

## PLACEMENT AUTOMATIQUE DES TOPONYMES

François LECORDIX, Corinne PLAZANET, François CHIRIE, Mathieu BARRAULT  
Institut Géographique National / France

ملخص: إن كتابة الأسماء الجغرافية على الخرائط أليما لازال قائما، ولم يتوصل الخبراء حتى الآن إلى حله تماما. يقدم هذا المقال آخر ماتوصل إليه في كتابة الأسماء الجغرافية على الخرائط بصفة آلية.

### INTRODUCTION

Parmi les nombreuses applications offertes par les SIG, l'une d'elle concerne la rédaction des cartes papiers traditionnelles par voie numérique. Une phase importante de cette rédaction concerne les écritures : afin d'avoir des résultats de bonne qualité cartographique, elle nécessite du personnel qualifié, et demande un temps non négligeable. En effet dans une base de données qui doit être indépendante de l'échelle de saisie, les toponymes sont rattachés uniquement aux objets géographiques qu'ils renseignent, et donc la position exacte sur la carte ne peut être enregistrée car elle est fonction de l'échelle de rédaction. Ainsi, pour la première carte expérimentale au 1:25 000 issue de la BD Topo, réalisée par l'Institut Géographique National, la feuille de Montpellier, il a fallu 150 heures pour obtenir, en mode manuel, la planche d'écritures comprenant environ 4000 textes.

Pour diminuer ces temps de réalisation de planches d'écritures, une première solution consiste à utiliser un des nombreux logiciels interactifs permettant de positionner des écritures : en simplifiant la manipulation des textes, un tel logiciel permet de réduire d'environ un tiers du temps la réalisation de la planche d'écritures, mais cette méthode nécessite toujours du personnel qualifié et les perspectives d'améliorations futures sont pratiquement nulles car cette solution est limitée par le facteur humain.

Une deuxième solution consiste à automatiser ce placement. Mais il n'existe actuellement sur le marché aucun logiciel permettant de réaliser ce positionnement, sur des cartes à n'importe quelle échelle, avec une qualité cartographique suffisante. En analysant les publications de recherche parues sur ce thème, deux approches nous ont été particulièrement utiles :

\* le système d'Ahn et Freeman (Ahn 1984, Ahn et Freeman 1983) : travaillant sur des canes à petite échelle, ce système cherche à positionner des noms de surfaces (états), de points (villes) et de lignes (grandes rivières). Le principe consiste à placer les toponymes dans l'ordre des "degrés de liberté" croissants, c'est-à-dire de placer d'abord les noms qui sont en conflit avec le plus grand nombre d'autres noms et de terminer par ceux qui en ont le moins. Pour les noms de points, on crée un graphe où les nœuds représentent les détails ponctuels à nommer de la carte. On définit un rectangle associé à chacun de ces nœuds par la réunion

des emprises des positions possibles du nom de ce nœud (Figure 1). Deux nœuds du graphe sont reliés si les rectangles se chevauchent. On décompose ensuite

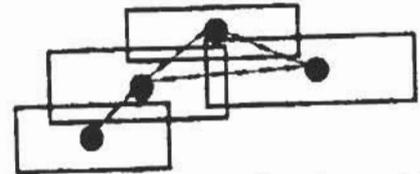


Figure 1 : Construction du graphe de superposition

le graphe en composantes connexes, que l'on traite séparément. Pour positionner le nom d'un nœud d'une composante, on travaille de manière séquentielle heuristique : on traite le nœud qui a le degré de liberté le plus faible et on choisit la position qui élimine le

moins de positions possibles pour les nœuds voisins. Pour certains blocages, il est possible de revenir en arrière sur certains choix.

\* la méthode de l'optimisation mathématique proposée par Cromley et Zoraster (Cromley 1985 et 1986, Zoraster 1986 et 1987) : le principe consiste à résoudre le problème précédent, non pas par une solution heuristique telle que l'ont proposée Ahn et Freeman avec les degrés de liberté, mais par une méthode mathématique d'optimisation. L'expression mathématique du problème se pose sous la forme :

soit  $i$  indice des  $n$  nœuds auxquels il faut affecter un nom,

$j$  indice des  $m$  positions possibles pour un nœud  $i$ ,

$X_{ij}$  position possible  $j$  pour un nom  $i$ , valeur valant 0 ou 1,

$P_{ij}$  sanction de la position possible  $j$  pour le nom  $i$ ,

il faut minimiser  $\sum_i \sum_j P_{ij} X_{ij}$

Il faut aussi respecter les contraintes de dénomination de tout point et de non chevauchement de toutes positions :

\*  $\forall i \quad \sum_j X_{ij} = 1$

\* si la position  $j$  du point  $i$  chevauche la position  $j'$  du point  $i'$  :  $X_{ij} + X_{i'j'} < 1$

Mais la limitation principale de ces recherches résulte

d'une prise en compte partielle, ou même inexistante, du fond cartographique qui s'avère pourtant un facteur déterminant du positionnement des écritures réalisé manuellement par un cartographe. D'autre part, comme l'ont souligné Doerschler et Freeman (Doerschler et Freeman 1992), la méthode de l'optimisation mathématique, si elle est intellectuellement satisfaisante, peut se révéler non économique et donc difficile à appliquer pour une carte complète qui représente un volume de données important. Nous avons donc recherché à automatiser ce placement en veillant à deux problèmes :

- \* prise en compte du fond cartographique pour une bonne qualité esthétique de placement.
- \* possibilité de traiter une carte complète qui peut contenir jusqu'à 5 000 écritures.

### DEMARCHE RETENUE

Afin de mettre en place les premiers principes de notre recherche, nous avons choisi de ne traiter, dans un premier temps, que les écritures horizontales attachées à des objets ponctuels ou surfaciques : noms de villes, communes et lieux-dits, sites touristiques, détails planimétriques ou orographiques, points cotés... Ces écritures représentent, en nombre, plus de la moitié des textes d'une carte et parfois même jusqu'à 80%. Les autres types d'écritures seront abordés ultérieurement dans nos recherches au vu des premiers résultats acquis.

Pour définir une bonne modélisation du problème de positionnement de ces écritures, rappelons d'abord les principales règles générales qui régissent le travail du cartographe et qu'il est indispensable de prendre en compte pour obtenir un résultat de qualité cartographique. Elles ont été analysées par Alinhac et Imhof (Alinhac 1964, Imhof 1975) et peuvent être classifiées en deux catégories :

\* Règles s'appliquant au nom pris isolément :

- R1) Les noms désignent les objets sans ambiguïté, donc il doit être plus proche de son objet que des autres objets.
- R2) Les noms doivent gêner le moins possible le reste du contenu de la carte et même éviter totalement certains détails (symboles ponctuels, carrefours et sinuosités de routes...)
- R3) En cas de choix après les deux règles précédentes, on choisit l'emplacement selon l'ordre de préférence de la figure 2: précisons que ce schéma est à comprendre en terme de zone préférentielle, et non comme une limitation à uniquement six positions strictes.

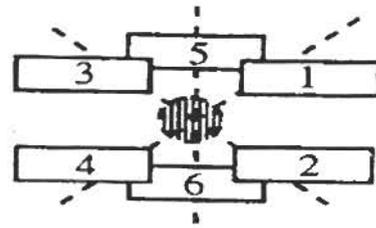


Figure 2 : Positions privilégiées

- R4) Les noms composés de plusieurs mots et qui désignent des lieux habités peuvent être disposés sur deux lignes.

- R5) En cas de présence d'une ligne planimétrique importante, le nom et l'objet à renseigner sont du même côté par rapport à cette ligne.

\* Règles s'appliquant aux noms entre eux :

- R6) Deux noms ne doivent pas se chevaucher ou se situer trop près l'un de l'autre.

- R7) Deux noms ne doivent pas former une figure géométrique (alignement).

Précisons que toutes ces règles sont très manichéennes, mais le cartographe réalise souvent un compromis entre ces différentes contraintes qui, dans la pratique, s'avèrent souvent contradictoires.

Afin d'obtenir un placement soigné, nous testons donc un nombre important de positions pour chaque objet à renseigner. Pour chacune de ces positions, nous procédons à une quantification de sa qualité en fonction des règles cartographiques R1 à R5 rappelées précédemment qui sont combinées suivant les souhaits du cartographe (par exemple privilégier la non mutilation du fond cartographique R2 ou privilégier l'ordre de préférence R3). Cette quantification permet d'affecter un coût à chaque position, appelé poids propre. Celui-ci traduit la qualité intrinsèque d'une position indépendamment des autres noms. Nous retenons alors, pour chaque nom, les  $m$  meilleures positions possibles, c'est-à-dire celles qui ont des valeurs de poids propre les plus faibles. Le facteur  $m$  est un paramètre du programme que nous fixons généralement à la valeur 30 qui se révèle être un bon compromis entre les problèmes de qualité nécessitant un choix très large de positions et les problèmes de volume de données informatiques imposant une limitation du nombre de ces positions. Il reste alors à sélectionner parmi ces  $m$  meilleures positions, celle à retenir en tenant compte des autres noms et des règles R6 à R7. Cette sélection peut se faire suivant différentes stratégies de placement : des méthodes heuristiques séquentielles voisines de celle de Ahn et Freeman ou des méthodes globales par optimisation mathématique.

## DESCRIPTON DES DIFFERENTES PHASES DE PLACEMENT

L'algorithme se décompose en trois étapes principales :

### 1) Initialisation des données

Cette première étape permet de déterminer les informations nécessaires pour la suite des traitements. Elle sert d'interface entre les informations de la base de données et les caractéristiques de la carte que l'on souhaite réaliser. Outre le chargement des données initiales (graphe et légende des objets à nommer et de la carte, paramètres géométriques de positionnement fixés par le cartographe), elle réalise le calcul d'informations complémentaires :

- \* abréviations des noms suivant les règles toponymiques de la carte (exemple : abréviation de Saint en St);

- \* détermination de la césure éventuelle en deux lignes (respect de R4) : impossible pour des toponymes très courts, obligatoire pour des toponymes trop longs, possibles entre les deux options précédentes;

- \* calcul du rectangle d'emprise des noms en fonction de la police et du corps d'écriture utilisés sur la carte.

Précisons que dans la suite des traitements géométriques, pour travailler sur un nom, nous utilisons uniquement son rectangle d'emprise;

- \* calcul de la zone d'influence associée à chaque objet suivant la figure 3 pour les objets surfaciques. Cette zone d'influence permet d'optimiser les traitements;

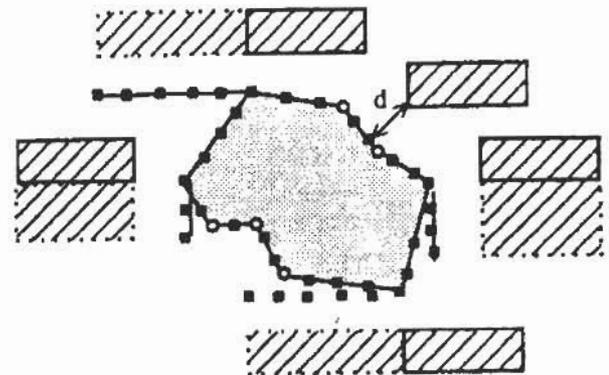
- \* détermination des interactions entre objets traduisant le chevauchement des zones d'influence

### 2) Détermination et quantification de positions pour chaque nom indépendamment des autres

Suivant le type d'objets à nommer, ponctuel ou surfacique, et son mode de positionnement souhaité, différentes positions sont déterminées :

\* **Objet surfacique :**

- **Positionnement extérieur :** à partir de la liste des points délimitant la surface on détermine les quatre points cardinaux; puis on détermine une nouvelle liste constituée de points répartis à équidistance sur la limite de la surface (la distance entre chaque point étant un paramètre du programme que l'on fixe généralement à 0.5 mm, ce qui correspond sensiblement à l'écart entre deux positions que teste visuellement le cartographe) et des points choisis sur les segments correspondant à la largeur et à la hauteur du nom à partir des quatre points cardinaux (figure 4). Les positions testées sont issues de ces points décalés vers l'extérieur d'une certaine valeur  $d$  qui est un paramètre du programme, pour placer les écritures en dehors de la surface



○ Points décrivant la surface    ■ Points équidistants testés

Figure 4 : Points pour un objet surfacique avec un positionnement extérieur

- **Positionnement mordant :** c'est le même principe que précédemment mais le décalage final se fait vers l'intérieur de la surface.

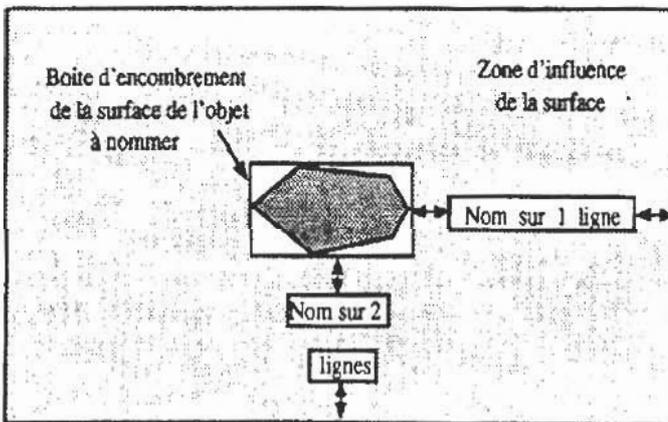


Figure 3 : Définition de la zone d'influence d'un objet

- Positionnement intérieur sur une maille : dans le cas où l'emprise du nom est plus petite que la boîte d'encombrement de la surface, les positions testées se trouvent sur les points d'intersection d'une maille dont le pas est un paramètre du programme. La taille de la maille est égale à la boîte d'encombrement de la surface moins le rectangle d'emprise du nom (Figure 5).

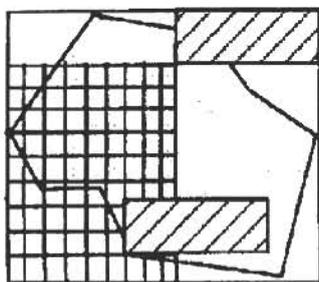


Figure 5 : Points pour un objet surfacique avec un positionnement intérieur sur maille

- Positionnement intérieur autour d'un centre : c'est le même principe de maille que précédemment, mais la maille est carrée et centrée sur le barycentre de la surface, décalé de la demie emprise du nom (figure 6).

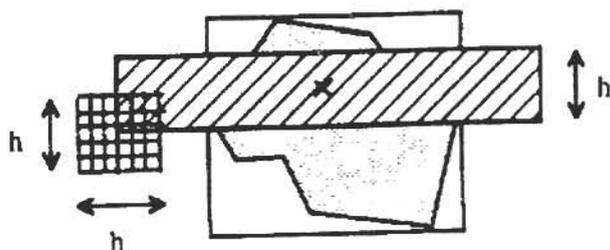


Figure 6 : Points pour un objet surfacique avec un positionnement intérieur autour d'un centre

\* Objet ponctuel :

- Positionnement attaché à un point d'accroche sans symbolisation : certains noms des bases de données IGN n'étant rattachés à aucun objet cartographique, il est nécessaire d'avoir la possibilité, de placer ce nom autour de ce point d'accroche de manière similaire au cas précédent (le barycentre est remplacé par le point d'accroche).

- Positionnement sur un symbole ponctuel circulaire : les positions testées sont réparties sur un cercle de rayon celui du symbole ponctuel plus un léger décalage de façon que les angles au centre les positions soient équipollents (la valeur du décalage et du nombre de partitionnement du cercle sont des paramètres du programme) (figure 7).

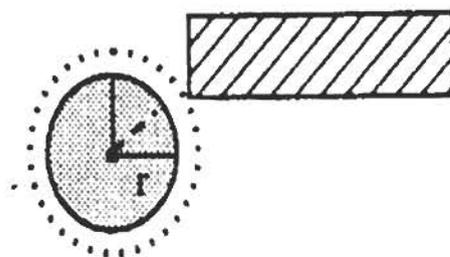


Figure 7 : Points pour un objet ponctuel

Une fois ces positions déterminées pour les différents types d'objets à nommer, le logiciel commence par réaliser un premier filtrage de certaines positions en fonction de la règle cartographique R1 (Principe de non ambiguïté). Cette élimination se fait par un calcul de distance : si la distance entre l'objet à renseigner et son nom est supérieure à la distance entre ce nom et le contour d'un objet voisin, on élimine cette position.

Ensuite, on détermine pour chaque position la quantification de sa qualité en calculant son poids propre. Ce poids propre est une combinaison linéaire des poids de mutilation, de direction et de distance :

\* Poids de mutilation (modélisation de la règle R2) : ce poids permet de prendre en compte le fond d'image de la carte et donc de déterminer les positions pour lesquelles le rectangle d'emprise du nom écrase le moins d'informations importantes de la carte. Le calcul se fait en mode maillé avec une résolution de 8 points par millimètre afin d'avoir une bonne finesse de maillage améliorant cette modélisation. La carte est rédigée avec une légende spécifique, identique à celle de la carte pour les symboles, mais les couleurs sont remplacées par des poids compris entre 0 et 9, et même 10 qui sont donc les valeurs des pixels de la carte en mode maillé. Une valeur de 0 correspond à des pixels que l'on peut écraser sans problème, une valeur de 9 à des pixels qu'il est très préjudiciable de masquer par le nom et la valeur de 10 correspond à des pixels qu'il est interdit de masquer (par exemple un détail ponctuel de la carte). On a alors le poids de mutilation d'une position :

Poids de mutilation (Position) =  $(\sum \text{pixels recouverts valeur des pixels}) / \text{nombre de pixels recouverts}$  et la position est interdite si la valeur d'un des pixels recouverts est égale à 10.

\* Poids de distance (modélisation des règles R1 et R5) : ce poids vise d'une part à rapprocher le nom du centre de l'objet, particulièrement pour une surface. Cette valeur est définie comme étant dix fois le rapport de la distance de la position au centre de l'objet sur la distance la plus élevée entre toutes

les positions possibles pour le nom et le centre de l'objet. D'autre part, ce poids est ajouté d'une valeur supplémentaire paramétrable si la position est, par rapport au centre de l'objet, de l'autre côté d'un élément linéaire significatif de la carte.

\* Poids de direction (modélisation de R3): suivant la zone où se trouve la position dans l'une des 8 orientations possible (nord, nord-est, est ...), le poids de direction est affecté d'une valeur comprise entre 0 et 10; ces valeurs sont des paramètres du programme.

Ces trois poids étant normalisés entre 0 et 10, les facteurs de combinaison linéaire doivent avoir leur somme égale à 1 et peuvent être fixés selon les préférences du cartographe : pour privilégier l'un des poids, il suffit d'augmenter la valeur du coefficient de ce poids et diminuer celles des deux autres.

Enfin, toutes les positions étant quantifiées ou éliminées, par un simple algorithme de tri on ne retient que les  $m$  positions ayant le poids propre le plus faible. Précisons qu'il faut choisir un compromis judicieux entre les paramètres du nombre maximum de positions retenues et du pas entre chaque position possible afin de ne pas retenir que des positions qui seront toutes éliminées par le placement d'un autre nom.

### 3) Stratégies de placement

Il faut maintenant retenir la position idéale en respectant la règle R6. Pour cela, on détermine d'abord les incompatibilités entre positions qui résultent, soit des recouvrements ou des trop grandes proximités d'une position écrite avec une position d'une autre écriture, soit d'une ambiguïté issue d'une position plus proche d'un autre objet qu'une position pour cet objet (Cf. Figure 8). Ce calcul est optimisé par l'utilisation des zones d'influence des objets. Pour continuer, ce dispose alors de différentes solutions :

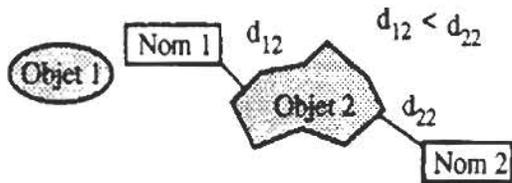


Figure 8 : Incompatibilités pour cause d'ambiguïtés

\* Méthodes de placement séquentiel : le but étant de placer le maximum de noms, il est nécessaire de choisir un bon ordre de placement et de sélectionner la meilleure position en tenant compte non seulement du poids propre, mais aussi des interactions. Pour l'ordre de placement, deux solutions ont été programmées :

- Ordre fixe suivant les degrés de contrainte décroissants : l'algorithme calcule d'abord pour chaque position son degré de contrainte comme étant le nombre de positions incompatibles avec celle-ci, puis le degré de

contrainte de chaque objet comme étant la moyenne, sur la liste des positions possibles de l'objet, des degrés de contrainte de chaque position de l'objet. En fait, comme les positions de bonne qualité, importent davantage que celles de mauvaise qualité, cette moyenne est pondérée par l'inverse du rang de la position dans la liste des positions possibles qui ont été classées suivant l'ordre croissant des poids propres. Une fois ces objets à nommer triés suivant le degré de contrainte décroissant, on choisit séquentiellement la position de chacun. Pour résoudre les blocages qui peuvent se produire, un retour en arrière de un niveau a été programmé.

- Ordre variable suivant les degrés de liberté croissants : la solution précédente étant trop rigide car le degré de contrainte varie en fait avec les choix réalisés au cours du placement, une deuxième approche consiste à recalculer, après le choix d'une position, le nouveau degré de liberté de chaque objet restant à nommer, c'est-à-dire le nombre de positions encore disponibles pour chaque objet.

Pour diminuer encore les probabilités de blocage qui sont favorisées par les méthodes séquentielles, le choix d'une position peut être fait non seulement en fonction du poids propre défini précédemment, mais en fonction aussi d'un poids d'interactions. Celui-ci traduit dans quelle mesure le choix de telle position générerait le placement des noms à venir. Cette valeur peut être définie comme dix fois le rapport du nombre de positions en interaction divisé par le maximum, sur l'ensemble des positions d'un objet, du nombre de positions en interaction. On définit alors le poids total d'une position comme une combinaison linéaire du poids propre et du poids d'interactions, les deux coefficients de la combinaison linéaire étant des paramètres du programme. Lors du placement séquentiel, on sélectionne alors la position qui a le poids total le plus faible. Précisons que plus le coefficient du poids d'interaction est fort, plus le nombre de blocage diminue, mais plus la qualité intrinsèque du placement risque d'être dégradée par diminution de l'influence du poids propre.

\* Méthodes par optimisation globale : pour résoudre le problème du choix de la meilleure position, en respectant les contraintes d'incompatibilité entre positions, une approche plus mathématique telle que l'ont proposée Cromley et Zoraster a été étudiée. Dans un premier temps, l'IGN a confié à la société Infosys la résolution du problème par des méthodes alternatives. L'étude s'est portée sur un extrait comprenant 340 noms à positionner, avec pour chacun au maximum une trentaine de positions possibles dont le poids propre est déjà calculé. Afin de diviser le problème en sous problèmes indépendants, on construit le

graphe des incompatibilités : deux noms sont connectés s'ils sont en interaction entre eux ou s'il existe un chemin de noms en interaction les reliant. Ce graphe est alors décomposé en composantes connexes ( sur l'essai fourni, il y en a 86). Rappelons que le problème à résoudre se pose sous la forme énoncée au début :

soit  $i$  indice des  $n$  nœuds auxquels il faut affecter un nom,

$j$  indice des  $m$  positions possibles pour un nœud  $i$ ,

$X_{ij}$  position possible  $j$  pour un nom  $i$ , valeur valant 0 ou 1,

$P_{ij}$  sanction de la position possible  $j$  pour le nom  $i$ ,

il faut minimiser  $\sum_i \sum_j P_{ij} X_{ij}$

avec : \*  $\forall i \quad \sum_j X_{ij} = 1$

\* si la position  $j$  du point  $i$  chevauche la position  $j'$  du point  $i'$  :  $X_{ij} + X_{i'j'} < 1$

- Résolution par recuit simulé : cette technique, qui s'est popularisée au cours des années 80, s'inspire d'une analogie avec la thermodynamique statistique: un système physique laissé à lui-même évolue vers un état d'énergie utile minimale, si sa température décroît lentement. Ce principe est utilisé, en métallurgie : un bloc de métal chauffé puis lentement refroidi tendra vers un état d'énergie minimale - et donc de résistance maximale. En fait, la température élevée du métal autorise des transitions moléculaires qui peuvent temporairement accroître localement l'énergie utile d'un fragment; la convergence vers une énergie minimale est statistique. Si l'on s'interdit ce type de transitions (ce qui est le cas dans la technique de trempe, en refroidissant brutalement l'échantillon), on observe au contraire que le métal reste "bloqué" à un niveau d'énergie plus élevé (ce qui accroît sa dureté.). Le même principe peut être appliqué en optimisation de grands systèmes, où le coût à minimiser est assimilé à une énergie et où on peut passer d'une solution à une autre par des transitions simples.

Le recuit simulé va effectuer un grand nombre de transitions dans un ordre aléatoire. Il pourrait sembler intuitif de ne choisir que des transitions qui diminuent le coût global (ce qui serait une trempe simulée, appelée descente de gradient en optimisation): en fait, la technique du recuit simulé consiste justement à admettre aussi des transitions défavorables et ceci sous le contrôle de la température. Ce n'est que lorsque la température atteindra zéro que les seules transitions favorables seront acceptées.

Modéliser en recuit simulé, c'est définir : les variables du système, le coût d'un état, les transitions d'un état à l'autre, la décroissance en température. Appliquée au placement des écritures, la métaphore se décline ainsi :

\* les variables du système sont les positions de chaque nom de ville:

\* le coût d'un état est la somme des poids de chacun des placements indiqués par les valeurs des variables.

Nous ajoutons à ce coût un surcoût très important lorsque deux placements produisent un chevauchement ou une ambiguïté:

\* la transition d'un état à l'autre se produit lorsqu'une des variables (position d'une ville) change de valeur,

\* la décroissance en température est exponentielle

On remarque que les incompatibilités ne sont pas systématiquement levées. Nous leur donnons simplement un coût très élevé, de 1000, à comparer au coût compris entre 0 et 10 de toutes les autres règles confondues. L'espoir - car le recuit simulé ne donne que des garanties statistiques d'optimum - est que le système finisse par atteindre un état de moindre coût en un temps raisonnable. En particulier, il faut que les incompatibilités aient toutes été résorbées. En pratique, cela se produit toujours pour l'essai étudié.

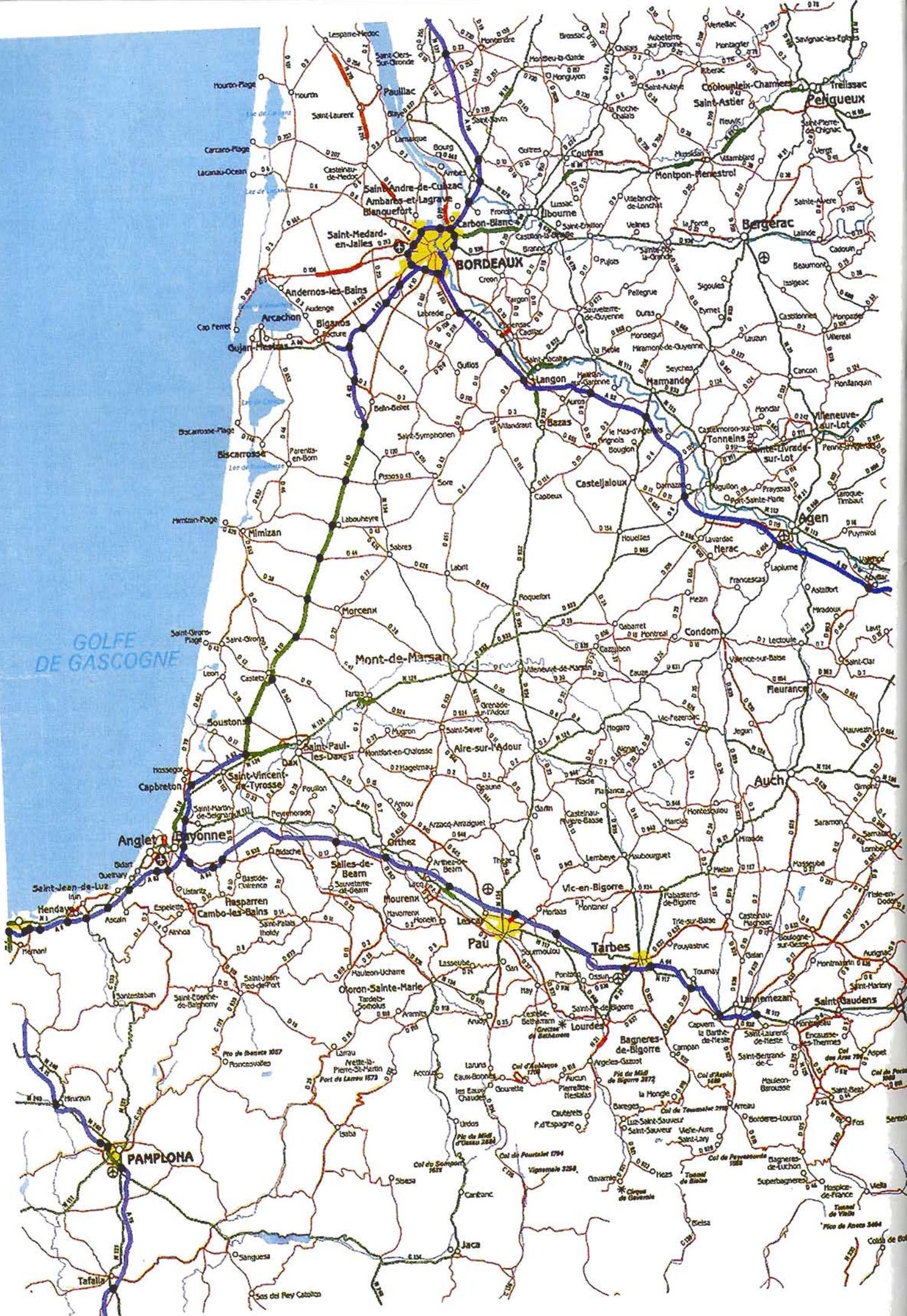
Avec cette algorithmique simple, nous obtenons sur l'essai de très bons résultats. Toutes les incompatibilités sont en général résorbées. Il est remarquable de constater la grande stabilité de l'algorithme par rapport au paramétrage du refroidissement.

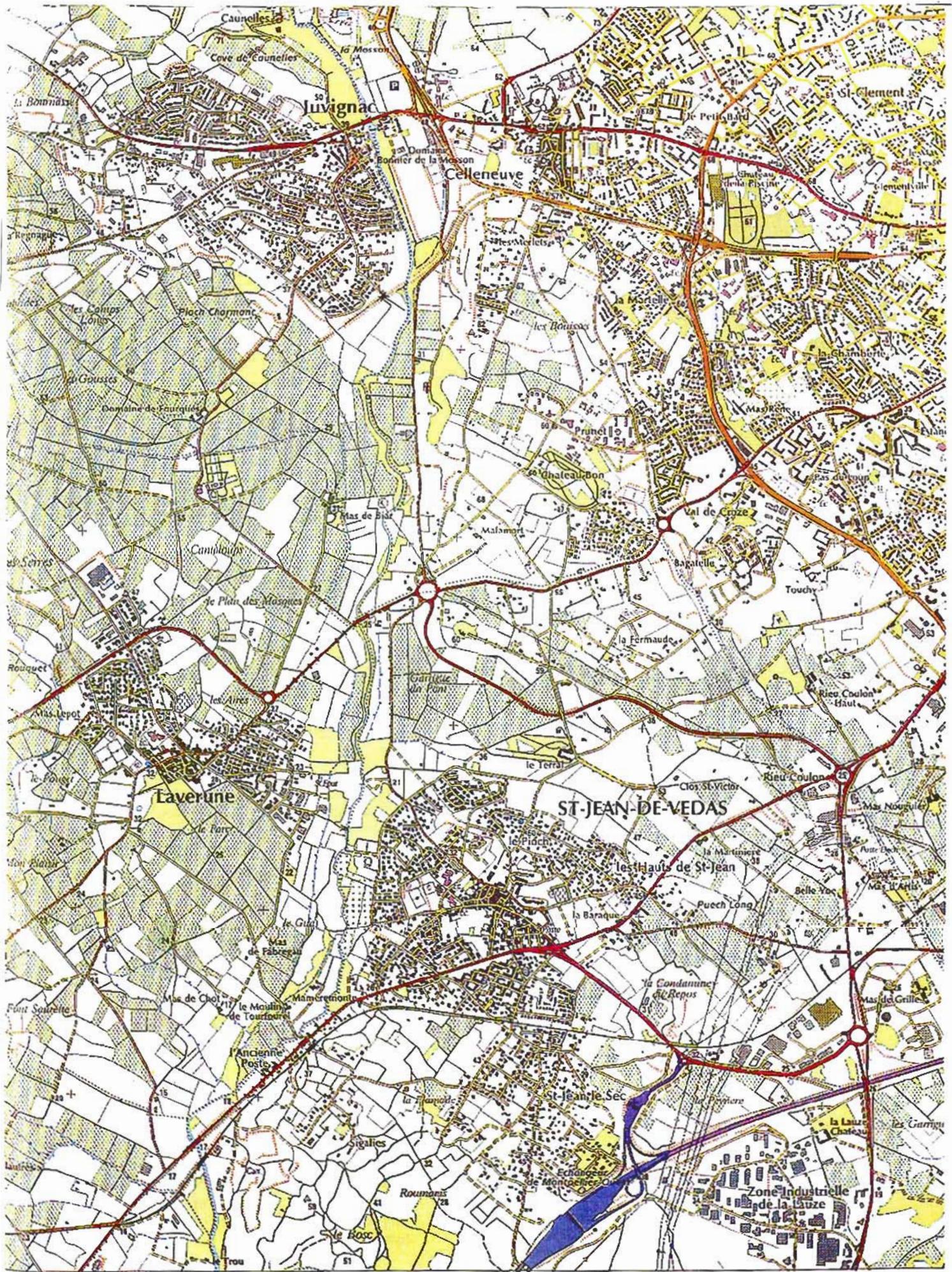
## EXPERIMENTATIONS

Le premier essai a porté sur une ébauche de carte au 1:100 000 pour des noms de commune (Ch—i, F. 1992). Aux vues des premiers résultats encourageants, les principes mis en œuvre ont été améliorés en 1993 et des tests sur des cartes réelles au 1:1 000 000 et au 1:25 000 ont été réalisés avec une estimation par un cartographe traditionnel du pourcentage de bon positionnement. La rédaction de la carte et de l'image des mutilations est faite avec le logiciel Mercator de Barco Graphics.

### 1) Essai au 1:1 000 000 sur une carte routière

Cet essai, assez semblable à ceux réalisés par Ahn et Frecman sur leur système AUTONAP, travaille principalement sur des noms de ville représentés en mode ponctuel. L'intérêt de cette échelle et de cette carte réside dans la forte densité toponymique qui permet alors de tester les stratégies de placement. Afin de se rendre compte de la qualité du résultat un extrait de la sortie cartographique est fourni en page suivante. Pour les 338 toponymes de la zone test, par une méthode séquentielle, 8 noms ne sont pas placés; grâce à l'utilisation de la méthode de recuit, outre l'amélioration globale du positionnement, 2 blocages entre noms sont résolus. Le temps d'exécution complet du programme est de 2





PLACEMENT AUTOMATIQUE DES ECRITURES SUR UN EXTRAIT DE LA CARTE DE MONTPELLIER AU 1/25000

minutes CPU sur une station DEC Alpha. Le pourcentage de bon positionnement validé par le cartographe est de l'ordre de 80%. Certaines erreurs peuvent être résolues aisément en améliorant la phase d'initialisation ( règles de césure par exemple).

Sur la sortie présentée, les travaux récents de placement automatique des numéros de routes et des kilométrages peuvent être aussi examinés. Les principes utilisés pour ces écritures sont très similaires à ceux utilisés pour les toponymes. Seules les écritures hydrographiques proviennent d'un placement interactif.

## 2) Essai au 1:25 000 sur une carte topographique

Si la densité toponymique sur ce type de carte est beaucoup plus faible que sur l'exemple précédent, cet essai s'avère cependant très intéressant à différents titres : il existe peu de travaux de placement automatique réalisés à cette échelle car la prise en compte du fond cartographique est primordiale pour un bon positionnement, les types d'objet à renseigner sont très variés (lieux-dits habités et non habités, bâtiments remarquables, points cotés ...), le nombre de cartes à réaliser à partir de la BD Topo est important (plus de 2000) et donc les gains de productivité seront d'autant plus considérables.

Sur ces données issues de la BD Topo, il existe une difficulté supplémentaire du fait de la structuration de la base: en effet, les toponymes de lieux-dits habités ne sont rattachés qu'à un point situé " au centre" des maisons renseignées: il est donc nécessaire de créer automatiquement une enveloppe des bâtiments, soit en mode vecteur suivant le schéma de la figure 12 lorsqu'il s'agit de quelques maisons, soit en mode maillé, en utilisant des procédés de généralisation (réduction d'échelle, érosion, dilatation, détournage et vectorisation) pour les villages et les villes qui sont constitués par un ensemble de maisons.

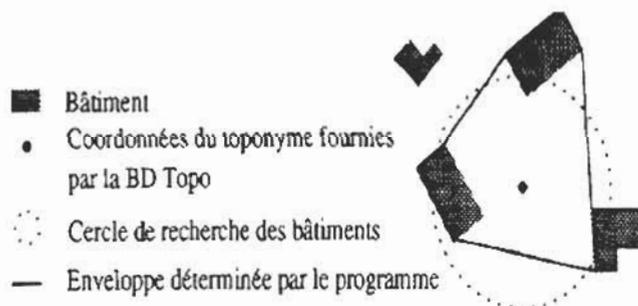


Figure 12 : Construction d'une enveloppe de bâtiment

Un premier essai a été réalisé sur un extrait de la carte de Montpellier, dont la sortie graphique est fournie en page précédente. Le résultat chiffré de positionnement bon ou acceptable pour les 252 écritures de l'extrait est de 90% (avec un pourcentage de 95% pour les points cotés qui représentent 50% des écritures horizontales d'une carte au 1:25 000, et de 86% pour les lieux-dits). Un deuxième essai en vraie grandeur a été

mené sur la TOP 25 du Ventoux, comprenant 2950 points et 1480 toponymes et désignations. Pour réaliser le positionnement de toutes les écritures de la carte, les traitements se sont faits en plusieurs étapes :

- \* placement interactif des écritures à disposition de surface (forêts...) en 16 heures;
- \* placement automatique des écritures horizontales sur une station DEC Alpha en moins d'une heure;
- \* corrections interactives de 10% de ces écritures en 12 heures;
- \* placement interactif des noms à disposition sur des objets linéaires (routes et hydrographie) en 12 heures.

## CONCLUSIONS ET PERSPECTIVES

Grâce à la quantification de la qualité du positionnement prenant en compte le fond cartographique, les recherches concernant le placement automatique des écritures horizontales fournissent des placements satisfaisants pour 80 à 90% d'entre elles. Ce pourcentage élevé rend possible l'utilisation de ces traitements en production, les 10 à 20% restants étant corrigés interactivement. L'industrialisation de ces recherches est en cours. En réutilisant certains principes de la quantification de la qualité des positions, les recherches sur les noms de routes ont abouti à des résultats similaires: 90% de positionnement satisfaisant pour un extrait de carte au 1:100 000. Les recherches sont fortement avancées pour d'autres types d'écritures : noms de rues, kilométrages, noms d'hydrographie. Tous ces travaux ne sont spécifiques ni à l'échelle de la carte, ni au pays cartographié. Ils peuvent donc être utilisés pour la cartographie numérique africaine.

Enfin, il faut observer que pour obtenir ces résultats, il a été nécessaire de concilier de nombreuses techniques qui sont parfois opposées utilisation simultanée du mode vecteur et du mode raster, prise en compte des contraintes cartographiques et informatiques.... C'est seulement à ce prix que les SIG permettront de rédiger des cartes avec une qualité cartographique digne des cartographes traditionnels.

## REFERENCES :

- Ahn J. (1984). Automatic Map Name Placement System. Image Processing Laboratory Technical Report 063. Electrical, Computer, and Systems Engineer Department, Rensselaer Polytechnic Instit., Troy, New York.

**Ahn J et Freeman H.** (1983), A Program for Automatic Name Placement. Proceedings, Auto-Carto VI, Vol. 2, p. 444-453.

**Alinhac G.** (1964) *Ecritures. Cartographie théorique et technique*, fascicule 1, chap IV.IGN Paris.

**Chirie F.** (1992). Programme de positionnement automatique des noms de communes. COGIT, IGN. Paris.

**Cromley R.C.** (1985). An LP Relaxation Procedure for Annotating Point Features Using Interactive Graphics. proceedings. AUTO-CARTO VII, p. 127-132.

**Cromley R.C.** (1986). A Spatial Location Analysis of the Point Annotation Problem. Proceedings, Second International Symposium on Spatial Data Handling, p. 38-49, Seattle, Washington.

**Doerschler J.S et Freeman H.**(1992) An Rule-Based System for Dense-Map Name Placement. Communications of the ACM, vol. 35, n.1.

**Imhof E.** (1975) Positionning Names on Maps. the American Cartographer, Vol.9, n.1, p.5-17.

**Zoraster S.** (1986), Integer programming applied to the map label placement problem. Cartographica Vol. 23. n. 3, p16-27.

**Zoraster S.** (1987). Practical experience with a map label placement program. Proceedings. Auto-Carto VIII. p.701-707