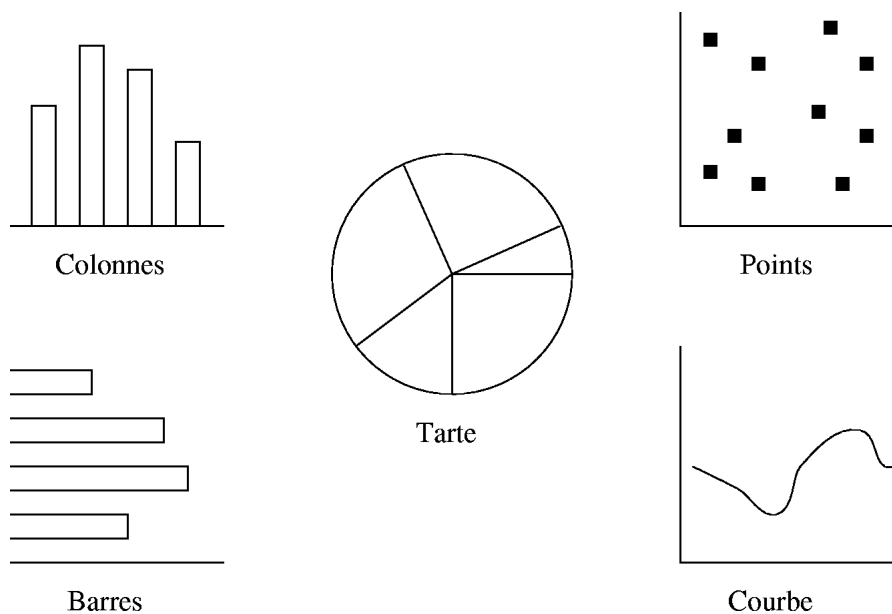


Figure 6: Les 5 graphiques de base selon Zelazny

##### 3.1.1 La décomposition

Dans une décomposition, on cherche à montrer des valeurs en tant que fraction d'un tout. La tarte, de par sa nature, est presque toujours le meilleur graphique pour montrer une décomposition.

##### 3.1.2 La position

---

<sup>1</sup>Zelazny les appelle “types de comparaison”, mais il affirme qu'ils correspondent directement aux messages que le rédacteur peut transmettre.

Ce type de message s'applique au classement d'éléments entre eux. Par exemple, comparer les profits de plusieurs entreprises. Le graphique à barres est le plus approprié pour ce type de message. Il rend la comparaison des valeurs facile (surtout si elles sont triées). Le graphique en colonnes est à éviter car il donne une impression d'évolution (on perçoit mieux le temps sur un axe horizontal).

### 3.1.3 L'évolution

Ce type de message vise à montrer des changements dans le temps, comme par exemple, la variation des profits d'une entreprise au fil des années. Le graphique en colonnes (données non-continues) et la courbe (données continues) sont appropriés pour ce type de message.

### 3.1.4 La répartition

C'est une répartition de fréquence dans laquelle on cherche à montrer combien d'échantillons se trouvent dans chacun des intervalles considérés. Les graphiques qui correspondent à ce type de message sont les colonnes et la courbe.

### 3.1.5 La corrélation

Une comparaison de corrélation montre la relation entre 2 variables. Elle permet de tester si le comportement d'une variable est lié à celui d'une autre. On utilise souvent le graphique en points (avec une courbe de régression) pour montrer une corrélation.

Cette étude est un très bon point de départ mais elle laisse beaucoup à désirer à deux points de vue: tout d'abord les 5 graphiques traités par l'auteur ne sont pas suffisants pour de vrais rapports, et ensuite les messages utilisés, bien que très utiles semblent manquer d'organisation (décomposition et position ont des points communs qui ne sont pas traités, évolution est trop générale, etc...). Nous avons donc étendu son étude à des graphiques plus complexes tout en organisant ses messages en une hiérarchie qui en illustre mieux les points communs.

La figure 7 présente cette hiérarchie et la figure 8 montre la différence entre 4 messages spécifiques de la branche "évolution": augmentation, diminution, stagnation et récapitulation. Les données décrivent la variation des profits (en millions de \$) d'une compagnie entre 1971 et 1978.

- comparaison
  - décomposition
  - position
- évolution
  - variation
    - \* augmentation
    - \* diminution
  - stagnation
  - récapitulation
- corrélation
  - forte
  - faible
- répartition
  - intervalles fixes (classes d'âge)
  - répartition désirée (uniforme, cloche, ...)

Figure 7: Hiérarchie des messages

<b>Année</b>	<b>Profits</b>
1971	10
1972	11
1973	11
1974	18
1975	16
1976	16
1977	15
1978	19

**AUGMENTATION:** Les profits de la compagnie ont augmenté de 8 millions entre 1971 et 1974, puis de 3 millions entre 1975 et 1978 après une chute de 2 millions.

**DIMINUTION:** Les profits de la compagnie ont diminué de 3 millions entre 1974 et 1977.

**STAGNATION:** Les profits de la compagnie ont stagné entre 1971 et 1973, puis entre 1975 et 1977 après une grosse augmentation en 1974.

**RÉCAPITULATION:** Les profits de la compagnie ont subi 3 ans d'augmentation, 2 ans de stagnation et 2 ans de diminution entre 1971 et 1978.

Figure 8: Adaptation du texte à l'intention du rédacteur

### 3.2 Effet

L'intention du rédacteur a un effet très important sur l'utilisation du texte et des graphiques. Cette influence n'est pas tout à fait la même dans un rapport subjectif que dans un rapport objectif.

Au niveau des rapports objectifs, l'effet de l'intention du rédacteur est assez claire. Certains graphiques sont mieux adaptés à certains messages (cf. Zelazny):

- décomposition: tarte
- évolution: colonnes ou courbe
- corrélation: points

Au niveau du texte, le contenu change complètement, comme on peut le voir à la figure 9. Cet exemple présente les mêmes données (profits d'une compagnie entre 1971 et 1976) à travers deux messages différents. La partie supérieure montre les données à travers un message d'évolution alors que celle du bas insiste plutôt sur la comparaison des années.

Pour ce qui est des rapports plus subjectifs, les effets sont beaucoup plus subtils et dépendent aussi du lecteur. Ce qui est certain, c'est qu'un graphique subjectif est plus difficile à réaliser qu'un texte subjectif car toute omission dans un graphique est suspecte. Dans un texte, l'omission est pratique courante (pour le rendre plus compact). Toutefois, comme on peut le voir à la figure 10, 2 graphiques similaires peuvent tout de même transmettre des messages très différents.

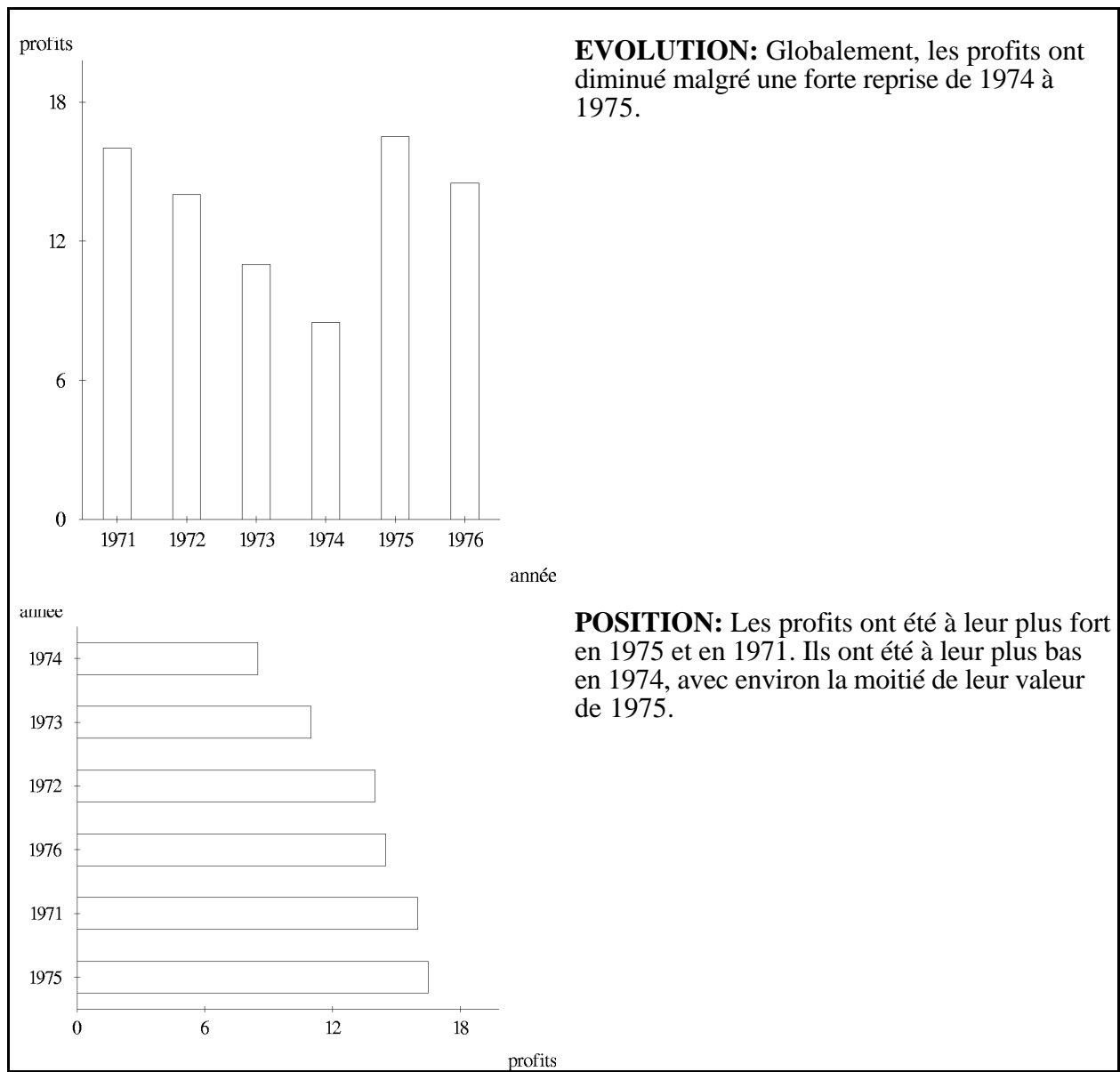


Figure 9: Effet des messages évolution et position



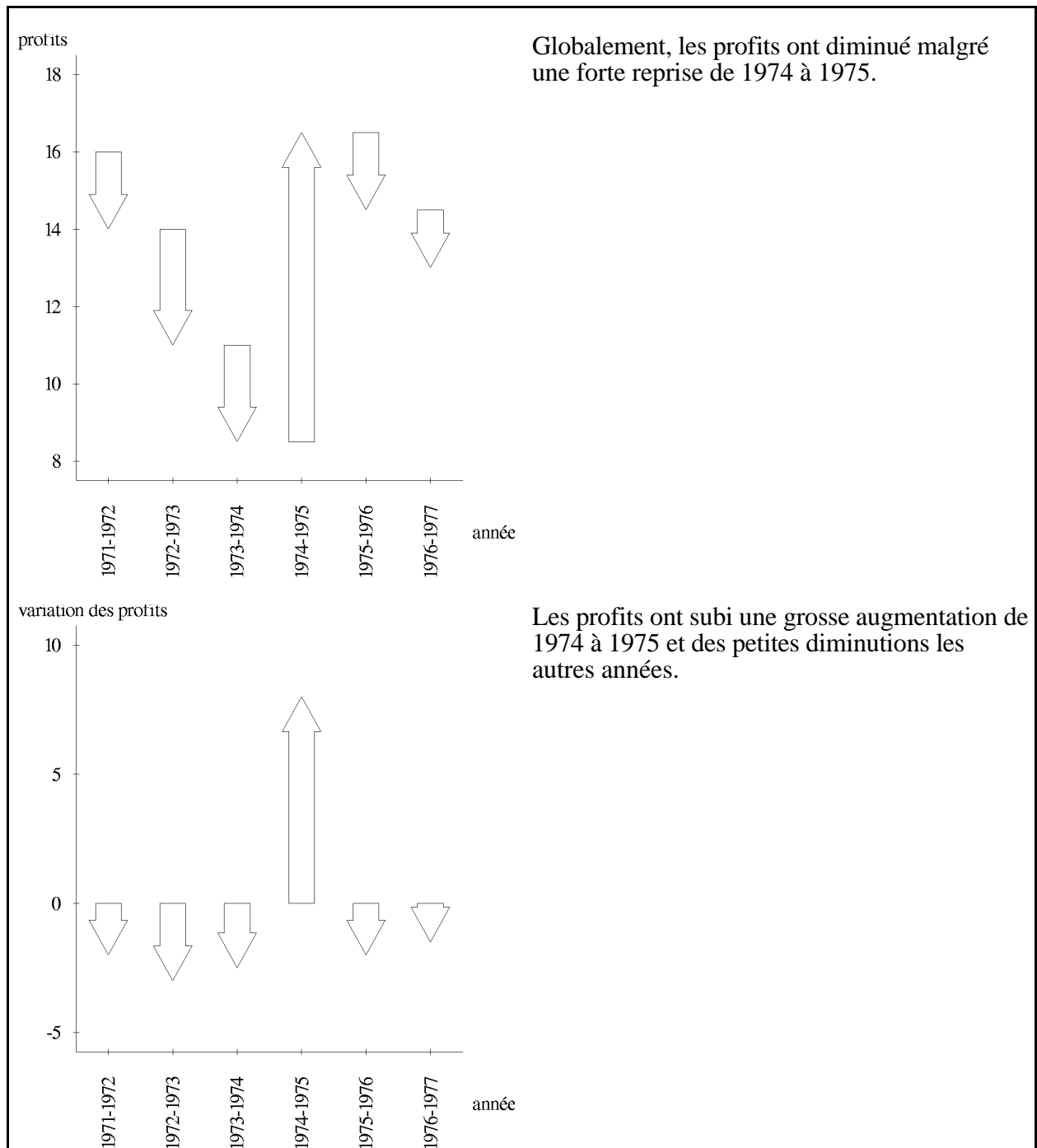


Figure 10: Intégration texte/graphique pour un message subjectif

La partie du haut insiste beaucoup sur les fluctuations et combine texte et graphique de façon conventionnelle. Le message qu'on y présente est honnête tant au niveau du texte que du graphique. Par contre, le mariage du texte et du graphique dans la partie du bas est assez spécial. Le message présenté dans le texte est extrêmement biaisé, à un tel point qu'il ne correspond plus beaucoup aux données. En effet, il insiste sur la grosse augmentation et parle d'une série de petites diminutions, oubliant cependant de mentionner que la somme de ces diminutions dépasse l'augmentation. Le choix d'un graphique pour accompagner ce texte est assez délicat: si on omet le graphique, le lecteur se doute qu'on lui cache des choses et si on en montre trop dans le graphique, le lecteur voit bien que notre texte n'est pas honnête. On a choisi de reprendre le graphique du haut en alignant les variations à 0, ce qui rend beaucoup plus difficile la sommation de toutes les petites diminutions, tout en présentant assez d'information pour ne pas soulever de doutes.

La figure 10 montre donc une intégration texte/graphique hors de l'ordinaire: utiliser le graphique pour rassurer un peu le lecteur tout en essayant de le déjouer en présentant les données de façon à appuyer le mieux possible le texte. On voit donc que les techniques d'intégration conventionnelles ne s'appliquent pas toujours aux cas subjectifs.

#### 4 L'implantation d'un prototype

Nous travaillons sur l'implantation d'un prototype intégrant les résultats théoriques présentés dans cet article. Ce système devra donc être en mesure de produire des graphiques et du texte pertinents et de les combiner.

Le processus de génération d'un rapport texte/graphiques fait intervenir un ensemble de modules différents, chacun capable de fournir son expertise aux autres. Selon nous, il est important d'isoler ces modules et de les rendre disponibles directement à l'utilisateur de notre système, plutôt que de lui fournir une "boîte noire" qui produit un rapport complet. La raison est simple: l'utilisateur peut avoir besoin de l'expertise de notre système pour faire un ou des choix bien particuliers sans nécessairement vouloir un texte complet.

Un premier module s'occupe de choisir les relations intéressantes à exprimer. Il s'agit de la première phase de la planification, le choix du contenu (ex: effet de la hausse des prix sur le taux de chômage).

Ensuite, un module détermine les schémas d'expression les plus appropriés pour une relation. Parmi les schémas d'expression les plus utiles on retrouve: un texte de présentation, un graphique accompagné d'un texte d'analyse, un texte qui relie plusieurs graphiques. C'est la seconde phase de planification, le choix de la structure.

Une fois le schéma identifié, un autre module s'occupe de préciser ses composantes. C'est à ce niveau-ci qu'on détermine quel graphique utiliser (tarte, courbe, etc...) et quelle information inclure dans le texte.

La phase suivante organise l'information à l'intérieur des composantes des schémas. Par exemple, c'est ici que l'on décide si les données seront triées sur les échelles des graphiques. C'est aussi à cette étape que l'ordre des idées du texte est déterminé.

Finalement, le module de réalisation (génération de surface) se charge de dessiner les graphiques (interne à notre système) et de produire le texte (utilisation projetée d'un générateur externe).

Le langage d'implantation que nous avons choisi est Prolog à cause de ses qualités au niveau de la gestion logique de bases de règles et de données relationnelles.

Le prototype travaille avec des données relationnelles en entrée et utilise des annotations pour exprimer les unités utilisées (prix en \$, poids en kg) ainsi que l'intention du rédacteur (s'il en a une) pour chacune des relations.

Le traitement des unités nécessite une certaine connaissance du monde. Celle-ci est encodée dans un graphe d'héritage multiple. À l'aide de ce graphe, dont on peut voir un extrait à la figure 11, nous pouvons organiser les propriétés des unités dans un ensemble de catégories générales (contraintes, type, organisation, etc...). Avec une telle organisation, il devient beaucoup plus simple d'exprimer des conditions dans le système. Par exemple, un choix d'échelle peut dépendre de la propriété "ordonnée" qui englobe par héritage les variables ordinales et quantitatives.

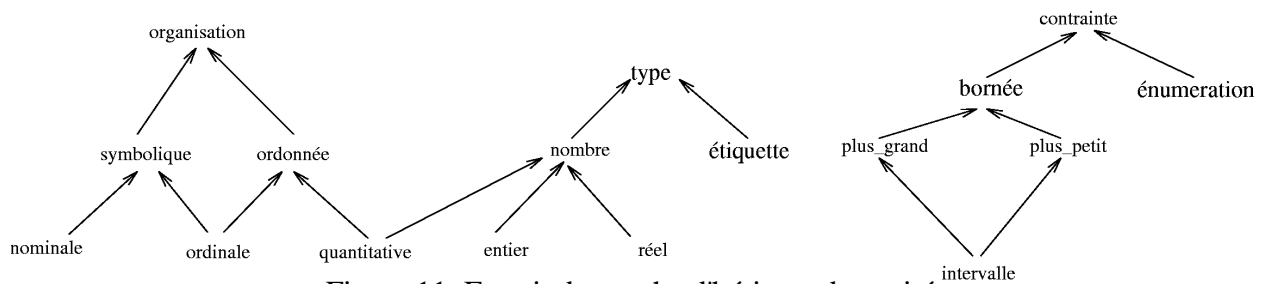


Figure 11: Extrait du graphe d'héritage des unités

De plus, pour pouvoir exprimer les unités de façon plus naturelle, on doit traiter leur décomposition. Cela peut être aussi simple que de dire "3 millions de \$" au lieu de "3 000 000 \$" ou un peu plus subtil, comme par exemple lorsqu'on décide d'approximer 129 jours par 4 mois et 1 semaine.

Présentement, le prototype réalise une partie des modules "experts" décrits auparavant. Les couches de haut niveau traitant le problème de la planification sont encore à l'état embryonnaire, mais les modules inférieurs, en particulier le choix et l'organisation interne des graphiques sont opérationnels et le module de réalisation a produit tous les exemples graphiques de cet article. Le texte n'est pas encore traité; pour cette étape, nous comptons interfacer notre système avec un générateur de texte déjà existant (Gagnon, 1993).

## 5 Conclusion

Les rapports statistiques sont une application intéressante de la génération intégrée de textes et de graphiques. En effet, ils cherchent à présenter une information très dense tout en s'assurant que le lecteur en retient les points essentiels. L'union des deux médias facilite énormément la tâche car les graphiques nous donnent une vue d'ensemble alors que le texte permet de cibler les détails intéressants. Un autre aspect intéressant de ce domaine d'application est le degré de liberté à notre disposition. Puisqu'on part de séries de nombres et que l'on veut produire un rapport cohérent, on a beaucoup de choix à faire. Ceci est autant un avantage (liberté) qu'un problème (complexité).

Les rapports sont très souvent biaisés par le point de vue de la personne qui les rédige. Il faut tenir compte de cet aspect du problème autant à un niveau théorique, pour reproduire le style des rapports qu'on imite, qu'au niveau pratique, pour produire un outil qui sera utilisé par les

rédacteurs. L'étude de l'effet de l'intention du rédacteur sur le format et le contenu du rapport est donc fondamentale.

Le système que nous proposons agira comme un expert capable de conseiller l'utilisateur dans la rédaction d'un rapport. Sa flexibilité lui permettra de produire un rapport complet sans aucune intervention ou de demander au rédacteur de faire des choix subjectifs aux moments opportuns.

## 6 Bibliographie

Bertin J., 1983, *Semiology of Graphics*, The University of Wisconsin Press, traduit par William J. Berg.

Gagnon M, 1993, *Expression de la localisation temporelle dans un générateur de texte*, Thèse de doctorat, Département d'informatique et de recherche opérationnelle, Université de Montréal, Publication 888.

Mackinlay J.D., 1986, *Automatic Design of Graphical Presentations*, Thèse de doctorat, Computer Science Department, Stanford University.

Paris C. L., 1991, The role of the user's domain knowledge in generation, *Computational Intelligence*, 7:71--93.

Tufte E. R., 1983, *The Visual Display of Quantitative Information*, Graphics Press.

Tufte E. R., 1990, *Envisioning Information*, Graphics Press.

Zelazny G., 1989, *Dites-le avec des graphiques*, InterEditions.