

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

**A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600**



Université d'Ottawa • University of Ottawa



UNIVERSITÉ D'OTTAWA

**Analyse de réseaux et intégration d'une approche déductive
aux systèmes d'information géographique.**

L'exemple du transport routier des matières dangereuses

Par

DENIS G. LEROUX
(B. Sc. Géographie, M. Sc. Géographie)

**THÈSE
PRÉSENTÉE À L'ÉCOLE DES ÉTUDES SUPÉRIEURES ET DE LA RECHERCHE**

**COMME EXIGENCE PARTIELLE
EN VUE DE L'OBTENTION DU GRADE DE
PHILOSOPHIAE DOCTOR (Ph. D.)**

**Département de géographie
Université d'Ottawa
Ottawa, Ontario
Canada**



**National Library
of Canada**

**Acquisitions and
Bibliographic Services**

**395 Wellington Street
Ottawa ON K1A 0N4
Canada**

**Bibliothèque nationale
du Canada**

**Acquisitions et
services bibliographiques**

**395, rue Wellington
Ottawa ON K1A 0N4
Canada**

Your file / Votre référence

Our file / Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced with the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-21005-7

Résumé

La présente thèse aborde le thème de la géographie des transports et plus spécifiquement celui de la recherche d'itinéraires sécuritaires pour le transport des matières dangereuses.

La très vaste majorité des algorithmes de recherche d'itinéraires optimaux ont recours à l'impédance comme variable d'optimisation. Cette approche est limitative, particulièrement dans le domaine du transport des matières dangereuses où un grand nombre de facteurs quantitatifs et qualitatifs doivent être pris en compte pour déterminer un itinéraire.

Cette recherche démontre que la théorie de la décision et les systèmes experts sont deux outils en mesure de rendre plus flexibles les modèles de recherche conventionnels en leur permettant de traiter les données qualitatives et quantitatives nécessaires à la sélection d'un itinéraire.

Un modèle de recherche d'itinéraire basé sur le couplage de l'algorithme de Dijkstra à un système expert est présenté. Ce modèle, intégré à un système d'information géographique, forme un système de support à la décision à référence spatiale dans le domaine du transport des matières dangereuses. La formalisation du modèle de recherche d'itinéraire et l'élaboration du système de support à la décision sont illustrées à l'aide d'un exemple sur la région d'Ottawa-Hull.

Abstract

The main topic of this thesis is transportation geography and more specifically the problem of safe path finding for the transportation of dangerous goods.

Most of the path finding algorithms use impedance as a unique variable for optimisation. This approach is limited, especially in the domain of dangerous goods, where quantitative and qualitative variables must be considered when selecting a path.

This research shows that decision theory and expert systems are among the best tools to make conventional path finding algorithms more flexible and able to deal with quantitative and qualitative data.

A path finding model based on the coupling of the Dijkstra algorithm and an expert system has been developed in this thesis. The integration of this model to a geographic information system creates a spatial decision support system for the transportation of dangerous goods. Formalisation of the safe path finding model and development of the decision support system are demonstrated by an example of the Ottawa-Hull area.

TABLE DES MATIÈRES

Résumé	i
Abstract	ii
.....	
Remerciements	iii
Table des matières	iv
Liste des figures	viii
Liste des tableaux	xi

Chapitre I

1.0 Introduction	1
1.1 Le sujet de la thèse	6
1.2 Justification du sujet	8
1.3 Objectifs de la thèse	9
1.4 Les limites de la thèse	10
1.5 Structure de la thèse	12

Chapitre II

2.0 Problématique	14
2.1 Géographie des transports	15
2.2 Déplacements dans les réseaux	19
2.2.1 Le plus court chemin	20
2.2.2 L'impédance	22
2.3 Les limitations de l'impédance	23
2.3.1 L'impédance et le transport des matières dangereuses	26
2.3.2 Les modèles d'itinéraires à objectif unique	28
2.3.3 Les modèles probabilistes à évaluation des risques	29
2.3.4 Modèles à objectifs multiples	33
2.3.5 Modèles en temps réel	36
2.4 Élément de solution: Les systèmes experts	40

Chapitre III

3.0 La théorie des graphes	41
3.1 Théorie des graphes et géographie	42

3.2 Les graphes en tant que modèle spatial	45
3.2.1 Éléments de base d'un graphe	47
3.2.2 La typologie des graphes	48
3.3 La représentation matricielle des graphes	50
3.3.1 La matrice d'incidence	51
3.3.2 La matrice de distance	53
3.3.3 Les listes d'adjacence	54
3.4 Les indices morphologiques et topologiques	56
3.4.1 Les indices de connectivité	56
3.4.2 Les indices d'accessibilité et de centralité	58
3.4.3 Les indices de forme et de développement	60
3.5 Modèles d'optimisation: affectations et circuits	62
3.5.1 L'affectation de ressources	62
3.5.2 Les circuits	65
3.6 Modèles d'optimisation: le plus court chemin	67
3.6.1 Plus court chemin: matrices et algorithme de la cascade	69
3.6.2 Plus court chemin: arbres de parcours	71
3.6.3 Plus court chemin: recherche heuristique	77

📖 Chapitre IV 📖

4.0 Systèmes d'information géographique et système de support à la décision à référence spatiale	81
4.1 Origine et développement des systèmes d'information géographique	83
4.2 Structure et composantes d'un système d'information géographique	85
4.3 Représentation des données dans un système d'information géographique	90
4.3.1 La structure de données matricielle	93
4.3.2 La structure de données vectorielle	94
4.3.3 Les structures de données hybrides	97
4.3.4 Les données d'attributs	99
4.4 Les systèmes d'information géographique en transport	101
4.4.1 Les systèmes d'information géographique en planification et en gestion des réseaux de transport	102
4.4.2 Les systèmes d'information géographique et la protection civile	105
4.4.3 Les systèmes d'information géographique en transport des matières dangereuses	108
4.5 Les systèmes de support à la décision à référence spatiale	110

📖 Chapitre V 📖

5.0 Systèmes experts et théorie de la décision	114
5.1 Théorie de la décision	115
5.1.1 Formalisation du problème de décision	116
5.1.2 Arbre de décision	117

5.2 Théorie de la décision:l'approche probabiliste	119
5.2.1 De l'événement indépendant aux probabilités conditionnelles	120
5.2.2 Probabilités subjectives	123
5.3 Théorie de la décision: l'approche qualitative	124
5.3.1 Raisonnement qualitatif: l'arithmétique d'intervalle	126
5.3.2 Raisonnement qualitatif: incertitude et ensembles flous	129
5.4 Les systèmes experts	131
5.4.1 Approche conventionnelle et approche par intelligence artificielle	132
5.4.2 Structure d'un système expert	135
5.4.3 Représentation des connaissances et inférence	136
5.4.4 Traitement de l'incertitude dans un système expert	142

📖 Chapitre VI 📖

6.0 Transport des matières dangereuses et risques	147
6.1 Les matières dangereuses	148
6.2 Le transport des matières dangereuses au Canada	150
6.2.1 Le cadre législatif	151
6.2.2 La classification des matières dangereuses	153
6.3 Profil statistique des accidents	156
6.3.1 Caractéristiques générales des accidents	157
6.3.2 Le transport routier	162
6.4 La notion de risque	174
6.4.1 Les types de risques	175
6.4.2 Évaluation et mitigation des risques	178

📖 Chapitre VII 📖

7.0 Formalisation du modèle de recherche d'itinéraire	183
7.1 Propriétés du modèle de recherche d'itinéraire	184
7.2 Formalisation du modèle de recherche d'itinéraire	187
7.2.1 Sélection dynamique des tronçons sécuritaires	189
7.2.2 "K" plus courts chemins: sélection à postériori de l'itinéraire sécuritaire	193
7.2.3 Sélection à priori des tronçons sécuritaires	197
7.3 Architecture du système de support à la décision à référence spatiale	200
7.4 Structure de la base de connaissances du système expert	203

📖 Chapitre VIII 📖

8.0 Élaboration du système de support à la décision à référence spatiale	207
8.1 Démarche d'implantation du système de support à la décision à référence spatiale .	208

8.2 Région à l'étude et sources de données	209
8.2.1 Les sources de données	212
8.2.2 Les outils essentiels au projet	216
8.3 Production des bases de données	219
8.3.1 Structuration des données cartographiques	219
8.3.2 Intégration des données d'attribut	226
8.3.3 Élaboration de la base de connaissances	230
8.4 Le système de support à la décision	236
8.4.1 NodArc	236
8.4.2 Structure de données de NodArc	240
8.4.3 Algorithme du plus court chemin	241
8.4.4 Recherche d'itinéraire	243
8.5 Évaluation du système: les sources d'erreurs	247

▣ Chapitre IX ▣

9.0 Conclusion	250
Bibliographie	260
Annexe 1	281
Annexe 2	305

LISTE DES FIGURES

<u>Figure</u>		<u>Page</u>
2.1	Chemins alternatifs sur une surface anisotropique	17
2.2	Calcul du chemin optimum, surface anisotropique	18
2.3	Représentation d'un réseau sous forme de graphe	19
2.4	Plus court chemin dans un réseau	21
2.5	Réseau à impédance	22
2.6	Indécision due à l'impédance	24
2.7	Processus d'évaluation des risques	30
2.8	Minimisation coûts-risques dans un modèle à objectifs multiples	35
2.9	Système de navigation en temps réel	38
3.1	Simplification d'un réseau à l'aide d'un graphe	45
3.2	Éléments de base d'un graphe	46
3.3	Types de graphe	48
3.4	Classification topologique des réseaux	50
3.5	Représentation matricielle des réseaux	51
3.6	Matrice de distance	53
3.7	Représentation par liste d'adjacence	55
3.8	Nombre cyclomatique et nombre de polygones	57
3.9	Affectation de ressources dans un réseau	64
3.10	Classification des algorithmes du plus court chemin	68
3.11	Algorithme de la cascade	70
3.12	Arbres de parcours et algorithme de Dijkstra	72
3.13	Parcours en profondeur et en largeur	74
3.14	Algorithme de l'escalade	78
4.1	Composantes d'un système d'information géographique	87
4.2	Principales fonctions des systèmes d'information géographique	88
4.3	Représentation de la réalité de terrain dans un SIG	91
4.4	Structure de données matricielle	93
4.5	Structure de données vectorielle	95
4.6	Mise en relation des attributs	99
4.7	Accessibilité des données de transport	103
4.8	Gestion des données et prise de décisions	112
4.9	Architecture d'un SSDRS	113
5.1	Arbre de décision	118
5.2	Ensembles flous	130
5.3	Recherche en intelligence artificielle	132

5.4	Approche conventionnelle de traitement	133
5.5	Approche par système expert du traitement	133
5.6	Structure générale d'un système expert	135
5.7	Représentation par réseau sémantique	139
5.8	Représentation par objets structurés	140
6.1	Pictogrammes d'identification des matières dangereuses	155
6.2	Évolution du nombre d'accidents	157
6.3	Répartition par phase de transport	158
6.4	Répartition par modes de transport	159
6.5	Répartition des accidents par classes de matières dangereuses	160
6.6	Répartition des accidents par division de recensement	161
6.7	Répartition des accidents par classes de matières dangereuses	163
6.8	Répartition des accidents routiers selon l'effet	164
6.9	Répartition des accidents routiers selon le secteur	164
6.10	Répartition des accidents routiers selon la contamination	165
6.11	Répartition quotidienne des accidents routiers	166
6.12a	Répartition horaire des accidents, tous modes confondus	166
6.12b	Répartition horaire des accidents routiers	166
6.13	Répartition des accidents routiers par division de recensement	167
6.14	Facteur ayant provoqué l'accident routier	170
6.15	Degré de sévérité des accidents routiers	170
6.16	Degré potentiel de sévérité des accidents routiers	171
6.17	Composantes de la perception du risque	176
6.18	Coûts de la réduction des risques	179
6.19	Mitigation des risques et compromis	180
7.1	Heuristique classique de l'algorithme de Dijkstra	190
7.2	Modèle de recherche d'itinéraire à l'aide d'un système expert	191
7.3	"K" plus courts chemins et système expert	194
7.4	Sélection à priori des tronçons sécuritaires	198
7.5	Limitations de la sélection à priori des tronçons sécuritaires	199
7.6	Structure du système de support à la décision à référence spatiale	201
7.7	Développement d'une base de connaissances	203
7.8	Hiérarchisation des faits menant au choix d'un itinéraire	205
8.1	Étapes de l'implantation du SSDRS	208
8.2	Région à l'étude	210
8.3	Répartition de la population	210
8.4	Grands axes routiers de la région	211
8.5	Réseau routier brute, région d'Ottawa-Hull	214
8.6	Architecture d'Arc/Info	217
8.7	Structuration des données cartographiques	220
8.8	Réseau routier simplifié, région d'Ottawa-Hull	222
8.9	Erreurs liées au mode de saisie	221
8.10	Erreurs de topologie	223
8.11	Discordances de superposition	225

8.12	Limites de la RMR et de la carte topographique	226
8.13	Intégration des données d'attributs	227
8.14	Rattachement des attributs	227
8.15	Densité de population, RMR d'Ottawa-Hull	228
8.16	Répartition des logements privés, RMR d'Ottawa-Hull	229
8.17	Répartition des tours d'habitations de + de 5 étages	229
8.18	Organigramme simplifié de la base de connaissances	231
8.19	NodArc, fenêtres d'affichage	237
8.20	NodArc, menu NODARC	238
8.21	NodArc, menu SYSTEM	238
8.22	NodArc, menu CONVERT	238
8.23	NodArc, menu NETWORK	239
8.24	NodArc, menu VIEW	239
8.25	NodArc, menu ROUTE	239
8.26	Intégration des données dans NodArc	240
8.27	Fenêtre de saisie des paramètres	243
8.28	Itinéraire retenu sans tenir compte de l'hippodrome	244
8.29	Fichier de sortie et table de routage	245
8.30	Trajet retenu en tenant compte de l'hippodrome	245
8.31	Fichier de sortie et table de routage	246
8.32	Secteur inaccessible le jour	247

LISTE DES TABLEAUX

<u>Tableau</u>		<u>Page</u>
6.1	Classification des matières dangereuses	154
6.2	Tableau de contingence, province/classe de matières	169
6.3	Tableau de contingence, sévérité/secteur	172
6.4	Tableau de contingence, sévérité potentielle/secteur	173
8.1	Inventaire des données pour la prise de décision	212

Chapitre I

Introduction

“Faire de la recherche, c’est apprendre à gérer l’incertitude et l’inconnu. C’est aussi faire des erreurs et apprendre à les identifier.”

-Daniel Latouche

Le présent projet de thèse s’inscrit dans un courant de pensée qui préconise l’emploi de modèles et de théories pour enrichir le domaine de l’analyse spatiale et aborder les problèmes de la géographie. Il est le résultat d’une influence subie par la lecture d’auteurs tels que BUNGE, HAGGETT et CHORLEY qui, en tant que géographes, ont marqué leur époque et en tant que penseurs, ont marqué des générations de géographes.

La révolution quantitative qui a pris place au cours des années 50 et 60 a profondément bouleversé la géographie en la faisant passer d’une approche descriptive et chorographique à une approche beaucoup plus analytique et systémique. La révolution quantitative a également permis de mettre en évidence l’existence d’une science de l’espace, une science qui témoigne

de l'interdépendance de la géographie et de la géométrie, de l'analyse spatiale et des "mathématiques de l'espace". BUNGE (1962) dans son livre intitulé *Theoretical Geography* fut l'un des premiers géographes à proposer une géographie théorique comme réponse méthodologique aux problèmes géographiques en prenant pour base la nature générale de la science.

En faisant de la géographie une science de la localisation (location science) capable de prédire une distribution spatiale, BUNGE introduisait non seulement les concepts de modèle, de théorie et de système mais également donnait aux mathématiques un rôle majeur dans le processus de modélisation. Contrairement aux autres géographes quantitatifs, BUNGE s'est intéressé à l'emploi des mathématiques non seulement pour leur aspect métrique ou statistique mais aussi pour leur aspect géométrique et topologique. Cette approche avait pour mérite de permettre une plus grande intégration de la géographie quantitative en analyse spatiale et de permettre l'expansion du mot "spatial" qui ne signifiait plus uniquement la distance mais aussi toutes les propriétés géométriques de la surface terrestre telles que la forme, la dimension, le pattern, la distance et la connectivité. Cette dernière propriété devait d'ailleurs, de pair avec la théorie des graphes, être à la source du domaine de la géographie des transports et de l'analyse de réseau en géographie.

BURTON (1963) devait également plaider pour une géographie axée sur la construction de modèles et l'établissement d'un ensemble de théories pouvant servir de base à la recherche géographique. BURTON y voyait là l'occasion pour les géographes quantitatifs d'établir une

nouvelle géographie et concluait en présentant la théorie des graphes comme un exemple de géographie théorique:

"It is conceivable that a body of useful theory could be built up around the application of graph theory to geographical problems. This is an example of what is meant by theoretical geography."
Burton (1963, p.21).

HAGGETT (1965) reprit cette idée en proposant cinq éléments fondamentaux autour desquels les théories et les modèles quantitatifs en géographie tant physique qu'humaine pourraient s'intégrer. Des cinq éléments proposés par HAGGETT, le premier et le plus fondamental est le *mouvement* dans l'espace, auquel se greffent les *réseaux* par lesquels les mouvements s'opèrent, puis les *noeuds* comme centres d'interaction ainsi que les *hiérarchies* et les *surfaces* pour exprimer le concept de compétition dans l'espace. Bien qu'HAGGETT ait proposé ces éléments dans le contexte global de la géographie théorique, il est surprenant de constater leur similarité avec les concepts de base de la géographie des transports. D'ailleurs, HAGGETT et CHORLEY devaient publier en 1969 un livre intitulé *Network Analysis in Geography*, l'un des premiers ouvrages à traiter de façon exhaustive de l'analyse de réseaux en géographie physique et humaine.

En plus d'extirper de sa torpeur le domaine de la géographie des transports et de l'analyse de réseau en géographie, les travaux des années cinquante et soixante ont contribué à populariser chez les géographes les termes analyse spatiale et modélisation. L'emploi des modèles en géographie a fait l'objet de nombreuses publications, notamment CHORLEY (1964), CHORLEY et HAGGETT (1967), HARVEY (1969), où la définition et la classification

des concepts de modèle et d'analyse spatiale occupaient une grande place. Comme l'approche analytique en géographie repose pour une bonne part sur ces notions et qu'elles sont partie intégrante de cette thèse, nous croyons important de définir dès maintenant ce que l'on entend par modèle et par analyse spatiale.

En s'inspirant des divers auteurs ayant traité du sujet et plus particulièrement de CHORLEY et HAGGETT (1967) il est possible de dégager 4 grandes caractéristiques des modèles. Les modèles sont des *systèmes simplifiés* car ils offrent une représentation simplifiée de la réalité. Les modèles sont aussi *sélectifs* dans la mesure où ils permettent de filtrer le "bruit des données" et de rejeter les éléments non essentiels à l'explication d'un phénomène. Ils sont également une *construction structurée* car ils exploitent les interrelations entre chaque composante d'un système. Finalement, les modèles sont *évocateurs* dans leur capacité de présenter des indices à leurs utilisateurs dans le but d'améliorer et d'étendre leur portée. Bref, l'objectif premier d'un modèle n'est pas de reproduire la réalité dans ses moindres détails mais de la comprendre, d'en proposer une interprétation; ce qui est complètement différent. Dans le cadre de cette thèse nous croyons important de préciser qu'en plus de ces quatre caractéristiques, *les modèles sont aussi un raisonnement sur les dimensions temporelle et spatiale de l'espace géographique.*

Quant à l'analyse spatiale, tout aussi fondamentale dans ce projet de thèse, elle a également pour origine le développement des méthodes quantitatives et statistiques des années cinquante. Toutefois, la définition de l'analyse spatiale demeure sujette à interprétation. On s'entend cependant pour dire que l'analyse spatiale concerne l'étude détaillée de la distribution

spatiale d'éléments ponctuels, linéaires et zonaux (Openshaw, 1991). Goodall (1987) est plus spécifique et la définit comme étant une description, sur une base systématique et rigoureuse, de la distribution spatiale d'un phénomène et une tentative d'identifier les facteurs contrôlant cette distribution. Dans le cadre de ce projet, nous considérons l'analyse spatiale comme un mode d'étude des phénomènes géographiques qui s'appuie sur l'interdépendance de la géographie, de la géométrie et des mathématiques.

Outre la révolution quantitative, un deuxième grand thème a également servi de catalyseur à cette thèse de doctorat. Le seul développement d'une géographie théorique et quantitative ne pouvait porter fruit sans l'appui technique d'outils capables d'opérationnaliser les concepts théoriques. Tout comme les autres sciences, la géographie est dépendante des nouveautés techniques pour son avancement, tout comme les nouveautés techniques sont tributaires des nouveautés théoriques. Le progrès de l'informatique est, sans aucun doute, le développement technologique qui a eu le plus d'impact sur la géographie.

L'émergence des outils informatiques de même que le caractère multidisciplinaire de la géographie ont eu très tôt pour effet de sensibiliser les géographes au problème de la disponibilité et de la valeur de l'information géographique. Au cours des années soixante-dix, la collecte, la saisie semi-automatisée et le transfert sous forme numérique de l'information constituaient l'une des principales tâches des spécialistes de l'analyse spatiale. Un tel travail ne pouvait se réaliser sans outils de travail adéquats et c'est à ce moment qu'est apparu le concept de système d'information géographique (SIG).

Conçus au départ strictement dans le but de manipuler des données spatiales, les SIG sont demeurés pendant plus d'une décennie qu'un simple outil de stockage. Ce n'est qu'au cours des années quatre-vingt et quatre-vingt-dix, une fois les grandes banques de données disponibles, que les besoins en analyse spatiale se sont fait sentir. Bien que jusqu'au début des années quatre-vingt-dix les SIG aient été particulièrement déficients en terme d'analyse spatiale, nous croyons qu'ils n'en demeurent pas moins la plate-forme idéale sur laquelle développer les modèles et les analyses spatiales nécessaires aux géographes. Cette thèse a donc pour but indirect de démontrer cette affirmation et ce, plus particulièrement dans le domaine de la géographie des transports.

1.1 Le sujet de la thèse

Bien que le présent ouvrage corresponde aux critères d'une thèse de doctorat, nous avons aussi voulu aborder son sujet sous la forme d'une "syn-thèse" qui nous permette de réunir en un tout divers concepts et approches desquels apparaîtra une nouvelle solution au problème exposé. Comme l'évoque le titre de la thèse, il sera question de transport. S'il est un thème imminemment géographique, c'est bien celui du transport. Le transport fait référence à la notion de distance, à la base même de la géographie. Le transport implique également une différenciation du lieu ainsi qu'une compétition entre ceux-ci, phénomène à la base même de la dynamique géographique.

Comme l'indique également le titre de la thèse, il sera question de déplacements dans les réseaux et plus particulièrement des modèles de calcul d'itinéraire. Développés dans le cadre des mathématiques discrètes et de la théorie des graphes, les algorithmes permettant de

déterminer le plus court chemin entre deux points d'un réseau visent un objectif unique, à savoir la minimisation de la distance parcourue ou la minimisation des coûts de transport. Cette approche purement mathématique, souvent employée en géographie, est pourtant mal adaptée aux problèmes exigeant une approche à objectifs multiples ou encore à critères multiples. C'est le cas notamment du transport des matières dangereuses où l'approche habituelle consiste à représenter le risque d'accidents par un coefficient de friction entre chaque point du réseau sans tenir compte de la dynamique engendrée par les facteurs de risque. La présente thèse abordera donc cet aspect de la théorie des graphes en démontrant les lacunes de cette approche dans le cas particulier du transport des matières dangereuses.

Étant donné qu'une thèse se doit d'apporter une réponse nouvelle au problème qu'elle aborde, cette recherche veut également démontrer qu'il est possible d'améliorer les modèles de calcul d'itinéraire issus de la théorie des graphes en leur adjoignant une capacité déductive par intégration à un système expert. Cette union d'un système expert avec l'approche algorithmique de la théorie des graphes constitue d'ailleurs le noyau central du projet de recherche. Dans une vision globale du modèle, nous verrons également comment cet ensemble pourrait bénéficier d'une fusion avec un système d'information géographique pour créer un système de support à la décision à référence spatiale (SSDRS) pour les intervenants impliqués dans le transport des matières dangereuses. Enfin, dans le but d'illustrer la méthodologie d'implantation d'un tel système, nous appliquerons à une base de données restreinte de la région d'Ottawa-Hull, le modèle élaboré au cours de cette recherche.

1.2 Justification du sujet

Du point de vue géographique, les recherches en transport se sont beaucoup plus concentrées, au cours des dernières années, sur les problèmes de flux dans les réseaux, les enquêtes origine-destination ainsi que l'allocation de ressources (location-allocation study). Le rejet de l'approche quantitative par beaucoup de géographes ainsi que l'absence de logiciels basés sur la théorie des graphes ont probablement contribué à cette situation. En soi, rouvrir le dossier conceptuel de la théorie des graphes tel que proposé lors de la révolution quantitative puis par la suite abandonné, justifie l'existence de cette thèse. Dans son intégration d'une approche déductive basée sur les systèmes experts à la démarche algorithmique traditionnelle, cette thèse se veut une contribution à l'usage de la théorie des graphes en géographie.

L'intérêt porté au transport des matières dangereuses n'est pas non plus un hasard. Le transport des matières dangereuses constitue une activité économique importante qui peut avoir, en cas de catastrophe, un impact sur la vie et la santé des gens, les biens publics et privés, la faune, la flore et l'environnement en général (OCDE, 1988). Il s'agit d'un sujet délicat puisqu'il fait référence à un sentiment de crainte chez la population en général. Le transport des matières dangereuses est aussi un sujet épineux en raison de la quantité croissante de matières dangereuses charroyées annuellement et des difficultés de prévoir les conséquences d'un accident les impliquant. Dans ce domaine, les géographes spécialisés en transport sont en mesure d'apporter une contribution grâce aux outils d'analyse spatiale dont ils ont la maîtrise.

1.3 Objectifs de la thèse

Le but général de cette thèse est de démontrer qu'il est possible d'introduire une capacité de raisonnement dans les modèles de recherche d'itinéraires en les fusionnant à un système expert. Le projet de recherche aborde quatre grands thèmes qui reflètent le sens général de notre démarche. Ces thèmes sont la théorie des graphes et l'approche algorithmique conventionnelle; les systèmes d'information géographique et les systèmes de support à la décision à référence spatiale; la théorie de la décision et les systèmes experts et enfin, le transport des matières dangereuses et la notion de risque. Ces grands thèmes forment l'assise de la thèse de laquelle il est possible de dégager les objectifs suivants:

Notre premier objectif général vise essentiellement à faire le point sur les modèles de recherche d'itinéraires basés sur la théorie des graphes pour en dégager les fondements théoriques nécessaires au projet ainsi que les faiblesses en ce qui a trait à l'établissement d'itinéraires à partir de critères quantitatifs et qualitatifs.

Notre deuxième objectif général est de dégager, du domaine de la théorie de la décision et des systèmes experts, les éléments susceptibles de contribuer à améliorer l'approche conventionnelle de recherche d'itinéraires en fusionnant un système expert aux algorithmes traditionnels.

Le troisième but de ce projet est d'explorer le domaine du transport des matières dangereuses, plus particulièrement en territoire canadien, pour mettre en évidence les principales caractéristiques de ces accidents et aussi, aborder la notion de risque et de sa mitigation.

Enfin, notre quatrième objectif est de formaliser un modèle de recherche d'itinéraire alliant algorithme traditionnel et système expert et d'en démontrer la démarche d'implantation à la région d'Ottawa-Hull.

1.4 Les limites de la thèse

Tout projet de recherche comporte des limitations et des choix a priori qui ont pour effet de privilégier certains aspects et d'en négliger d'autres. Bâtir un projet de recherche, c'est faire des choix et les assumer. Ce projet ne fait pas exception à la règle. En abordant ce projet, nous avons délibérément omis de nous attarder à certaines perspectives, notamment en ce qui concerne le transport des matières dangereuses.

Le présent projet de thèse n'aborde pas, entre autres, les aspects légaux et éthiques du transport des matières dangereuses. La mise en place d'un cadre légal très strict en ce qui concerne ce type de transport pourrait, s'il est bien adapté au contexte canadien, permettre de minimiser les risques liés au transport de ces matières. Cependant, un cadre légal n'a aucune valeur s'il n'est pas réellement mis en oeuvre sur le terrain. Dans un contexte économique difficile, les transporteurs ont souvent tendance à faire passer la marge de profit avant les questions de sécurité.

Nous avons également omis délibérément les questions éthiques qui ont trait à l'établissement d'itinéraires de transport désignés. En concentrant le transport de matières dangereuses dans un corridor donné, n'augmente-t-on pas le risque encouru par la population vivant dans ce secteur? Comment justifier l'accroissement de danger pour cette population? Ne devrait-on pas plutôt répartir sur l'ensemble du réseau le transport des produits dangereux pour ainsi distribuer

également le risque sur tous les citoyens? Quels critères doit-on prendre en compte pour qualifier une population de vulnérable à ce type de transport et en exclure d'autres? Autant de questions qui pourraient faire l'objet d'une autre thèse et que nous avons dû omettre pour nous concentrer sur les aspects spatiaux du problème.

Du point de vue technique, certaines limitations sont également inévitables. A notre connaissance, aucun système expert n'a été couplé à un système d'information géographique de façon *réellement fonctionnelle*. Bien que suggéré par plusieurs auteurs (nous en ferons état dans les prochains chapitres), aucun produit n'a encore été mis sur le marché. Pour pallier à ce problème, la seule solution consiste à développer des interfaces entre les systèmes existants pour faciliter la communication entre chacune de ces "boîtes noires". Pour ce projet, le logiciel de système d'information géographique ARC/INFO PC^{md} et le système expert CLIPS^{md} ont été mis à profit et ce non sans problèmes.

Le but de la thèse n'étant pas de développer un logiciel mais de démontrer le potentiel du couplage théorie des graphes-systèmes experts et de surcroît n'étant pas nous-même programmeur professionnel, nous avons tenté de réduire au minimum la démarche de programmation. Malgré cela, plusieurs centaines d'heures ont été consacrées à cette tâche, ce qui nous laisse croire que pour en arriver à un produit commercialisable, beaucoup de travail serait à faire.

1.5 Structure de la thèse

La présente thèse se divise en neuf chapitres. Le premier chapitre introduit le contexte général de la recherche, le sujet, les objectifs et les limites de la thèse. La problématique de la recherche fait l'objet du deuxième chapitre. La géographie des transports et les limitations des modèles actuels de recherche d'itinéraires basés sur l'impédance y sont abordés. Le chapitre 3 est consacré entièrement à la théorie des graphes et à son utilisation en géographie. Une attention particulière est portée au modèle d'optimisation qu'est l'algorithme du plus court chemin.

Le quatrième chapitre présente le domaine des systèmes d'information géographique (SIG) et des systèmes de support à la décision à référence spatiale (SSDRS). Nous nous attarderons particulièrement à décrire la structure des SIG ainsi que leurs apports en géographie des transports. La théorie de la décision et les systèmes experts sont le sujet du chapitre 5 où l'on aborde, entre autres, le raisonnement qualitatif et la représentation des connaissances dans un système expert. Quant au chapitre 6, il présente la situation du transport des matières dangereuses au Canada, à partir d'un profil statistique des accidents. La notion de risque est également discutée à l'intérieur de ce chapitre.

Les chapitres 7 et 8 exposent les résultats de nos réflexions. Dans un premier temps, le chapitre 7 est consacré à la formalisation du nouveau modèle de recherche d'itinéraire et de ses variantes de même qu'à la présentation de la structure du SSDRS. Dans un deuxième temps, le chapitre 8 présente la méthodologie d'implantation d'un tel système dans le contexte d'une région urbaine comme Ottawa-Hull.

Enfin, le chapitre 9 apporte une conclusion générale sur la thèse et présente également quelques avenues de développements futurs.

Chapitre II

Problématique

"Et pourtant elle se meut"

- Galilée

Cette repartie que l'on attribue à Galilée lors de sa convocation au tribunal de l'Inquisition en 1633 démontre bien une des caractéristiques fondamentales de la Nature: le mouvement. Que ce soit dans ses plus infimes retranchements, tels que l'atome, ou à l'échelle du cosmos, le mouvement est omniprésent et continu dans la nature. Si l'on affirme souvent que la nature a horreur du vide, elle a tout autant horreur de l'immobilisme. Chez l'humain, le premier témoignage d'une activité de transport organisée a probablement été le sentier tracé et entretenu par nos ancêtres préhistoriques. Le sentier était d'une importance vitale pour ces communautés que ce soit pour la chasse, la cueillette, les communications et la guerre entre tribus. Au fil de l'évolution humaine, l'invention de la roue et l'utilisation du cheval comme moyen de traction ont amené le développement des premières routes carrossables et ainsi

l'expansion des capacités de transport. Plus puissant que l'humain, le cheval permettait de transporter plus loin des charges plus lourdes. La mise sur pied d'infrastructures routières adaptées à ce nouveau moyen de transport allait modifier radicalement la société.

L'avènement de la révolution industrielle a également contribué à modifier le rapport de l'humain face au transport. Tout d'abord, l'invention de la machine à vapeur puis du moteur à explosion a décuplé la capacité de se déplacer tant en vitesse qu'en terme de charge. De façon plus subtile, la révolution industrielle a également engendré un besoin de transport en introduisant le concept de spécialisation. En concentrant dans une région donnée les usines de fabrication des produits, l'approvisionnement en matières premières des usines et la distribution des produits manufacturés vers les consommateurs reposaient obligatoirement sur une capacité de transport accrue. La spécialisation des activités humaines a aussi eu pour effet de spécialiser l'occupation du territoire en aires fonctionnelles donc d'accroître les déplacements. Quiconque affronte quotidiennement les embouteillages aux heures de pointes réalise l'importance voir même l'impact négatif que peut avoir le phénomène des transports sur nos vies. Il n'est donc pas étonnant que ce soit développé en géographie un axe de recherche qui s'attarde au domaine des transports.

2.1 Géographie des transports

Les phénomènes typiquement géographiques que sont les échanges, les transformations, les transferts et les flux dans les réseaux ont été abordés de deux façons dans le passé. Une première approche, moins répandue dans la géographie contemporaine mais fort populaire avant les années cinquante, s'appuyait sur une méthode énumérative et descriptive inspirée de

la géographie régionale. L'historique du développement des réseaux et leur impact sur les régions y occupaient une place majeure (à titre d'exemple, voir WOLKOWITSCH, 1973 ainsi que BAMFORD et ROBINSON, 1978).

Ce type d'études donnait un portrait de la situation, un peu comme les statistiques descriptives, mais n'abordait d'aucune manière la modélisation des réseaux ou l'étude de leur morphologie et de leur mécanique interne. Cette géographie des transports se limitait généralement à amasser des données sur la longueur, la demande et la capacité des réseaux, à les localiser et les décrire sans s'attarder à expliquer le pourquoi et le comment. La vague quantitative a cependant fait apparaître une deuxième approche de la géographie des transports, vague qui s'est manifestée, dans un premier temps, dans les études de localisation (location studies).

C'est avec la popularisation des modèles issus de l'école allemande du XIX^{ième} siècle, notamment ceux de Von Thünen et de Christaller, puis du modèle de Weber sur la localisation industrielle minimisant les coûts de transport que la notion de distance et de transport a pris sa place dans la modélisation géographique. Cependant, tout comme le modèle gravitationnel, ces modèles avaient comme principal défaut d'associer l'espace géographique à une plaine isotropique loin de correspondre à la réalité de terrain. Malgré leur apport indéniable en géographie théorique, cet inconvénient explique pour une bonne part la perte de vitesse qu'ont vite connue ces modèles chez les géographes. Néanmoins, ceux-ci ont été un premier pas important dans l'établissement d'une modélisation des transports et dans le concept du déplacement au moindre coût (least cost path).

C'est à partir de cette idée du déplacement à moindre coûts ou de minimisation de la friction que WARNTZ (1961) a apporté la solution au problème de la détermination du chemin optimum sur une surface anisotropique. WARNTZ s'est attaqué à ce problème en utilisant l'exemple d'un avion transatlantique dont le trajet optimum pouvait être déterminé en prenant comme critères la vitesse et la direction des vents sur une surface continue échantillonnée à l'aide d'une grille (matrice). Dès 1954, August Losch avait remarqué qu'en transformant la surface anisotropique sous forme de matrice dont chaque cellule contient la valeur de friction associée à son passage, le problème du chemin optimum devenait analogue à celui de la réfraction de la lumière traversant des milieux de densités différentes (WAGNER *et al*, 1993). Voir figure 2.1

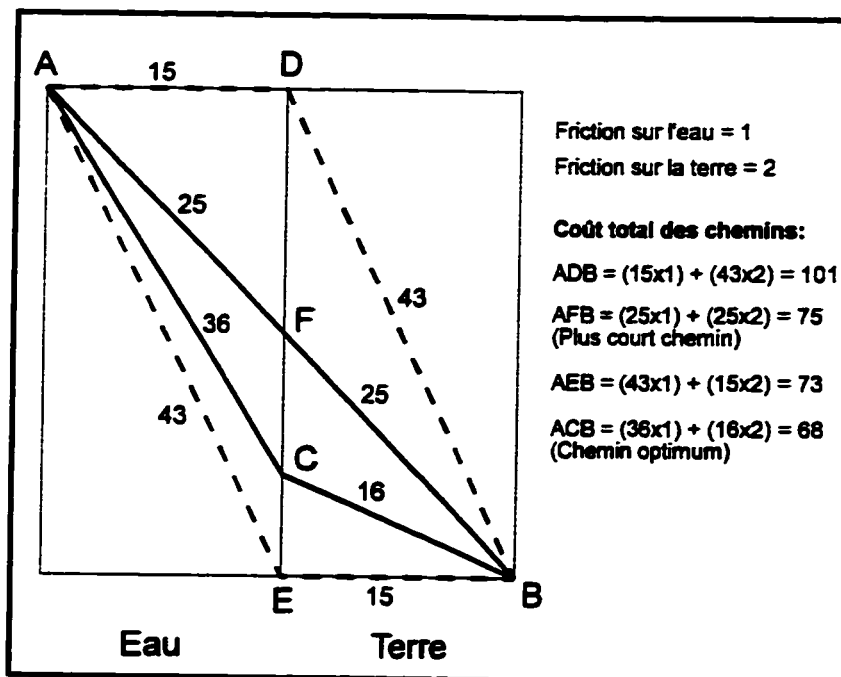


Figure 2.1: Chemins alternatifs sur une surface anisotropique.
D'après Lowe et Moryadas (1975, p. 19)

En se basant sur l'observation de Losch et sur le problème du transatlantique WARNTZ développa une solution en deux étapes. D'une matrice de départ contenant les valeurs de

résistance au déplacement, une matrice de coûts de déplacements de la cellule d'origine à toutes les autres cellules de la matrice est créée. À la seconde étape, la matrice de coûts de déplacements est balayée et le chemin de moindre résistance de la cellule d'origine à la cellule de destination est retenu (figure 2.2).

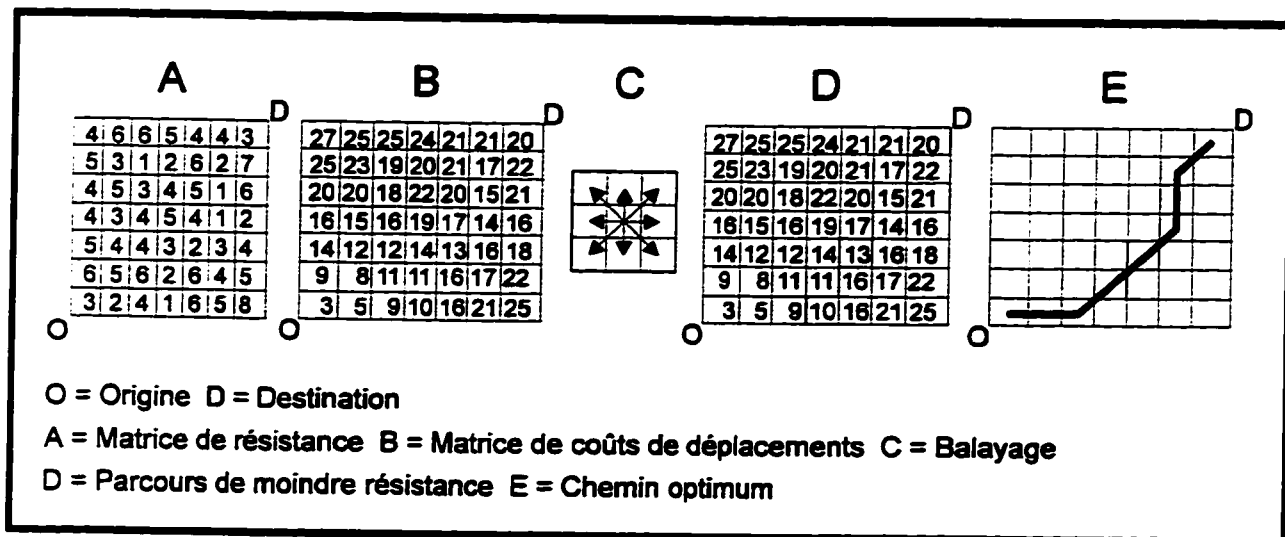


Figure 2.2: Calcul du chemin optimum, surface anisotropique

Cette solution au problème du chemin optimum présente certains avantages. Elle permet de considérer l'ensemble du territoire à l'étude et se prête bien aux applications en géographie physique où l'emploi de cette méthode, en conjonction avec un modèle numérique de terrain (MNT), sert entre autres à modéliser l'écoulement des eaux. Les militaires l'utilisent également de façon courante dans leurs systèmes de combats tactiques (DOUGLAS, 1994). Cependant, cette solution pose problème lorsqu'il s'agit de déplacements sur une infrastructure de transport linéaire comme les routes. On doit alors se tourner vers une représentation vectorielle du phénomène en remplaçant la matrice par un graphe.

2.2 Déplacements dans les réseaux¹

Dès le milieu du siècle les géographes ont été tenté d'aborder le déplacement dans les réseaux en mode vectoriel. Contrairement à la matrice qui, même à résolution fine, n'est jamais conforme à la réalité des réseaux, la structure vectorielle s'ajuste mieux aux phénomènes linéaires, qu'il s'agisse de courbes, de méandres ou de boucles. Le calcul de la distance, fondamental dans les réseaux, est également plus précis dans l'approche vectorielle. Un autre avantage du mode vectoriel est qu'il se prête bien à la représentation des réseaux sous forme de graphe (figure 2.3).

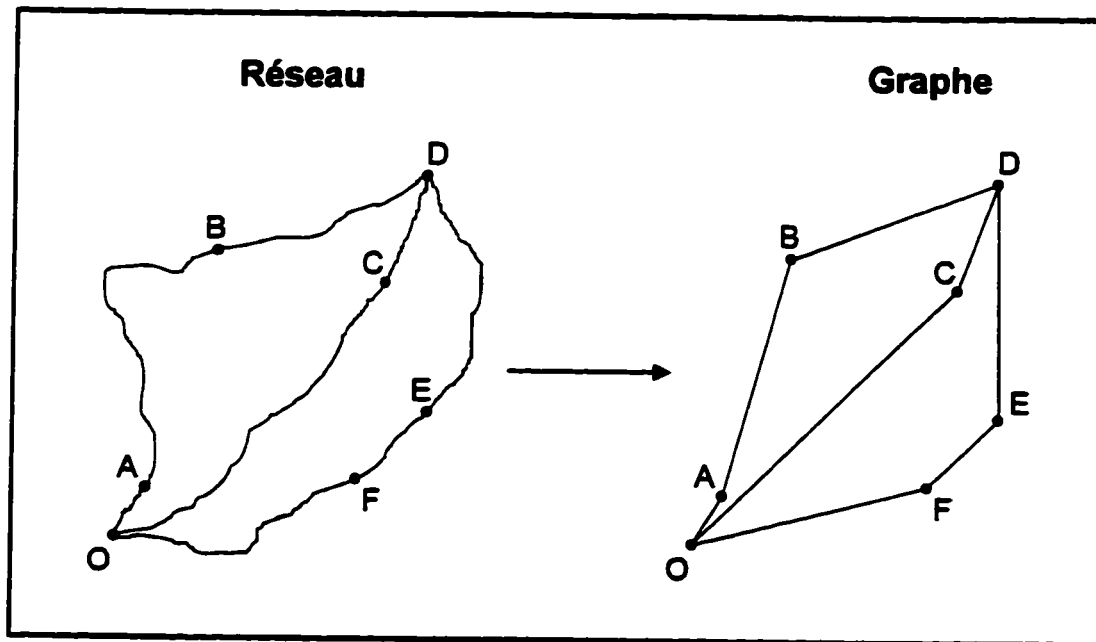


Figure 2.3 Représentation d'un réseau sous forme de graphe

L'intérêt de cette représentation réside dans sa capacité de simplifier les réseaux complexes tout en conservant l'information de départ. Par analogie, on peut comparer le graphe d'un

¹La présente section n'a pour but que d'introduire le lecteur à la théorie des graphes dans le cadre de la problématique de la thèse. Le chapitre 3 traite de façon plus approfondie de cette théorie.

réseau à une carte; ils sont tous deux un modèle graphique qui synthétise une réalité de terrain parfois fort complexe.

Toutefois, l'avantage majeur de l'emploi des graphes en transport repose sur l'accès qu'ils donnent à la théorie des graphes, théorie qui s'appuie elle-même sur une base mathématique solide. La présence des mathématiques a ouvert la porte au développement de nombreux modèles dont plusieurs ont trait à l'optimisation du déplacement dans les réseaux. Cela est notamment le cas de l'algorithme du calcul du plus court chemin.

2.2.1 Le plus court chemin

La solution menant à la détermination du plus court chemin est une solution d'optimisation qui fait appel à la programmation dynamique. Cette dernière est une technique d'optimisation systématique d'une séquence de décisions dans laquelle l'optimalité du résultat final (dans ce cas-ci le plus court chemin) dépend de l'optimalité de chaque décision individuelle effectuée précédemment (WERNER, 1985, p. 46).

La figure 2.4 nous permet d'illustrer ce processus dynamique. Supposons que l'on veuille déterminer le plus court chemin entre les points "O" (origine) et "D" (destination). Bien que le tronçon de route O-A puisse paraître être le choix optimal (segment le plus court, donc minimisation de la distance) lors de la première itération, il doit être rejeté puisque les itérations ultérieures vont démontrer que cette décision ne conduit pas au chemin le plus court entre "O" et "D". Le tronçon de route O-C (bien que plus long que O-A), doit être retenu pour parvenir à la solution optimale. Déterminer le plus court chemin est donc un processus itératif et

dynamique basé sur un processus décisionnel dont le résultat est remis en cause à chaque itération jusqu'à l'obtention du résultat optimal. Avec ce genre d'algorithme, toutes les combinaisons possibles doivent être examinées avant d'affirmer que l'on a bel et bien le chemin le plus court.

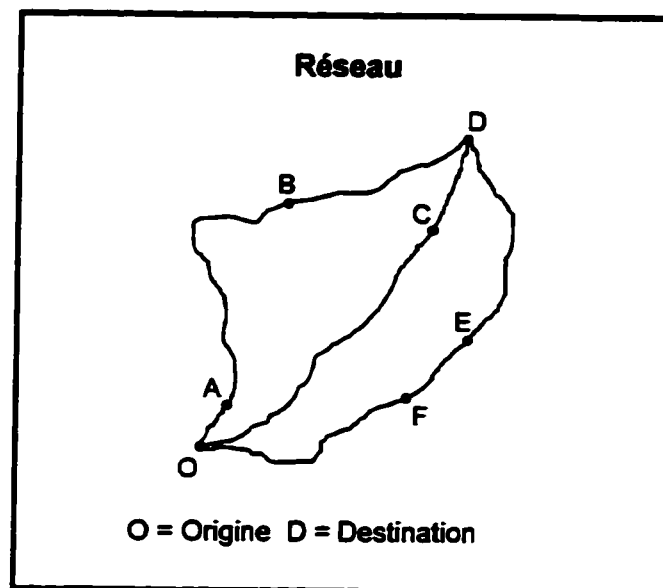


Figure 2.4: Plus court chemin dans un réseau

Bien que la distance cartésienne serve généralement de métrique, les géographes ont développé au fil des ans plusieurs variations à la distance cartésienne. On peut, par exemple, calculer le chemin le plus court à partir d'une distance-temps, d'une distance-coûts ou encore d'une distance psychologique (LOWE et MORYADAS, 1975). La distance perd alors son sens unique de séparation entre deux points dans l'espace pour devenir une notion de friction, de résistance, mieux connu en physique sous le terme d'impédance.

2.2.2 L'impédance

L'impédance est un phénomène de physique électrique équivalant à la résistance électrique entre deux points d'un conducteur. L'analogie avec un réseau de transport est parfaite. En substituant l'impédance à la distance dans l'algorithme du plus court chemin on ne calcule plus le chemin le plus court mais le chemin optimum de moindre résistance entre chaque point du réseau. Plus l'impédance sera élevée dans un tronçon du réseau, moins il y aura de chance que ce segment soit retenu dans le calcul du chemin optimum.

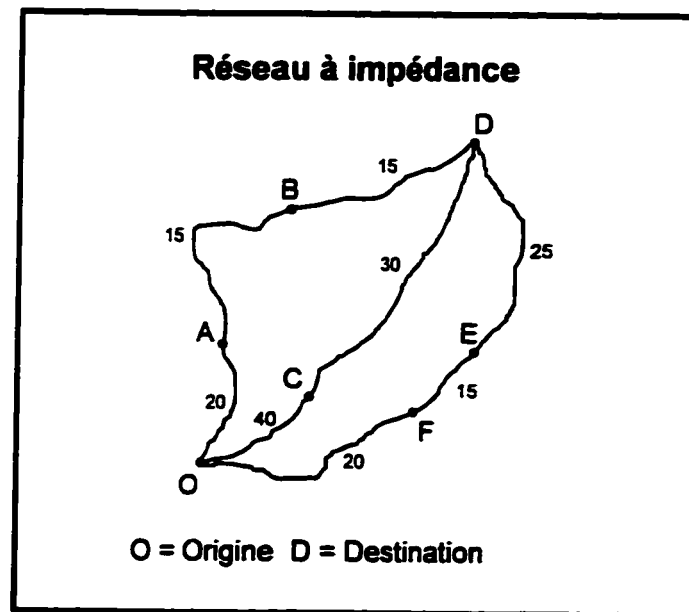


Figure 2.5: Réseau à impédance

Dans le réseau de la figure 2.5, bien que la route la plus courte soit **O-C-D**, la résistance est telle dans les segments **O-C** et **C-D** que l'itinéraire **O-A-B-D** lui est préférable.

L'impédance étant une variable dérivée, c'est-à-dire calculée à partir de la combinaison de plusieurs autres variables, son emploi a été très populaire en transport. De plus, l'emploi de

l'impédance ne modifie en rien les algorithmes de calcul d'itinéraires de la théorie des graphes puisqu'il ne s'agit que de substituer l'impédance, une valeur numérique, à la distance, une autre valeur numérique. D'ailleurs, le système d'information géographique ARC/INFO^{md}, le logiciel le plus vendu dans le monde, utilise le concept d'impédance dans tous ses modèles de transport. Cette approche comporte cependant, à notre avis, des faiblesses qui en limite l'application dans certains domaines.

2.3 Les limitations de l'impédance

Dans le cas des modèles relativement simples ou peu de variables entrent en compte, l'impédance peut tout à fait bien répondre aux besoins. La situation est différente lorsqu'il s'agit de phénomènes complexes où plusieurs variables doivent être considérées. Supposons un réseau quelconque pour lequel on possède, pour chaque tronçon, une série de variables $V_1, V_2, V_3 \dots V_n$. On peut alors définir l'impédance pour un tronçon donné comme étant la combinaison des variables V_1 à V_n pour ce tronçon.

$$\text{Impédance} = \{ V_1, V_2, V_3 \dots V_n \}$$

Pour la plupart des réseaux, surtout si les variables du modèles sont de type binaire (présence/absence), il arrive assez souvent que deux ou plusieurs tronçons obtiennent la même valeur d'impédance. Ce genre de situation empêche ou du moins réduit sérieusement la justesse du processus décisionnel pour une itération donnée. En fait il devient impossible pour l'algorithme de discerner entre les tronçons (figure 2.6). Si cette condition se répète pour

plusieurs itérations, comme c'est souvent le cas, le résultat final peut s'écarter sensiblement du résultat optimum².

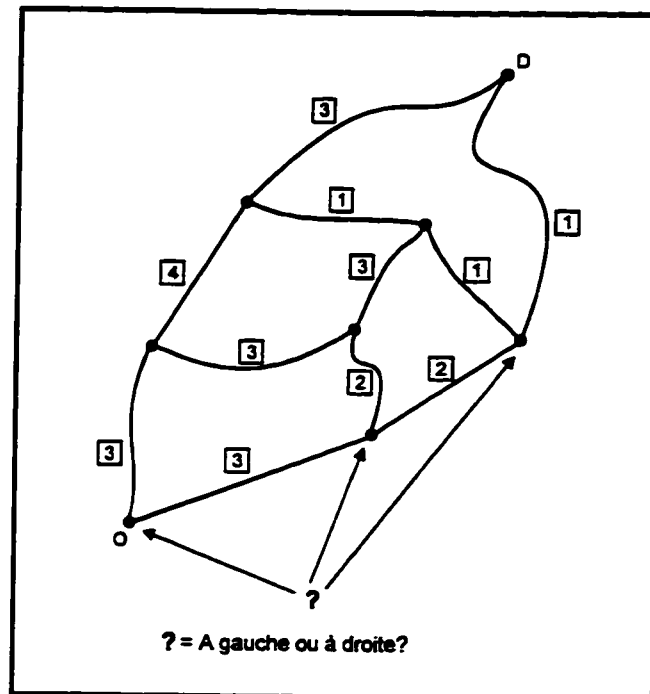


Figure 2.6: Indécision due à l'impédance.

Une première raison pouvant expliquer une telle situation est le manque de capacité discriminante des variables ayant servi au calcul de l'impédance. Dans ce cas, pour remédier au problème, l'utilisateur n'aurait qu'à refaire un choix plus judicieux de variables et recalculer l'impédance pour chaque tronçon. Une autre raison réside dans le concept même d'impédance. En réduisant à une valeur numérique unique une combinaison de variables, on perd nécessairement une partie de l'information. C'est un peu comme tenter de décrire un individu à

² Dans une situation de ce genre, certains algorithmes vont alors utiliser la distance cartésienne en plus de l'impédance pour permettre la discrimination.

l'aide de son poids uniquement. Deux personnes de même poids peuvent avoir un physique totalement différent!

L'autre lacune de la méthode de l'impédance consiste en son absence de mécanisme permettant de "raisonner" sur les caractéristiques de chaque tronçon. En éliminant l'impédance et en conservant l'accès aux variables lors du calcul de l'itinéraire, il serait alors possible d'appliquer des conditions du type "IF <condition A> THEN <action B>". Par exemple, lors du calcul du plus court chemin, une condition du type "IF <route non-pavée> THEN <camions prohibés>" pourrait être incluse dans le calcul. En plus d'introduire une capacité de raisonnement dans les modèles, une telle approche permettrait d'introduire des données non-quantitatives dans l'établissement des itinéraires.

Cette possibilité d'inclure des variables non-quantitatives pourrait également ouvrir la porte à de nouveaux modèles de calcul d'itinéraires plus flexibles, voire même basés sur la perception de l'espace par l'être humain. COUCLELIS (1986) et MARK (1993) ont démontré que, chez l'humain, l'orientation dans l'espace et dans les réseaux répond à des connaissances procédurales ainsi qu'à un contexte perceptuel qui s'apparente à une métaphore de l'espace (MARK, 1993).

Dans ce type de carte mentale, les individus vont associer des repères géographiques à une action quelconque. Par exemple, il peut s'agir d'une procédure du type "*au centre d'achat je dois effectuer un virage à gauche puis, à la prochaine intersection avec des feux de circulation, un virage à droite*". Un tel modèle mental repose, pour une bonne part, sur des données

verbales ou qualitatives qui ne se prêtent que difficilement à une traduction quantitative. Les modèles de calcul d'itinéraires traditionnels fondés sur la distance ou l'impédance sont de toute évidence inaptes à exprimer une telle complexité.

2.3.1 L'impédance et le transport des matières dangereuses

Le domaine du transport des matières dangereuses étant multidisciplinaire, les solutions aux problèmes impliquent généralement un chevauchement des sciences sociales, environnementales et mathématiques. Un grand nombre de facteurs doivent être considérés pour minimiser les risques encourus par la population sans pénaliser indûment les transporteurs. La recherche d'un itinéraire optimum est donc un compromis entre les risques encourus par la population et les coûts de transports élevés que peut entraîner un détour de la marchandise.

Certaines variables, comme la population pouvant potentiellement être affectée lors d'un accident ou les pertes monétaires encourues à la suite d'un accident sont relativement faciles à quantifier. Par contre, le concept de vulnérabilité de la population, certains critères sociaux, l'utilisation du sol le long des tronçons de même que certaines lois des transports sont beaucoup plus difficiles à quantifier et à pondérer, donc à traduire sous forme d'impédance.

Malgré ces inconvénients, l'approche probabiliste est demeurée couramment employée en transport des matières dangereuse, en partie due à l'influence des ingénieurs aux traditions mathématiques bien établies et en partie en raison de la facilité d'associer impédance et probabilité d'occurrence d'un accident pour un tronçon donné. Plus la probabilité d'accident

pour un tronçon sera élevée, moins il y aura de chance que ce tronçon soit retenu dans le calcul de l'itinéraire.

Or, comme l'ont mentionné GATRELL et VINCENT (1991), les risques technologiques sont relativement rares et de nature imprévisible. L'emploi des probabilités pour quantifier ce risque est ardu, spécialement pour les approches basées sur la fréquence relative des accidents (GATRELL et VINCENT, 1991, p. 152). Les probabilités obtenues sont généralement extrêmement faibles et à toutes fins identiques pour la plupart des tronçons, ce qui enlève tout pouvoir discriminant à l'impédance. A titre d'exemple, HARWOOD *et al* (1993) présentent un modèle probabiliste dans lequel la probabilité qu'un accident avec déversement se produise varie de 0.04 à 0.48 (selon le type de route) par million-de-camions par kilomètre. (HARWOOD *et al*, 1993, p. 196). Redistribué sur une base d'un camion par kilomètre de route on obtient des probabilités qui varient de 0.00000004 à 0.00000048 par tronçon d'un kilomètre.

Un autre inconvénient de l'approche probabiliste réside dans sa nécessité d'avoir une définition complète du problème. Comme l'indique FEDRA (1989, p 242):

"The classical methods of operations research and control engineering, which require a complete and quantitative definition of the problem from the outset, are certainly insufficient."
(FEDRA, 1989, p. 242)

Pour que les résultats soient valables, il faut que l'on ait en mains les données complètes (probabilité d'accident, de déversement, de vulnérabilité, etc.) pour chaque tronçon en vue d'établir les risques segment par segment. Dans une situation réelle, il est plutôt rare que l'on possède toutes les données nécessaires à ce calcul. On doit alors généraliser l'information

disponible à l'ensemble du réseau, ce qui veut dire que plusieurs tronçons auront une valeur d'impédance identique.

2.3.2 Les modèles d'itinéraires à objectif unique

Bien qu'un grand nombre de facteurs doivent être considérés pour minimiser les risques lors du transport des matières dangereuses, certains modèles ont été développés en ne tenant compte que d'un nombre très restreint de facteurs. Les modèles de GLICKMAN (1983), IVANCIE (1984) et PIJAWKA (1985) définissent le risque lié au transport comme étant la population exposée le long de l'itinéraire. Dans ces modèles, l'impédance de chaque tronçon est calculée en multipliant la population adjacente au tronçon par sa longueur.

Ces modèles prennent pour acquis que le risque du transport augmente si la population affectée augmente. L'avantage majeur de ce type de modèle est leur grande simplicité. Ils exigent peu de données hormis les décomptes de population généralement assez faciles à obtenir. Par contre, ils manquent de subtilité à bien des égards.

Tout d'abord, l'ensemble de la population est considéré comme étant à risque égal, ce qui n'est évidemment pas le cas. Ils ne tiennent pas compte de l'utilisation du sol que l'on retrouve de part et d'autre des tronçons. Aucune nuance n'est donc faite sur le type de construction et le type de population qui s'y trouve. Enfin, le seul objectif de ce type de modèle est de minimiser les risques sans tenir compte des facteurs économiques. L'itinéraire choisi peut donc être démesurément long pour minimiser les risques sans aucun compromis sur les coûts engendrés

par un tel détour. Toutefois, ces modèles ont le mérite d'avoir ouvert la porte à d'autres modèles en mesure de tenir compte d'un plus grand nombre de critères.

2.3.3 Les modèles probabilistes à évaluation des risques

L'analyse et l'estimation des risques occupe une place centrale dans le transport des matières dangereuses depuis plus de vingt ans. Dès 1971, le *National Transportation Safety Board* des États-Unis insistait sur la nécessité de développer des modèles d'itinéraires basés sur l'évaluation des risques (LIST *et al*, 1991, p. 101). Plusieurs des techniques analytiques développées pendant cette période ont pour origine les modèles probabilistes mis au point par l'industrie de l'énergie nucléaire.

A la fin des années soixante-dix, ANG *et al* (1979) ont été les premiers à suggérer une méthodologie d'évaluation des risques incorporant une approche systématique et synthétique des facteurs causant les risques. Leur apport principal consistait en la décomposition en étapes distinctes de l'analyse de risque, comme procèdent les ingénieurs de l'industrie nucléaire pour établir les scénarios d'accidents. Toute une série d'étapes doivent être franchies pour parvenir à déterminer le risque pour chaque tronçon, risque qui s'exprime sous forme de probabilité (impédance). Plusieurs chercheurs ont repris la suggestion de ANG *et al* pour en améliorer la méthodologie, notamment VAN AERDE *et al* (1989).

Le modèle RISKMOD développé par VAN AERDE *et al* est typique de l'approche à évaluation des risques et comporte cinq étapes dans l'établissement du risque lié au transport des matières dangereuses (figure 2.7). Chaque tronçon du réseau doit être évalué en fonction de

ces cinq phases qui sont suffisamment structurées pour être traduites sous forme de programme informatique, ce qui en assure l'automatisation. La première étape de la démarche

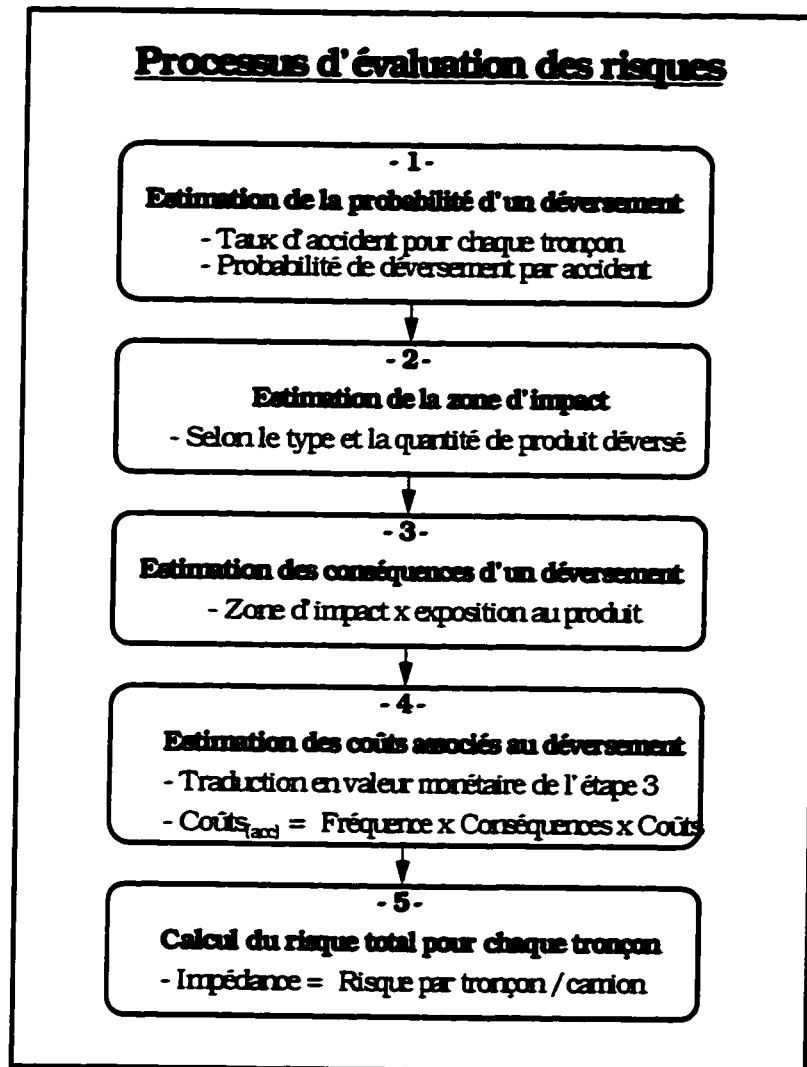


Figure 2.7: Processus d'évaluation des risques
D'après VAN AERDE *et al* (1989)

consiste à évaluer la probabilité qu'un déversement se produise pour un tronçon donné. Cette probabilité se définit comme étant le produit de la probabilité d'un accident multiplié par la probabilité conditionnelle qu'un déversement se produise:

$$P_{(\text{déversement})} = P_{(\text{accident})} \times P_{(\text{déversement} | \text{accident})}$$

A l'étape suivante, on procède à l'estimation de la zone potentiellement affectée par un déversement. Il s'agit d'évaluer l'étendue, en fonction de divers scénarios de déversement et en tenant compte du type de produit (gaz, explosif, liquide, etc) et de la quantité déversée, de la zone d'impact.

La troisième étape de l'évaluation concerne l'estimation des conséquences d'un déversement. En considérant la zone d'impact et le degré d'exposition à la substance déversée, le nombre de gens à être évacués, le nombre de blessés, le nombre de morts ainsi que l'impact sur l'environnement sont estimés. A ce stade, les résultats obtenus peuvent s'écarter sensiblement de la réalité, surtout en milieu urbain, puisqu'il est très difficile de chiffrer une population qui se déplace constamment en raison des migrations urbaines quotidiennes (aires de travail, de repos, de loisirs, etc).

La phase d'estimation des coûts associés au déversement est probablement l'étape la plus discutable de ce type de modèle. Cette étape consiste à traduire en valeurs monétaires les conséquences d'un déversement puisque le risque final s'exprime en coûts pour la société. Parmi ces estimations, certaines comme les dommages aux propriétés ou les coûts d'évacuation par personne sont relativement faciles à calculer pour chaque tronçon. Par contre, l'évaluation des coûts relatifs aux dommages à la faune aquatique (déversement dans une rivière) nous apparaît être de nature spéculative puisqu'il s'agit d'une approximation subjective faite par les utilisateurs du modèle.

Selon nous, là où le modèle présente sa faiblesse la plus importante, c'est dans son estimation (nécessaire) des coûts associés à la perte de vie humaine ainsi qu'aux blessures:

"[...] given a population exposure density of 500 people/km² and a hypothetical fatality value of 1,000,000\$ per life, as shown in Table 5."

- VAN AERDE *et al*, (1989), p. 332.

Bien que les auteurs indiquent qu'il s'agit d'une valeur hypothétique et en principe identique pour tous les individus, une telle approche ne pourrait être implantée dans une situation réelle de gestion du transport des matières dangereuses. Il y a fort à parier que la population en général réagirait négativement au monnayage d'une vie humaine surtout en raison de la possibilité d'introduire une catégorisation de la population³. Il faut cependant admettre que cette approche peut être utile aux transporteurs de matières dangereuses ainsi qu'à leurs assureurs pour établir le montant approximatif des poursuites légales en cas d'accidents avec perte de vie.

Finalement, à la dernière étape du modèle à évaluation des risques, on calcule l'impédance de chaque tronçon de route à partir des critères établis aux étapes précédentes ainsi qu'à partir des données spécifiques à chaque tronçon (taux d'accidents, décompte de population, nombre de passages de camions, etc.). Par la suite, le chemin optimum minimisant les risques est calculé à partir de l'algorithme standard de la théorie des graphes.

³Ce type d'approche qui consiste à associer une valeur monétaire aux pertes humaines tire son origine de l'approche "ingénieuriste" de l'industrie nucléaire (les auteurs du modèle sont ingénieurs). On peut supposer que les scénarios développés pour cette industrie devaient considérer les éventuelles pertes humaines et les coûts des poursuites légales que les familles des défunts intenteraient contre l'industrie fautive.

Les modèles à évaluation des risques présentent des avantages majeurs comparativement aux modèles à objectif unique. Ils tiennent compte de plusieurs facteurs et sont plus conformes à la réalité du transport des matières dangereuses. Cependant, la recherche de l'itinéraire optimum se fait uniquement sur l'impédance, la valeur unique issue du calcul du processus d'évaluation des risques. L'itinéraire choisi est donc déterminé à partir d'un objectif unique: la minimisation de l'impédance. Il faut donc se tourner vers les modèles à objectifs multiples pour introduire la notion de compromis entre les risques et les coûts de transport dans l'établissement des itinéraires.

2.3.4 Modèles à objectifs multiples

SHOBRY (1981) et ROBBINS (1981) ont été les premiers à présenter des modèles de calcul d'itinéraires à objectifs multiples pour le transport des matières dangereuses (LIST *et al*, 1991). Typiquement, ces modèles vont déterminer des itinéraires qui présentent un compromis entre deux objectifs. Il peut s'agir de minimiser simultanément les coûts de transport et les risques liés au transport des matières dangereuses. La méthode consiste à attribuer deux valeurs d'impédance par tronçon, une pour chaque objectif, et d'utiliser un algorithme du plus court chemin modifié pour obtenir une série de solutions optimales de Pareto (c'est-à-dire une série de solutions non-inférieures ou non-dominantes).

Supposons un modèle de transport dans lequel on désire minimiser simultanément les coûts de transport et les risques liés aux matières dangereuses. Si l'on représente les coûts de transport par $C_{(coûts)}$ et les risques par $R_{(risques)}$, on peut alors définir la somme des coûts engendrés pour traverser chaque tronçon du réseau par:

$$C_{(coûts)} = \left(\sum_i \sum_j k_{ij} X_{ij} \right)$$

où: (i,j) = tronçon de i à j .
 K_{ij} = Coûts du transport pour le tronçon (i,j)
 X_{ij} = 1 si le tronçon (i,j) fait partie de l'itinéraire
 0 si le tronçon (i,j) ne fait pas partie de l'itinéraire

(D'après CURRENT *et al*, 1990)

De la même manière on peut définir la somme des risques potentiels pour le transport des matières dangereuses:

$$R_{(risques)} = \left(\sum_i \sum_j d_{ij} X_{ij} \right)$$

où: (i,j) = tronçon de i à j .
 d_{ij} = Risque associé au tronçon (i,j)
 X_{ij} = 1 si le tronçon (i,j) fait partie de l'itinéraire
 0 si le tronçon (i,j) ne fait pas partie de l'itinéraire

(D'après CURRENT *et al*, 1990)

Le problème consiste alors à minimiser simultanément, lors du calcul de l'itinéraire, les valeurs de $C_{(coûts)}$ et de $R_{(risques)}$.

Minimiser CR = ($C_{(coûts)}$, $R_{(risques)}$)

Une série d'itinéraires présentant une solution non-inférieure sera alors déterminée. Chacune de ces solutions représentera un compromis plus ou moins grand en regard des deux objectifs (figure 2.8). Par exemple, sur la figure 2.8, le point "A" représente un itinéraire qui minimise les

risques mais à un coût de transport élevé. Par contre, l'itinéraire "C" minimise les coûts mais est plus risqué. Dans cet exemple, l'itinéraire "B" est un compromis entre les coûts de transport et les risques liés aux matières dangereuses.

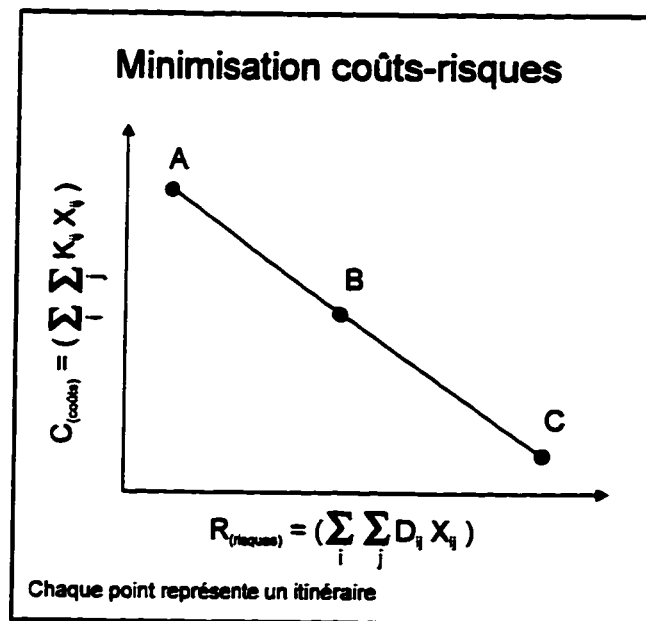


Figure 2.8: Minimisation coûts-risques dans un modèle à objectifs multiples. (D'après CURRENT et al, 1990, p. 189)

En plus de leur flexibilité, ces modèles ont également pour avantage d'offrir comme solution non pas un mais plusieurs itinéraires (dépendamment du nombre de solutions non-inférieures possibles) présentant des alternatives en terme compromis sur les objectifs. Les responsables de la sécurité publique ainsi que les transporteurs sont alors en mesure d'étudier différents scénarios de transport et de faire un choix qui satisfasse les deux parties.

Cependant, l'emploi de deux ou plusieurs valeurs d'impédance par tronçon implique un plus grand effort de pondération des tronçons. La plupart de ces modèles utilisent la méthodologie d'évaluation des risques avec tous les inconvénients que cela comporte (voir section 2.3.3). La solution de compromis n'élimine pas également le problème intrinsèque de l'impédance qui coupe l'accès direct aux variables et à toute possibilité de raisonnement sur celles-ci.

Les modèles à objectifs multiples ont aussi pour inconvénient de ne pas donner d'information explicite sur la nature du compromis entre les différents objectifs. Ce compromis peut être quantifié (pondéré) mais en aucun cas on ne peut savoir ce qui a été sacrifié dans le choix d'un itinéraire. Pour parvenir à un équilibre entre deux objectifs, une série d'éléments mineurs ou un seul élément majeur peuvent avoir été sacrifiés pour en arriver au même résultat global. Dans le cas du transport des matières dangereuses, cette zone grise du modèle peut conduire à des décisions "douteuses" pour certains tronçons de l'itinéraire.

Enfin, comme l'indique CURRENT *et al* (1990), le modèle à objectifs multiples mène parfois à un cul-de-sac lorsqu'appliqué à des réseaux complexes. Dans ces conditions, le problème devient intraitable avec cette approche puisque dans les pires conditions le nombre de solutions non-inférieures croît exponentiellement par rapport au nombre de noeuds du réseau (CURRENT *et al*, 1990, p. 188).

2.3.5 Modèles en temps réel

Les mesures de sécurité en matière de transport des substances dangereuses sont souvent abordées dans l'optique des règlements et du contrôle comme par exemple la détermination

préalable d'un itinéraire que les transporteurs devront respecter. Il faut toutefois admettre que l'itinéraire le plus sécuritaire ou le règlement le plus strict n'ont pas de valeur si aucun moyen n'est mis de l'avant pour les faire respecter. C'est en partie ce raisonnement qui est à l'origine des systèmes de navigation en temps réel pour le transport des matières dangereuses.

Mieux connus sous le vocable anglais de "*Vehicle Tracking Systems*" ou de "*Car Navigation Systems*", cette technologie est apparue commercialement dès la fin des années quatre-vingt (WHITE, 1991, p. 116). Conçue au départ pour guider les véhicules automobiles en région urbaine, l'idée d'adapter la technologie au domaine des matières dangereuses a pris corps bien que l'infrastructure nécessaire à son implantation en limite fortement l'application. L'approche consiste à faire le repérage continu du véhicule par surveillance satellitaire puis de relayer à une station de guidage et de répartition la position du véhicule dont le trajet sera calculé à partir de toute une série de paramètres routiers (BEROGGI, 1994).

Les quatre éléments composant les modèles en temps réel sont le système de positionnement global (Global Positioning System), la station de guidage et son répartiteur, le réseau de communication véhicule-station de guidage et le système de collecte des paramètres routiers (figure 2.9). A chaque instant, un satellite en orbite effectue le repérage du camion dont la position est retransmise à la station de guidage. De la station de guidage, un répartiteur s'occupe d'établir l'itinéraire le plus sécuritaire et retransmet au conducteur du véhicule, sous forme de carte numérique ou verbalement, la route à prendre pour se rendre à destination.

En plus de prendre en compte les critères de sécurité comme le font les autres modèles de calcul d'itinéraires, le modèle en temps réel incorpore également des paramètres routiers tels que les embouteillages, les conditions météorologiques, les routes fermées, etc. Il s'agit donc

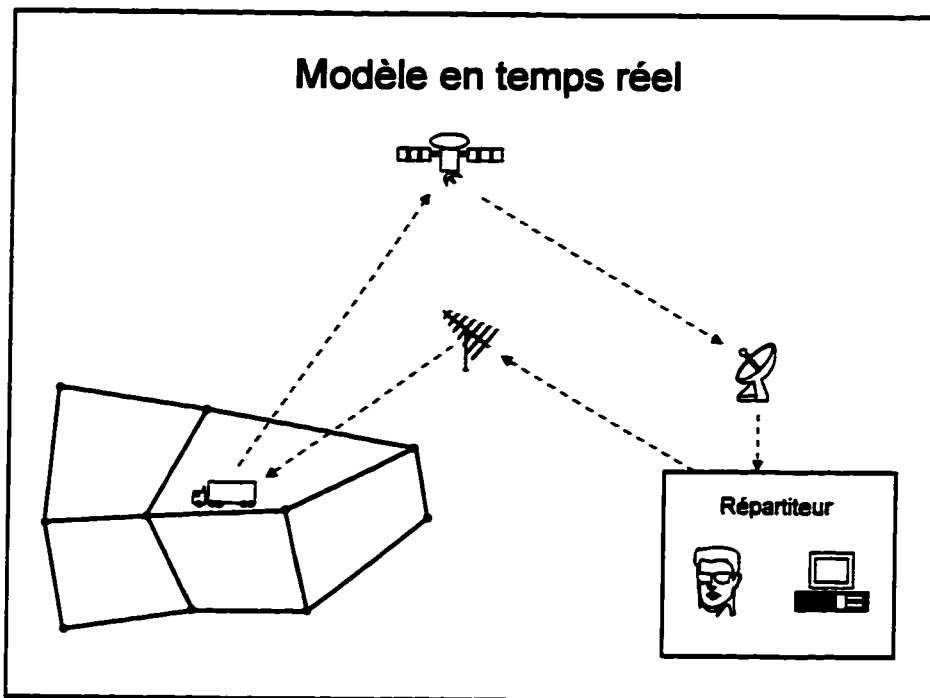


Figure 2.9: Système de navigation en temps réel
(D'après BEROGGI, 1994, p. 509)

d'une approche dynamique de guidage et de repérage puisqu'en tout temps le répartiteur est en mesure de savoir si le transporteur respecte l'itinéraire qui lui est assigné et qu'en retour le même transporteur est pris en charge si un problème d'itinéraire se pose. Ce dernier aspect est important puisqu'en milieu urbain il est plutôt fréquent que des rues soient fermées pour diverses raisons, rendant ainsi l'itinéraire de départ inadéquat.

Un des avantages des modèles en temps réel est qu'ils permettent un contrôle plus efficace du flux de matières dangereuses. Il est possible de savoir à tout moment où sont les cargaisons et si les transporteurs respectent les itinéraires établis. Le contrôle en temps réel est également plus fidèle à la réalité de terrain puisqu'il s'ajuste aux conditions changeantes du trafic routier contrairement aux modèles "traditionnels" où plusieurs itinéraires alternatifs doivent être développés pour pallier aux imprévus (routes bloquées, construction, etc)

Il faut cependant admettre que les modèles en temps réel ne sont pas très répandus à l'heure actuelle surtout en raison des infrastructures qu'ils nécessitent (lien satellitaire, station de guidage, répartiteur, réseau de télécommunication, station de collecte des données routières, etc.) et des coûts élevés que cela comporte. Si l'on considère que pour la seule province de l'Ontario il y avait, en 1988, 158 000 camions de plus de 4500kg; équiper l'ensemble de cette flotte de système de guidage-repérage serait tout à fait prohibitif pour les transporteurs (GORYS, 1991, p. 20).

Un autre facteur limitant l'application des systèmes en temps réel est le problème de la collecte des paramètres routiers et la mise-à-jour continue des ceux-ci. La gestion à l'échelle régionale ou provinciale des capteurs de données et des bases de données en temps réel, la validation des données ainsi que le calcul de centaines d'itinéraires en simultané implique le traitement de millions d'observations donc l'emploi de systèmes informatiques très performants et coûteux. Le traitement en temps réel nous apparaît donc comme une solution à long terme, solution qui exige quand même une amélioration des modèles actuels de calcul d'itinéraires.

2.4 Élément de solution: Les systèmes experts

Le chapitre 2 a permis de démontrer que le concept d'impédance impose une limite aux modèles de calcul d'itinéraires actuels. L'impédance restreint la capacité de discrimination des modèles en assimilant à une valeur unique l'ensemble des variables décrivant les tronçons routiers. Le raisonnement du type "IF-THEN" sur les variables, raisonnement à la base de tout système de prise de décision, est incompatible avec cette approche puisqu'il doit s'effectuer sur l'ensemble des variables. Ces lacunes sont particulièrement évidentes dans le cas du transport des matières dangereuses où une foule de facteurs quantitatifs et non-quantitatifs doivent être pris en compte pour établir un itinéraire sécuritaire.

Dans le but de pallier à ces problèmes, nous croyons que l'ajout d'un système expert aux modèles de calcul d'itinéraires conventionnels serait un moyen efficace d'inclure une approche déductive dans le calcul d'itinéraire et de contourner les limitations de l'impédance. Un tel système expert, couplé à un système d'information géographique formerait un système de support à la décision à référence spatiale (SSDRS) efficace en analyse spatiale et plus particulièrement en gestion du transport des matières dangereuses. Il permettrait d'établir des itinéraires à partir d'une base de connaissances sur laquelle seraient appliquées des règles d'inférence simulant la démarche d'un expert humain. Enfin, en plus d'établir un itinéraire, ce type de système serait en mesure d'expliquer à l'utilisateur le raisonnement ayant conduit à la sélection d'un itinéraire donné, ce qui constitue un avantage certain dans le cas du transport des matières dangereuses.

Chapitre III

La théorie des graphes

"Lorsque l'homme se meut, la géométrie le guide."

- Léon Lalanne, 1863

C'est le mathématicien suisse Leonhard Euler, né à Bâle en 1707, qui fut le premier en 1736 à rédiger un mémoire sur la théorie des graphes, mémoire qui devait être publié dans le recueil des publications de l'Académie des Sciences de Saint-Petersbourg (ORE, 1970, p. 21). Euler était reconnu comme étant un scientifique de génie dont la production très diversifiée (200 volumes) touchait à la fois la théologie, les langues, la médecine, les mathématiques, la physique et l'astronomie. Ce n'est que par distractions mathématiques qu'il aborda le domaine de la théorie des graphes en publiant son mémoire sur le *Problème des ponts de Koenigsberg*.

La ville de Koenigsberg (aujourd'hui Kaliningrad) était à l'époque divisée en plusieurs quartiers reliés entre eux par sept ponts. Euler avait présenté le problème de la façon suivante: Est-il

possible de trouver un chemin qui permet de visiter chaque quartier de la ville et de revenir à son point de départ en traversant chaque pont une seule fois? (ORE, 1970, p. 22). Euler devait démontrer mathématiquement dans son mémoire que dans le cas de Koenigsberg, la ville ne pouvait être parcourue en un cycle simple; c'est-à-dire que quel que soit le sommet choisi comme point de départ, il était impossible d'y revenir sans emprunter au moins une fois un tronçon de route déjà emprunté. Dans la suite de son mémoire, Euler exposait le problème des cycles de façon plus générale en indiquant les conditions nécessaires à l'obtention d'un cycle simple.

Suite aux travaux d'Euler, la théorie des graphes est devenue partie intégrante du domaine des mathématiques discrètes et, dès le XIX^{ième} siècle, la théorie était appliquée aux circuits électriques et aux diagrammes moléculaires. La théorie des graphes constitue d'ailleurs l'un des domaines des mathématiques qui présente le plus d'applications pratiques. On l'utilise entre autres en architecture, en chimie, en électricité, en génétique, dans les systèmes de contrôle, en linguistique et bien entendu en géographie. Ce n'est cependant qu'avec les travaux de KÖNIG en 1936 qu'une première synthèse du domaine de la théorie des graphes est devenue accessible.

3.1 Théorie des graphes et géographie

Ce n'est pas un géographe mais un mathématicien qui fut le premier à présenter une application de la théorie des graphes en géographie et plus particulièrement en cartographie. CAYLEY (1879) publia dans les *Actes de la Société Royale de Géographie* un article sur le problème des quatre couleurs. Ce problème est fondé sur l'hypothèse que toute carte peut être

coloriée avec quatre teintes seulement sans qu'il y ait juxtaposition d'une même teinte. Plusieurs décennies se sont écoulées avant que la théorie des graphes ne refasse surface en géographie physique avec les travaux de HORTON (1945) sur les ratios de bifurcation des systèmes fluviaux, travaux auxquels STRAHLER (1952) devait apporter des modifications.

L'emploi de la théorie des graphes en tant qu'outil de modélisation a cependant fait son apparition en géographie au début des années soixante avec la publication de GARRISON sur la connectivité du réseau autoroutier des États-Unis (GARRISON, 1960). GARRISON y présentait une approche nouvelle fondée sur une étude morphologique et topologique des réseaux. L'année suivante, NYSTUEN et DACEY publiaient le résultat de leurs travaux sur la structure des régions nodales américaines dans lesquels ils mettaient à profit les matrices de connexité de la théorie des graphes comme méthode de régionalisation (NYSTUEN et DACEY, 1961). Le plaidoyer de BUNGE sur la nécessité d'une géographie théorique où la théorie des graphes devait occuper une place majeure a par la suite été publié en 1962 (BUNGE, 1962). KANSKY en 1963 s'attaqua au domaine de la simulation des réseaux et développa, à partir des indices de la théorie des graphes, un modèle mathématique de l'expansion du réseau ferroviaire sicilien (KANSKY, 1963).

Au milieu des années soixante, HAGGETT (1965) publia *Locational Analysis in Human Geography* dans lequel deux chapitres traitaient spécifiquement de la théorie des graphes. HAGGETT récidivait avec CHORLEY en 1967 pour lancer *Models in Geography* dans lequel le chapitre sur les modèles de réseaux en géographie démontrait les capacités de modélisation de la théorie (CHORLEY et HAGGETT, 1967). Ce n'est toutefois qu'à la fin des années

soixante qu'une oeuvre importante dans le domaine des analyses de réseau en géographie *Network Analysis in Geography* fut publiée par le duo HAGGETT et CHORLEY (HAGGETT et CHORLEY, 1969). Avec ce livre, HAGGETT et CHORLEY ont pour mérite d'avoir présenté le potentiel de la théorie des graphes à la fois en géographie physique et humaine. Leur étude détaillée de la théorie y était appuyée tout autant par des exemples d'hydrologie que de transport, abordant à la fois l'équilibre des flux dans un réseau hydrographique et son équivalent dans un réseau routier.

Suite à l'abondance de publications des années soixante, les années soixante-dix et quatre-vingt ont été plus calme ou du moins sans progrès majeur dans le domaine. Le champ de l'analyse spatiale prend un essor notable à cette époque et la théorie des graphes y est généralement présentée comme une approche parmi d'autres. C'est le cas de ABLER *et al* (1971). DALTON *et al* (1973) ainsi que BRIGGS (1974) font exception en publiant deux ouvrages portant principalement sur les indices d'accessibilité et de connectivité des réseaux. Au début des années quatre-vingt, SELKIRK aborde le sujet dans le contexte de l'analyse spatiale, un peu comme ABLER *et al* l'avaient fait la décennie précédente (SELKIRK, 1982). WERNER est l'un des rares, au cours des années quatre-vingt, à s'attaquer au sujet de l'optimisation des réseaux et de la modélisation des flux via la programmation linéaire (WERNER, 1985).

Bien que le nombre de publications portant sur la théorie des graphes ait considérablement diminué au cours des années quatre-vingt, son emploi en géographie n'est pas disparu pour autant. Les notions de topologie qui sont à la base des structures de données spatiales

vectérielles des systèmes d'information géographique (SIG) en sont un exemple puisqu'elles proviennent directement de la théorie des graphes. Bon nombre de fonctions des SIG sont également redevables du développement de la théorie des graphes des dernières décennies.

3.2 Les graphes en tant que modèle spatial

Quiconque a eu la chance de survoler à haute altitude une région urbaine a pu constater la complexité des réseaux de transport et de communication qui tissent le territoire. Aborder un phénomène aussi complexe exige d'un modèle qu'il simplifie le système tout en préservant l'information pertinente.

Les graphes répondent à ce besoin en réduisant un réseau en un ensemble de noeuds et de liens auxquels les postulats de géométrie et de topologie peuvent être appliqués (figure 3.1).

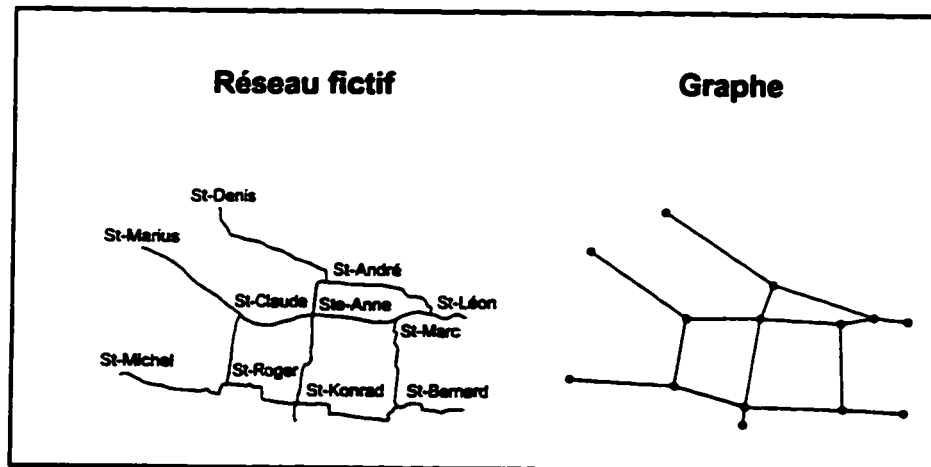


Figure 3.1: Simplification d'un réseau à l'aide d'un graphe

Bien qu'en apparence fort simple (voir simpliste), les graphes ont servis de base aux mathématiciens pour développer toute une gamme d'outils mathématiques qui permettent d'étudier les réseaux dans leurs composantes individuelles ou dans leur ensemble, rendant

ainsi possible les comparaisons. Les graphes sont donc des modèles qui représentent deux phénomènes spatiaux fondamentaux, les localisations et les interrelations, phénomènes qui ouvrent la porte à une analyse *statique* ou *dynamique* des réseaux.

3.2.1 Éléments de base d'un graphe

Un graphe est un *diagramme composé de points appelés **sommets** reliés entre eux par des lignes appelées **arêtes***. Chaque arête rejoint nécessairement deux sommets (WILSON et WATKINS, 1990, p. 8). Lorsque trois arêtes ou plus sont reliées entre elles pour former un polygone, le graphe renferme alors une ***région*** (figure 3.2).

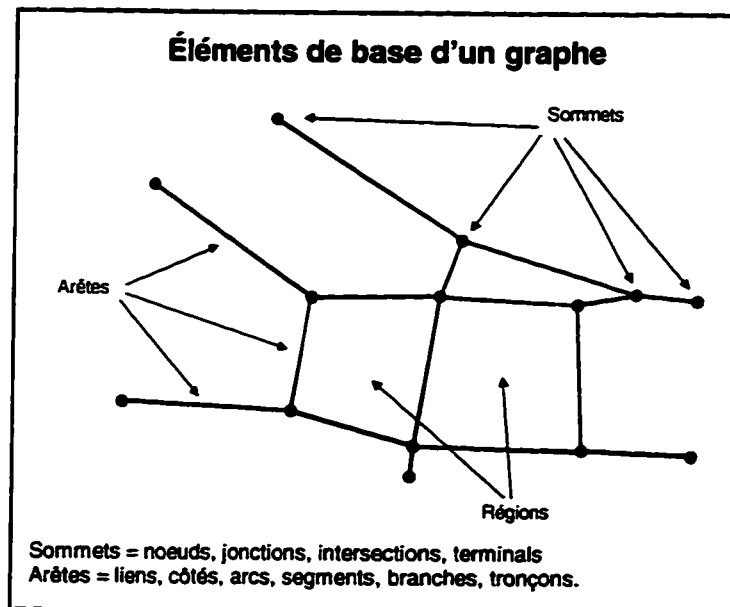


Figure 3.2: Éléments de base d'un graphe

La théorie des graphes ne possède pas de terminologie standardisée, ce qui fait que certains auteurs parleront de ***noeuds***, de ***jonctions***, d'***intersections*** ou de ***terminals*** pour décrire les sommets d'un graphe. De la même manière, les termes ***liens***, ***côtés***, ***arcs***, ***segments***, ***branches***, ou ***tronçons*** sont utilisés comme synonyme du mot arête. Tous ces termes sont

acceptables à condition de les utiliser de façon consistante. En se référant à ce vocabulaire, il est possible de définir cinq propriétés de base des graphes:

- ❑ Un graphe possède toujours un nombre fini de sommets
- ❑ Chaque arête se termine par un sommet à ses extrémités
- ❑ Chaque arête relie deux sommets différents
- ❑ Un seul arc peut relier une même paire de sommets¹
- ❑ Aucune distinction n'est faite entre un sommet d'origine et un sommet de destination²

Les trois premiers énoncés doivent être respectés en toutes circonstances. Un graphe ne peut, par exemple, avoir un nombre indéfini de noeuds ou posséder des arêtes formant des boucles (même noeud de départ et d'arrivée). Cependant, certaines situations vont exiger d'outrepasser les deux derniers énoncés. Ces situations se produisent régulièrement dans le cas des réseaux routiers.

Bien que les graphes permettent d'illustrer, un peu comme une carte, des réseaux complexes, ce mode graphique est inadéquat pour aborder les aspects mathématiques de la théorie. Une définition symbolique et plus formelle a donc été établie.

Un graphe **G** consiste en un ensemble fini et non-vide de sommets noté **N** et un ensemble d'arêtes reliant ces sommets noté **A**. La notation **G(N,A)** représente un graphe **G** composé de **N** sommets et de **A** arêtes. De même, la notation **(i,j)** indique une arête reliant le sommet $i \in N$ avec le sommet $j \in N$.

(TEODOROVIC, 1986)

¹ Dans la pratique, cette règle peut être transgressée puisque dans certaines circonstances il est possible que plus d'une route rejoignent les mêmes villes.

² Cette règle ne s'applique que dans le cas d'un graphe non-dirigé, c'est-à-dire sans sens unique. Dans la pratique, l'origine et la destination sont considérées dans les modèles.

3.2.2 La typologie des graphes

Une des fonctions des graphes est de qualifier les réseaux qu'ils représentent de façon à reconnaître facilement certaines situations données. Pour effectuer ce survol de la typologie des réseaux, nous nous sommes référés à WILSON et WATKINS (1990), HAGGETT et CHORLEY (1969) ainsi qu'à WILSON (1972). La présente section ne constitue pas un inventaire exhaustif de la typologie des graphes, inventaire qui pourrait s'étendre sur plusieurs pages. Le lecteur peut se référer aux auteurs ci-haut mentionnés pour plus de détails.

La forme la plus élémentaire de graphe est le **graphe nul**, c'est-à-dire un graphe composé de "N" sommets et de 0 (zéro) arête $G(N,0)$ (figure 3.3a). Un graphe est **fini** si chacun de ses sommets est relié aux autres par une arête, sans qu'il s'agisse nécessairement d'un lien direct (figure 3.3b). Lorsque chaque sommet d'un graphe est relié à tous les autres sommets par un

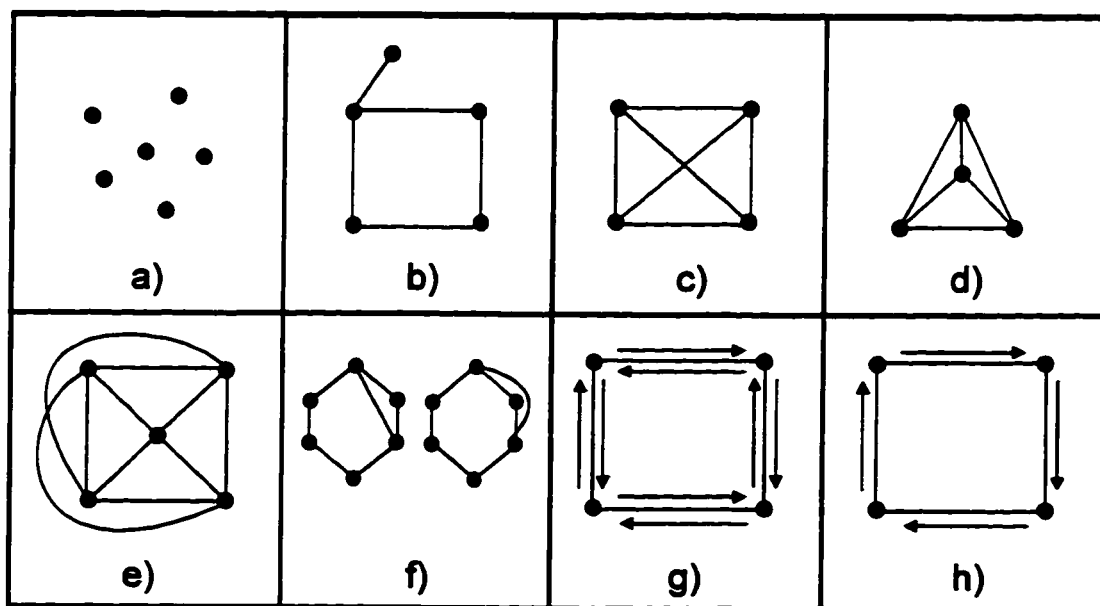


Figure 3.3: Types de graphe

lien direct, on a alors un **graphe complet** (figure 3.3c). Un graphe complet à "N" sommets possède toujours $\frac{1}{2}N(N-1)$ arêtes (WILSON, 1972, p. 16).

Un graphe ou un réseau sera **planaire** s'il peut être tracé dans un plan de façon à ce que ses arêtes n'aient pas d'autres intersections que les sommets (ORE, 1970, p. 9) (figure 3.3d). Un réseau hydrographique ou un réseau de routes de campagne sont des exemples de graphes planaires. Par opposition, les graphes **non-planaires** ne peuvent être représentés dans un plan sans qu'il y ait création d'intersections ne correspondant à aucun sommet (figure 3.3e). Un réseau d'autoroutes en milieu urbain est un exemple typique de graphe non-planaire. Les autoroutes se croisent par superposition (les viaducs) sans qu'il n'y ait de carrefour.

Les graphes étant une illustration d'une structure topologique, c'est-à-dire d'un ensemble d'interrelations, c'est principalement l'agencement du réseau qu'ils reflètent. On dira donc de deux graphes qu'ils sont **isomorphes** s'ils représentent la même situation topologique, c'est-à-dire les mêmes liens entre les sommets (figure 3.3f). Les graphes permettent également de décrire les flux dans les réseaux. Un graphe dont chacune des arêtes peut être parcourue dans les deux directions est un graphe **non-orienté**, contrairement à un graphe ne contenant que des sens uniques que l'on appellera alors **digraphe** (figure 3.3g et 3.3h). Dans une situation réelle de transport, il est commun de rencontrer des **graphes mixtes**, c'est-à-dire des graphes qui possèdent des sens uniques et des arêtes bidirectionnelles.

Les géographes ont également développé leur propre classification des graphes en fonction des phénomènes généralement étudiés en géographie. Cette classification que l'on doit à

HAGGETT et CHORLEY (1969, p. 7) est essentiellement basée sur une approche topologique (figure 3.4). Outre la division planaire et non-planaire, on y retrouve les chemins, les arbres et les circuits. Cette classification couvre donc principalement les systèmes linéaires et arborescents.

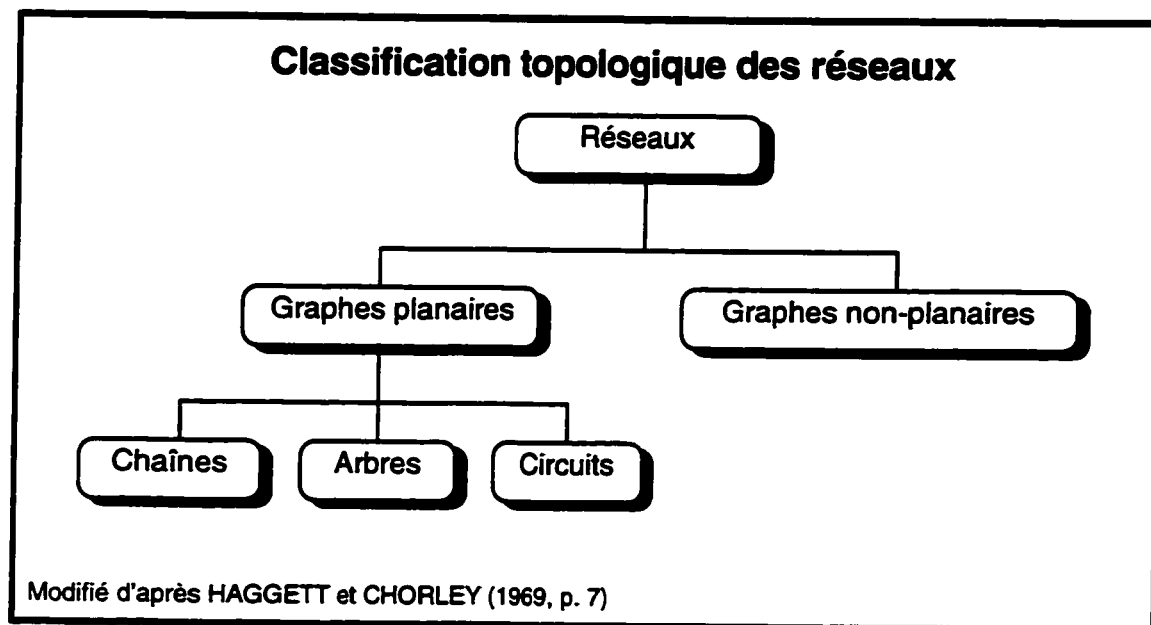


Figure 3.4 Classification topologique des réseaux

Dans cet arrangement, les chaînes font référence à des itinéraires ne passant qu'une seule fois par sommet, les arbres aux réseaux hydrographiques et les circuits à des trajets dont les points de départ et d'arrivée sont les mêmes.

3.3 La représentation matricielle des graphes

Bien qu'il soit possible de calculer manuellement certains indices topologiques d'un réseau à partir de sa représentation graphique, l'opération devient vite fastidieuse voire impossible dans

le cas d'un réseau complexe. La solution à la manipulation des réseaux complexes est leur transformation sous forme matricielle. La reproduction sous cette forme donne l'accès aux outils d'algèbre matricielle et facilite le traitement informatique des réseaux.

3.3.1 La matrice d'incidence

La représentation d'un réseau sous forme de matrice d'incidence (ou d'adjacence) consiste à réduire le graphe en une matrice binaire (0 ou 1) dont chaque cellule indiquera la présence d'un lien direct entre deux sommets par un 1 et son absence par un 0 (figure 3.5). En pratique, il

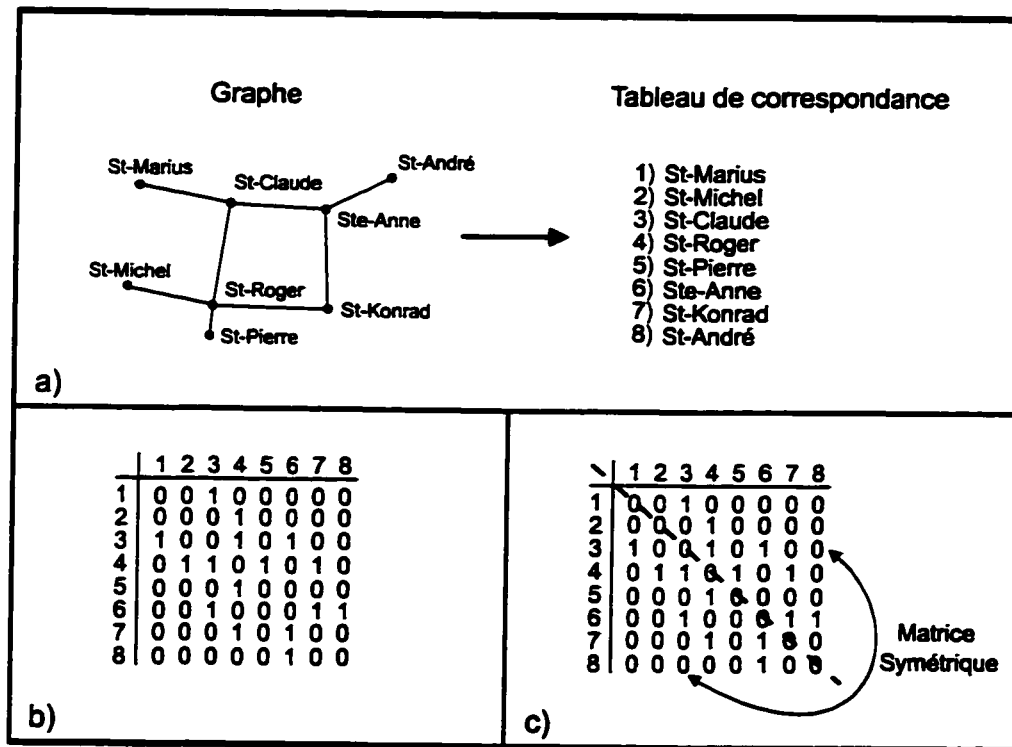


Figure 3.5: Représentation matricielle des réseaux

faut dans un premier temps créer un tableau de correspondance à "N" entrés ou chaque sommet se verra associer un identificateur entier qui servira d'indice en ligne et en colonne pour la matrice (figure 3.5a).

Une fois le tableau de correspondance créé, on procède à la construction de la matrice d'incidence avec les identificateurs comme indice. Pour un réseau de "N" sommets on créera une matrice carrée de "N" lignes par "N" colonnes. L'existence d'un lien direct entre 2 sommets sera marquée par un 1 et son absence par 0 (figure 3.5b). La diagonale de la matrice est initialisée à zéro puisqu'un sommet ne peut avoir de lien avec lui-même.

Dans le cas d'un graphe non-dirigé, la matrice d'incidence sera symétrique de part et d'autre de la diagonale étant donné que la présence du couple (i,j) est nécessairement accompagnée de celle du couple (j,i) (figure 3.5c). Dans le cas d'un digraphe, comme un réseau hydrographique par exemple, la matrice sera asymétrique puisque la présence du couple (i,j) interdit celle du couple (j,i) (à moins que la rivière ne coule dans les deux directions à la fois!).

Les matrices d'incidence sont très couramment employées en analyse de réseaux. Elles servent au calcul de bon nombre d'indices d'accessibilité et de connectivité. En représentant sous forme binaire les réseaux, les matrices d'incidence permettent de comparer des réseaux différents, de quantifier leur structure interne et d'en évaluer l'efficacité. Par exemple, en utilisant le calcul matriciel, il est possible de connaître facilement le nombre de chemins indirects entre deux sommets. Il suffit d'élever à la puissance 2,3,4... "p" la matrice d'incidence pour découvrir le nombre chemins de 2,3,4... ou "p" arêtes reliant les sommets du réseau. L'emploi des matrices d'incidence présente cependant un inconvénient. Dans certaines situations, il ne suffit pas de savoir qu'il existe un lien entre deux sommets, il faut pouvoir quantifier ce lien.

3.3.2 La matrice de distance

La matrice de distance d'un réseau se construit de façon similaire à la matrice d'incidence. La différence entre les deux matrices réside dans l'information stockée. Alors que la matrice d'incidence n'indique que la présence/absence de lien entre deux sommets, la matrice de distance indique la distance à parcourir pour relier ces deux sommets. La matrice de distance $d(i,j)$ contiendra, pour chaque couple (i,j) ayant un *lien direct*, la distance les séparant. Pour tout couple (i,j) n'ayant pas de lien direct, une distance infinie ∞ sera stockée dans la matrice (figure 3.6a).

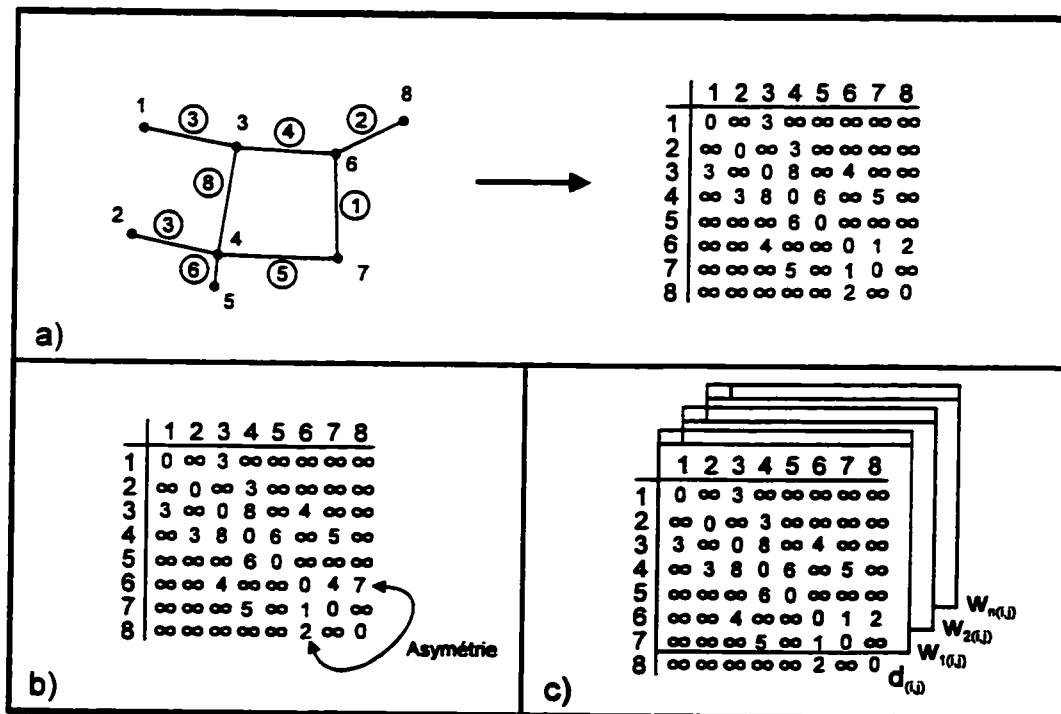


Figure 3.6: Matrice de distance

La matrice de distance est nécessairement symétrique $d(i,j) = d(j,i)$ puisqu'elle doit satisfaire à la règle d'inégalité du triangle $d_{ij} \leq d_{ik} + d_{kj}$ (BUCKLEY et HARARY, 1990, p. 127). Ce type de matrice est aussi fort utile lorsque l'on est en présence d'un *réseau pondéré* (à impédance).

Les valeurs de distance de la matrice peuvent être remplacées par des valeurs d'impédance tel que discuté au chapitre 2 (figure 3.6b). La matrice de distance peut alors être symétrique ou asymétrique puisque l'impédance du couple $w(i,j)$ peut être différente du couple $w(j,i)$. A titre d'exemple, la modélisation du trafic routier aux heures de pointe suppose l'emploi d'une matrice asymétrique en raison du déséquilibre des flux routiers. Il est également possible de modéliser des phénomènes plus complexes à l'aide de ce type de matrice en transformant le tableau en cube (figure 3.6c). L'ajout d'une troisième dimension permet ainsi d'associer plusieurs valeurs d'impédance (w_1 à w_n) ou d'attributs à chaque arête du réseau. C'est ce type d'approche qui est utilisée pour les modèles d'itinéraires à objectifs multiples.

3.3.3 Les listes d'adjacence

La représentation d'un réseau à l'aide d'une matrice offre l'avantage de donner accès directement à l'information contenu dans chaque cellule de la grille. A l'aide des indices ligne/colonne on peut savoir immédiatement s'il existe un lien direct entre deux sommets donnés. L'inconvénient de cette méthode est que la taille des matrices devient rapidement un problème. Un réseau de 100 sommets peut être représenté avec une matrice de 10 000 cellules, tandis qu'un réseau de 1000 sommets en exigera 1 000 000. Si l'on considère également que cette matrice est composée, en majeure partie, de cellules avec valeur 0, la redondance est évidente et l'approche inefficace. De plus, la mémoire des ordinateurs étant limitée, on ne peut employer plusieurs matrices de grande dimension sans que tôt ou tard les limites de la machine ne soient atteintes. Pour pallier ce problème, la représentation des réseaux par liste d'adjacence a été développée.

Dans le cas des listes d'adjacence, seuls les liens existant entre deux sommets sont stockés dans la matrice qui prend alors la forme d'une séquence de liens (figure 3.7). La redondance est ainsi réduite et un réseau de "N" sommets, peut être représenté avec un maximum de $(N^2 - N)/2$ cellules (BUCKLEY et HARARY, 1990, p. 226).

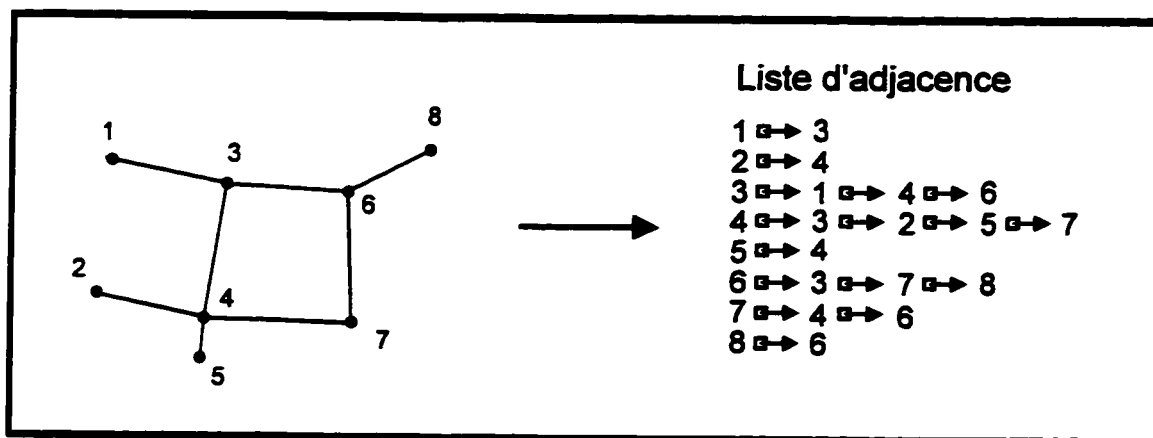


Figure 3.7: Représentation par liste d'adjacence

Par contre, le stockage séquentiel des données oblige à parcourir l'ensemble de la liste pour obtenir la valeur d'un lien particulier, ce qui peut être effectué directement dans une matrice à partir des indices. De plus, l'ordre d'apparition des arêtes dans les listes d'adjacence affecte l'ordre dans lequel les arêtes seront traitées par les algorithmes. Les gains réalisés en supprimant la redondance peuvent donc, dans certains cas extrêmes, être annihilés par ces inconvénients. La solution idéale consiste donc en une approche qui élimine la redondance tout en gardant un accès direct aux données.

La représentation matricielle a tout de même pour mérite d'avoir ouvert la porte à deux types d'approches dans l'étude des réseaux. Une première approche dont l'objectif est d'analyser la

morphologie et la topologie des réseaux à partir d'une série d'indices d'accessibilité, de centralité, de connectivité et de forme. La deuxième approche concerne surtout les modèles d'optimisation, c'est-à-dire la recherche de chemins optimaux, de localisation optimum et de maximisation des flux dans les réseaux. Le reste du chapitre sera donc consacré à ces deux approches.

3.4 Les indices morphologiques et topologiques

Même si le propos de cette thèse, à savoir le calcul d'itinéraire, fait partie des modèles d'optimisation, il est impossible d'aborder l'usage de la théorie des graphes en géographie sans en présenter brièvement les indices morphologiques et topologiques. Ces indices, qui sont à la base de la théorie des graphes, ont été développés pour permettre la comparaison des réseaux en ce qui concerne leur connectivité, leur forme et leur accessibilité. Ils sont aussi fort utiles pour aborder l'évolution temporelle des réseaux en quantifiant leur degré de développement. Ces indices peuvent se diviser en trois grands groupes avec tout d'abord les indices se rapportant à la connectivité, puis ceux décrivant la forme et le degré de développement et enfin les indices d'accessibilité et de centralité.

3.4.1 Les indices de connectivité

Les indices de connectivité considèrent principalement les propriétés topologiques des réseaux et ne tiennent compte dans leur calcul que du nombre de sommets, d'arêtes, de régions et de sous-graphes. Ils donnent donc une description globale des réseaux, un peu comme le font les statistiques descriptives en statistiques. Plusieurs indices de connectivité découlent de l'équation qu'EULER publia en 1736, équation qui ne fait que mettre en lumière les rapports

existant entre les différentes composantes d'un graphe. Selon EULER, dans un graphe planaire, on retrouve toujours le rapport suivant:

$$S + R = A + 2$$

où: S = nombre de sommets
 R = nombre de régions (ou polygones)
 A = nombre d'arêtes

Un premier indice issu de cette équation est le *nombre cyclomatique* (μ), qui permet de calculer le nombre de circuit³ ou de régions dans un graphe (SELKIRK, 1982, p. 162).

$$\mu = A - S + P$$

où: S = nombre de sommets
 A = nombre d'arêtes
 P = nombre de sous-graphes

Le nombre cyclomatique est particulièrement utile en géographie, notamment dans les systèmes d'information géographique vectoriel et en cartographie puisqu'il permet de connaître très facilement le nombre de polygones que renferme une carte et de faciliter la construction de la topologie (figure 3.8).

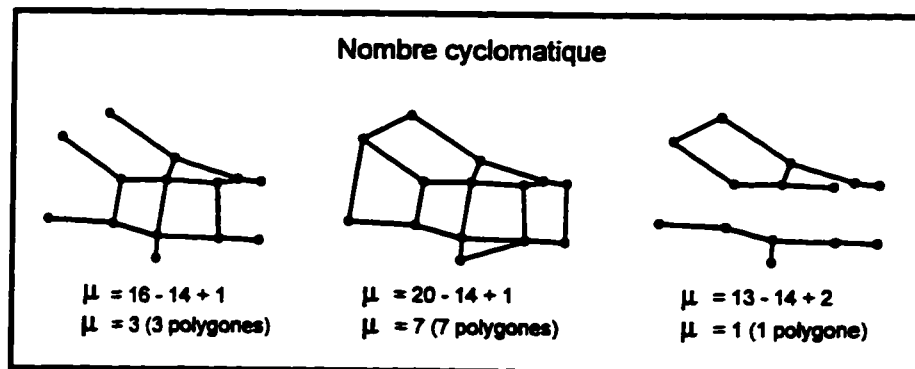


Figure 3.8: Nombre cyclomatique et nombre de polygones

³Un circuit est un itinéraire qui débute et se termine au même sommet d'un réseau pour former une région fermée (polygone).

Parmi les autres indices de connectivité, on retrouve l'*indice alpha* (α) qui permet de calculer le nombre maximum de circuits possibles dans un réseau en fonction du nombre de sommets qu'il contient ainsi que les *indices beta* (β) et *gamma* (γ) qui expriment le degré de connectivité absolu et relatif (SELKIRK, 1982, p. 166).

$$\alpha = \frac{\mu}{(2S - 5)}$$

$$\beta = \frac{A}{S}$$

$$\gamma = \frac{A}{3(S - 2)}$$

3.4.2 Les indices d'accessibilité et de centralité

En géographie, l'usage de la théorie des graphes est souvent associé à la mesure de l'efficacité des réseaux routiers ou de communications. Il est donc primordial d'avoir des outils de mesure de la centralité et de l'accessibilité des sommets qui correspondent le plus souvent à des villes. On retrouve généralement un lien direct entre l'accessibilité et la centralité d'une ville et son dynamisme économique. La *nodalité* est un premier outil de mesure de l'accessibilité d'un sommet.

La nodalité se calcule pour chaque sommet d'un réseau et correspond à la somme des arêtes qui y sont reliées. On distingue la nodalité des intrants (Indegree) et la nodalité des extrants (Outdegree) (KNOKE et KUKLINSKI, 1982, p. 45). La première équivaut à la somme des lignes de la matrice d'incidence et la seconde à la somme des colonnes. En plus de déterminer

l'accessibilité d'un sommet, le rapport intrants/extrants permet de définir le "type de relation" qu'entretient un sommet avec ses voisins. En substituant des valeurs de flux à la matrice d'incidence, on peut établir un score pour chaque sommet (ville)⁴.

$$Nodalité_{(intrants)} = \sum_{i=1}^N Z_{ij}$$

$$Nodalité_{(extrants)} = \sum_{j=1}^N Z_{ij}$$

où: N = nombre de sommets
 Z_{ij} = matrice d'incidence ou de pondération

La position qu'occupe un sommet dans un réseau peut également prendre beaucoup d'importance dans certaines situations. Plusieurs indices ont pour objectif de mesurer la centralité d'un sommet dans l'ensemble du réseau. Ces indices nécessitent la construction d'une matrice de distance, distance qui doit être le plus court chemin entre chaque sommet du réseau.

L'*indice de Shimmel* indique la distance minimale qui doit être parcourue, à partir d'un sommet donné, pour rejoindre tous les autres sommets. Plus la valeur obtenue est faible, meilleur est la position du sommet dans le réseau (SELKIRK, 1982, p. 170). L'indice de Shimmel se calcul en effectuant la sommation, en ligne, de la matrice de distance. De même ordre, le *nombre de König* (nombre associé) correspond à la valeur la plus élevée, en ligne, de la matrice de distance, pour chaque sommet. Enfin, l'*indice de centralité de Leavitt* est une variante de

⁴Ce type d'indice appliqué à un réseau de communication tel qu'INTERNET permettrait de connaître l'apport de chaque site au réseau global. On peut supposer que certains sites ne font que soutirer de l'information alors que d'autres sont essentiellement des pourvoyeurs.

l'indice de Shimmel avec comme pondération la somme en ligne et en colonne de la matrice de distance (DAUPHINÉ, 1979).

$$Shimmel = \sum_{j=1}^N d(X_i, X_j)$$

$$Leavitt = \frac{\sum_{i=1}^N \sum_{j=1}^N d(X_i, X_j)}{\sum_{i=1}^N d(X_i, X_j)}$$

où: N = nombre de sommets
 $d(X_i, X_j)$ = le chemin le plus court entre X_i et X_j

3.4.3 Les indices de forme et de développement

Les indices de forme et de développement servent principalement à caractériser la configuration de l'ensemble d'un réseau. Ils ont leur utilité surtout dans l'étude temporelle de l'évolution des réseaux, lorsqu'il est nécessaire de comparer divers stades de développement d'un même réseau. Parmi ces indices on retrouve l'*indice pi* (π) qui est le rapport entre la longueur total du réseau et son *diamètre*, ce dernier étant le chemin le plus court entre les deux sommets les plus éloignés du réseau (DAUPHINÉ, 1979). L'*indice de densité* correspond au rapport entre le nombre d'arêtes d'un réseau et le nombre maximal d'arêtes possible (KNOKE et KUKLINSKI, 1982, p. 45).

Contrairement à l'indice de densité, l'*indice de cohésion* mesure le degré de relations réciproques (de liens bidirectionnels) dans un réseau et tient compte de l'orientation des arêtes.

Un réseau très fortement orienté (un réseau hydrographique par exemple) aura un indice de cohésion de 0.0 tandis qu'un graphe parfaitement bidirectionnel obtiendra un indice de 1.0 (KNOKE et KUKLINSKI, 1982, p. 50). Finalement, l'*indice de dispersion* se calcule en faisant la somme de tous les plus courts chemins pour tout le réseau, c'est-à-dire la somme de tous les indices Shimmel (SELKIRK, 1982, p. 171). L'ajout d'un sommet en un point stratégique d'un réseau peut parfois avoir un impact important sur la valeur de l'indice de dispersion, ce qui permet de mettre en valeur la relation topologique qui existe entre chaque noeud et le reste du réseau.

$$\pi = \frac{\sum_{i=1}^N \sum_{j=1}^N d(X_i, X_j)}{\text{Diamètre}}$$

où: N = nombre de sommets
 Diamètre = Diamètre du réseau
 $d(X_i, X_j)$ = longueur de l'arête entre X_i et X_j (0 si aucun lien)

$$\text{Densité} = \frac{\sum_{i=1}^N \sum_{j=1}^N Z_{ij}}{N^2 - N}$$

où: N = nombre de sommets
 Z_{ij} = matrice d'incidence

$$\text{Cohesion} = \frac{\sum_{i=1}^N \sum_{j=1}^N (Z_{ij} + Z_{ji})}{\frac{N^2 - N}{2}}$$

où: N = nombre de sommets
 $Z_{ij} + Z_{ji}$ = Matrice d'incidence (Égale 1 si les deux éléments égalent à 1, autrement 0)

$$Dispersion = \sum_{i=1}^N \sum_{j=1}^N d(X_i, X_j)$$

où: N = nombre de sommets
 $d(X_i, X_j)$ = le chemin le plus court entre X_i et X_j

3.5 Modèles d'optimisation: Affectations et circuits

Les modèles d'optimisation en analyse de réseau ont été développés pour la plupart dans le cadre de la recherche opérationnelle étant donné leur grand potentiel d'application sur le terrain. L'objectif de ces modèles étant de fournir une solution optimum pour un problème d'équilibre donné, l'intérêt économique y joue un rôle important. Parmi les différents modèles d'optimisation, on retrouve l'affectation de ressources mieux connue sous le terme de "location-allocation study", l'établissement de circuits (itinéraires de livraison) et le calcul du chemin optimum. Ces trois types de modèles nécessitent une approche de programmation dynamique contrairement aux indices de la théorie des graphes qui aborde les réseaux de façon statique.

3.5.1 L'affectation de ressources

Les modèles d'affectation de ressources ont fait leur apparition en géographie au cours des années soixante, notamment avec YEATES (1963) concernant le problème de la délimitation des districts scolaires, mais c'est à HITCHCOCK (1941) que l'on doit la présentation du modèle classique de programmation linéaire comme solution à ce problème. L'affectation de ressources implique habituellement quatre composantes: **1)** des ressources ou des services à être distribués dans un réseau **2)** Un ou plusieurs point(s) de distribution de ces ressources ou services **3)** Un objectif (par exemple livrer la marchandise ou les services) **4)** Un ensemble de

contraintes qui limitent la capacité à atteindre l'objectif (distance maximale pouvant être parcourue, délai d'intervention, etc).

Dans le contexte des réseaux, l'affectation de ressources consiste à assigner chaque arête du réseau à un point de distribution de marchandises ou de services de façon à minimiser les contraintes (par exemple la distance à parcourir). Supposons l'exemple de la distribution du courrier en milieu urbain. L'objectif consiste à localiser dans le quartier un nombre " p_j " de boîtes postales, au coûts " f_j " par boîte, pour desservir " n " clients tout en tenant compte de la distance à parcourir " d_{ij} " par les clients pour se rendre à la boîte postale ainsi que des contraintes du parcours " w_{ij} " (par exemple l'absence de trottoir). L'équation d'affectation prendra alors la forme suivante:

$$MIN_{(A)} = \sum_{i=1}^n \sum_{j=1}^p w_{ij} \times d_{ij} + \sum_{j=1}^p f_j$$

(D'après MILLER, 1993, p. 534)

où: n = nombre de clients à desservir
 p = nombre de sites de distribution
 f_j = coûts d'implantation des sites de distribution
 d_{ij} = distance entre le client n_i et le site de distribution p_j
 w_{ij} = facteur de pondération de la distance d_{ij}

Il s'agit nettement d'un processus dynamique puisque l'emplacement d'une boîte aura nécessairement un impact sur la distance à parcourir pour atteindre les autres boîtes. D'une situation de départ généralement arbitraire, " p " boîtes sont distribuées dans le réseau et à chaque itération du calcul, la configuration n'est retenue que si elle permet de réduire la valeur de " A ". Le processus se répète jusqu'à ce que la solution optimum soit atteinte, c'est-à-dire que la valeur de " A " cesse de diminuer. A la fin du processus, on connaîtra le nombre de sites de

distribution nécessaires et leur position dans le réseau, la clientèle qui sera affectée à chacun des sites et pour chaque arête du réseau, le site de distribution auquel elle est rattachée (figure 3.9).

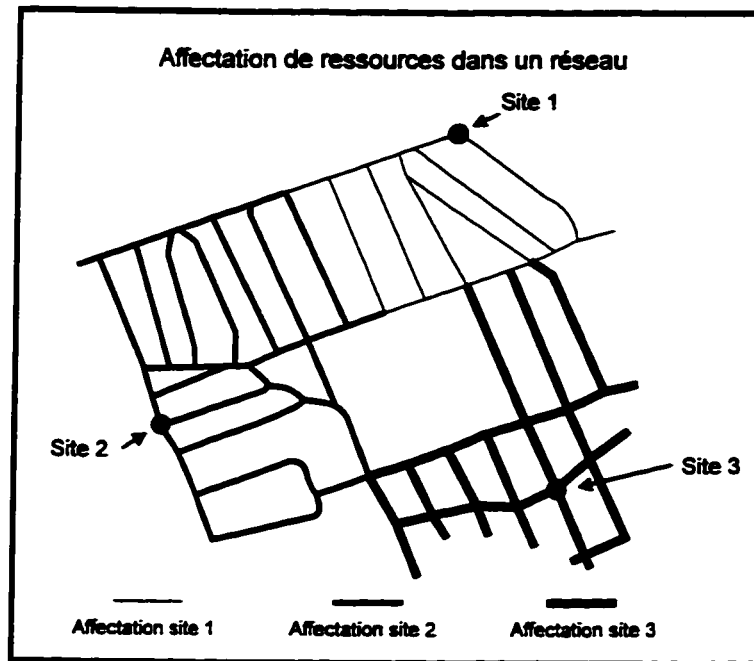


Figure 3.9: Affectation de ressources dans un réseau

Il existe plusieurs variantes au modèle d'affectation de ressources. Dans plusieurs cas, les sites de distribution sont connus et le traitement consiste à déterminer la section du réseau qui sera rattachée à chaque site à partir du calcul du plus court chemin. Par exemple, pour délimiter les zones de transport scolaire on calculera, à partir de la position de l'école, la partie du réseau qui se trouve à moins d'un kilomètre de marche par le plus court chemin. Au-delà de cette zone, le transport motorisé sera implanté. En recherche opérationnelle, l'affectation de ressources ne représente qu'une partie du problème de la distribution de marchandises ou de

services à la clientèle. L'autre aspect concerne l'établissement d'itinéraires optimaux de livraison.

3.5.2 Les circuits

Comme nous l'avons souligné en début de chapitre, la théorie des graphes est née d'un problème relevant du domaine des circuits. En démontrant qu'il n'était pas possible de parcourir la ville de Königsberg sans passer deux fois par le même pont, EULER fut le premier à s'attaquer à un problème courant aujourd'hui: l'établissement d'itinéraires optimaux de livraison. On appelle cycle ou circuit tout chemin dans un réseau qui débute et se termine au même point. On dira d'un réseau qu'il est *eulérien* s'il est possible de le parcourir et de revenir au point de départ en passant qu'une seule fois par chaque arête. Un graphe sera *hamiltonien* si l'on peut le parcourir en ne passant qu'une fois par chaque sommet (ORE, 1970, p. 29). Évidemment, ce ne sont pas tous les réseaux qui sont eulériens ou hamiltoniens. Pour établir des itinéraires qui tendent vers ces solutions parfaites, des modèles d'optimisation ont été développés. Dans le cas des circuits hamiltoniens, il s'agit du *problème du voyageur de commerce* (Traveling Salesman Problem TSP) et pour les circuits eulériens, du *problème du postier chinois* (Chinese Postman Problem CPP).⁵

Le modèle d'optimisation du voyageur de commerce consiste à définir un itinéraire optimum passant par "N" sommets (ou points de livraison) de façon à ce qu'en partant du sommet 1 on

⁵L'origine de ce nom remonte à la première publication sur le sujet faite par M.G. GUAN dans la revue *Chinese Mathematics* en 1962. GUAN y présentait le problème à partir de l'exemple d'un postier chinois désireux d'éviter les déplacements inutiles (BUCKLEY et HARARY, 1990, p. 272).

visite N-1 sommets pour revenir au sommet 1. Si l'on prend comme données de départ une matrice de distance symétrique, le coût d'un circuit "i" sera:

$$MIN \text{ Circuit}_{(H)} = \sum_{j=1}^{N-1} d_{i_j, i_{j+1}} + d_{i_N, i_1}$$

où: N = nombre de sommets
 (i_1, i_2, \dots, i_n) = permutation des entiers de 1 à N indiquant l'ordre dans lequel les sommets sont visités dans la $i^{\text{ème}}$ itération (circuit).

(D'après HAGGETT *et al*, 1977, p. 78)

L'inconvénient de cette approche est le nombre de circuits possibles à examiner ou si l'on préfère le nombre d'itérations potentiellement nécessaires pour parvenir à la solution optimum. Dans le cas d'une matrice symétrique, le nombre d'itérations s'établit à $\frac{1}{2}(N-1)!$, soit 5040 circuits hamiltoniens possibles pour un simple réseau de 15 sommets (HAGGETT *et al*, 1977, p. 78). Ce type d'approche exclue donc l'établissement de circuits ayant plus de quelques dizaines de points de livraison. Pour des problèmes plus complexes d'autres modèles ont été développés, notamment celui de SHEN LIN qui subdivise le réseau en sous-réseaux de taille réduite, et qui conduit à une approximation plutôt qu'à une solution optimum (ABLER *et al*, 1971, p. 276).

Le modèle du postier chinois est similaire au modèle du voyageur de commerce. La distance de chaque arête est pris en compte et à chaque itération, il y a permutation de la séquence des arêtes visitées. Le nombre de passages dans chaque arête, pour un circuit donné, est utilisé pour pondérer la distance et ainsi favoriser une solution optimum qui se rapproche le plus d'un circuit eulérien parfait.

$$MIN\ Circuit_{(eulerien)} = \sum_{A(i,j)=1}^{NA} S_{ij} \times d(X_i, X_j)$$

(D'après TEODOROVIC, 1986, p. 69)

où: NA = nombre d'arêtes dans le réseau
 $A(i,j)$ = Arête du sommet "i" au sommet "j"
 S_{ij} = nombre de passages dans l'arête (i,j) pour une itération donnée
 $d(X_i, X_j)$ = longueur de l'arête (i,j)

Qu'il s'agisse des modèles d'affectation, de l'établissement de circuits dans un réseau ou encore de certains indices morphologiques, tous ont en commun une caractéristique fondamentale: déterminer le plus court chemin ou le chemin optimum dans un réseau. L'algorithme du plus court chemin est certainement le plus utile en théorie des graphes et il occupe d'ailleurs une place importante dans cette thèse. Le reste du présent chapitre y sera donc consacré.

3.6 Modèles d'optimisation: le plus court chemin

L'algorithme du plus court chemin est probablement l'un des plus étudiés en analyse de réseau si l'on considère la très nombreuse littérature sur le sujet. Dès l'apparition des premiers ordinateurs, des routines de calcul de plus en plus sophistiquées sont apparues comme solution au problème du déplacement dans les réseaux complexes. Parmi les pionniers à s'être intéressés à ce problème, notons DIJKSTRA (1959), DANTZIG (1960) ainsi que FARBEY, LAND et MURCHLAND (1967). Plusieurs autres auteurs ont aussi contribué au domaine et le lecteur peut se référer à TEODOROVIC (1986), SHAPIRO (1987) et SEDGEWICK (1990) pour plus de détails. La revue systématique de la totalité des modèles dépassant le cadre de cette

recherche, nous avons donc procédé à un regroupement des approches pour en dégager l'essentiel et faciliter la présentation. La figure 3.10 résume cette classification.

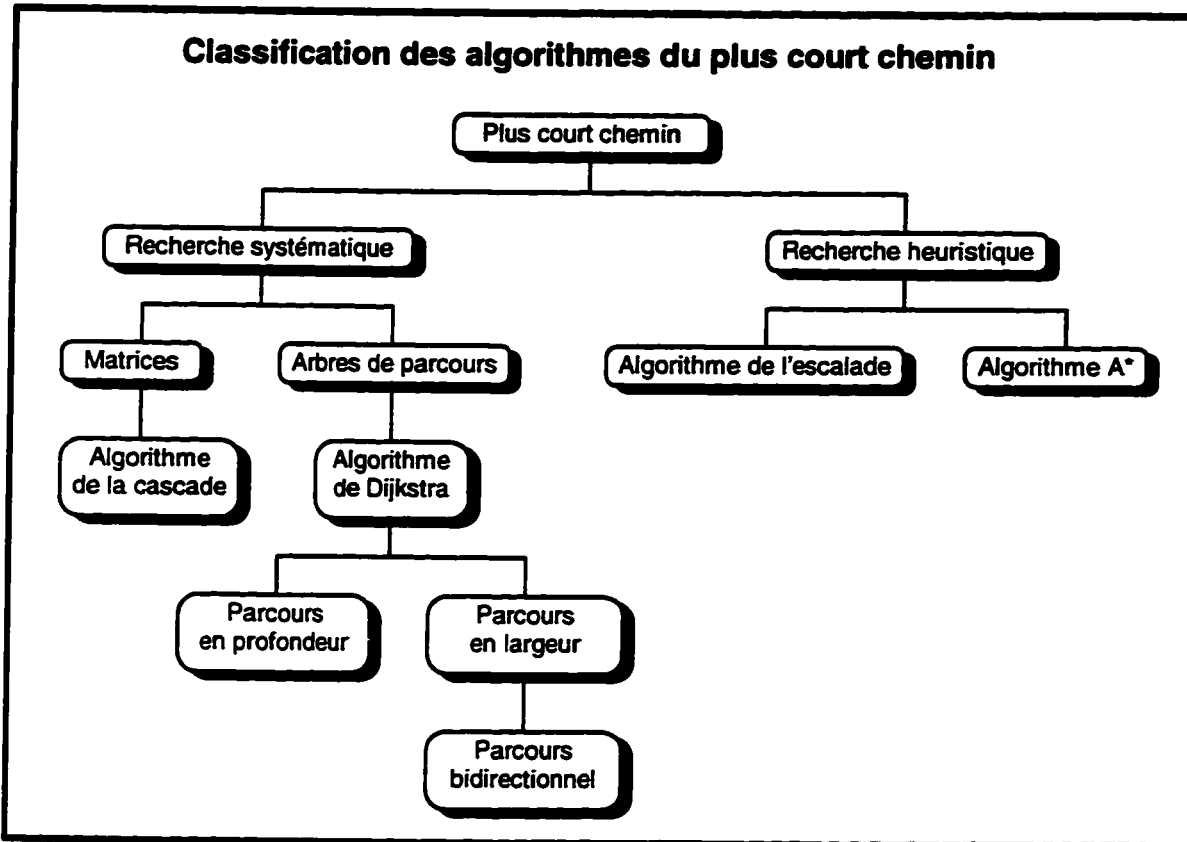


Figure 3.10: Classification des algorithmes du plus court chemin

On peut distinguer deux grandes approches dans la recherche du chemin le plus court dans un réseau. Une première approche qui opère selon une méthode de recherche systématique, chaque arête du réseau étant visitée selon un ordre préétabli. Cette approche se divise en deux grands types de modèles; les modèles matriciels avec notamment l'*algorithme de la cascade* et les modèles à *arbres de parcours* avec l'*algorithme de Dijkstra*. Pour ces derniers, les arbres peuvent être traversés en profondeur ou en largeur ou encore de façon bidirectionnelle. La deuxième approche introduit le concept de sélection des arêtes dans les algorithmes à partir

d'une *recherche heuristique simple (algorithme de l'escalade)* ou plus sophistiquée (*algorithme A**).

3.6.1 Plus court chemin: matrices et algorithme de la cascade

L'algorithme de la cascade est l'une des méthodes les plus simples de calculer le plus court chemin entre chaque paire de sommets d'un réseau. FARBEY *et al* (1967) en ont donné une bonne description, de même que HAGGETT et CHORLEY (1969). Comme son nom l'indique, l'algorithme de la cascade consiste à traverser la matrice de distance de départ en utilisant la diagonale comme pivot et à calculer par comparaisons successives le chemin le plus court. Le résultat final sera une matrice de distance dans laquelle chaque cellule renfermera la longueur du chemin le plus court pour chaque paire de sommets respectifs. Le fonctionnement général de l'algorithme est le suivant (figure 3.11).

D'une situation de départ suivante:

- Réseau orienté, mixte ou bidirectionnel (figure 3.11 a)
 - Matrice de distance contenant la longueur de chaque arête $d(i,j)$ (figure 3.11b)
 - L'absence d'arête entre deux sommets (i,j) est représentée dans la matrice par la valeur $d(i,j) = \infty$ (figure 3.11b)
 - La diagonale est remplie de zéros pour satisfaire la condition $d(i,j) = 0$ si $i = j$
1. Choisir successivement chacun des éléments de la diagonale comme pivot pour établir les axes de calcul. Établir deux axes perpendiculaires à partir du pivot (figure 3.11c).
 2. Sélectionner tour à tour chacun des éléments en ligne et en colonne de façon à ce que pour chaque $d(i,j)$, $i \neq j$ (figure 3.11c).
 3. Faire la somme successivement des éléments en ligne du pivot avec les éléments en colonne du pivot et remplacer les valeurs à leur intersection ligne-colonne si la somme est plus petite que la valeur actuelle (figure 3.11e)
 4. Répéter les opérations pour toute la matrice (figure 3.11f,g,h et l).

Le résultat final est une matrice des plus courts chemins élémentaires telle que démontrée à la figure 3.11i. Plusieurs indices de la théorie des graphes nécessitent ce type de matrice (indices de shimbel, de Leavitt, etc) pour leurs calculs. La grande simplicité de cet algorithme cache cependant plusieurs défauts majeurs.

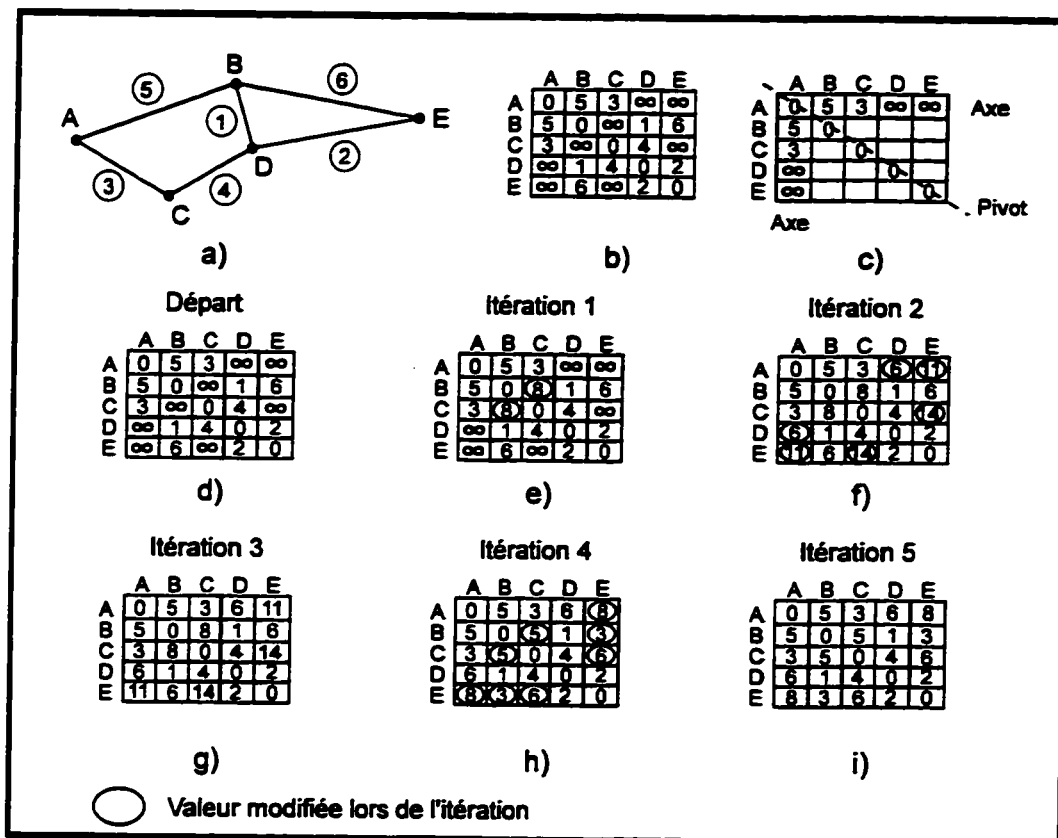


Figure 3.11: Algorithme de la cascade

Exception faite du calcul d'indices, il est plutôt rare qu'une application exige le calcul des plus courts chemins pour toutes les paires de sommets d'un réseau. Avec l'algorithme de la cascade, même si l'on désire connaître le chemin pour deux sommets particuliers, l'ensemble de la matrice doit être traité ce qui implique beaucoup de calculs inutiles. Bien que les algorithmes du plus court chemin exigent tous une somme considérable de calculs, la méthode de la cascade est particulièrement exigeante à cet égard. Dans sa version originelle, la

cascade nécessite N^3 ($N =$ nombre de sommets) opérations d'addition-comparaison-permutation pour solutionner un graphe. Pour un réseau de 1000 sommets, c'est 1,000,000,000 d'opérations, ce qui fait de la méthode de la cascade la méthode la moins efficace et la plus lente de toutes.

Un autre handicap de cette approche est qu'elle permet de trouver facilement la longueur du plus court chemin mais rend très difficile le stockage en mémoire de la séquence de sommets qui composent le parcours du chemin le plus court. On connaît donc la distance à parcourir... sans savoir par où passer! Hormis les applications de calcul d'indices, on doit donc se tourner vers une autre méthode plus efficace pour le calcul du chemin optimum. C'est dans cette optique que la méthode des arbres de parcours a été développée.

3.6.2 Plus court chemin: arbres de parcours

Pour palier aux déficiences de la méthode matricielle une alternative a été développée par DIJKSTRA (1959). L'algorithme de DIJKSTRA consiste à faire la recherche du plus court chemin entre deux sommets en progressant du sommet de départ vers les autres sommets tout en visitant les sommets adjacents de façon à construire un arbre de parcours. Cette approche est basée sur le fait qu'il est possible de représenter n'importe quel graphe par un arbre dont chacun des embranchements correspond à une route possible entre deux sommets. En d'autres termes, l'arbre de parcours est une autre forme de représentation graphique d'une matrice d'incidence ou d'une liste d'adjacence. Il suffit de parcourir cet arbre en minimisant la distance à parcourir à chaque embranchement pour que, tôt ou tard, le chemin le plus court entre deux sommets particuliers soit découvert. Contrairement à l'algorithme de la cascade qui

utilise une matrice d'incidence, les arbres de parcours opèrent à partir d'une liste d'adjacence.

Le fonctionnement général de l'algorithme est le suivant (figure 3.12):

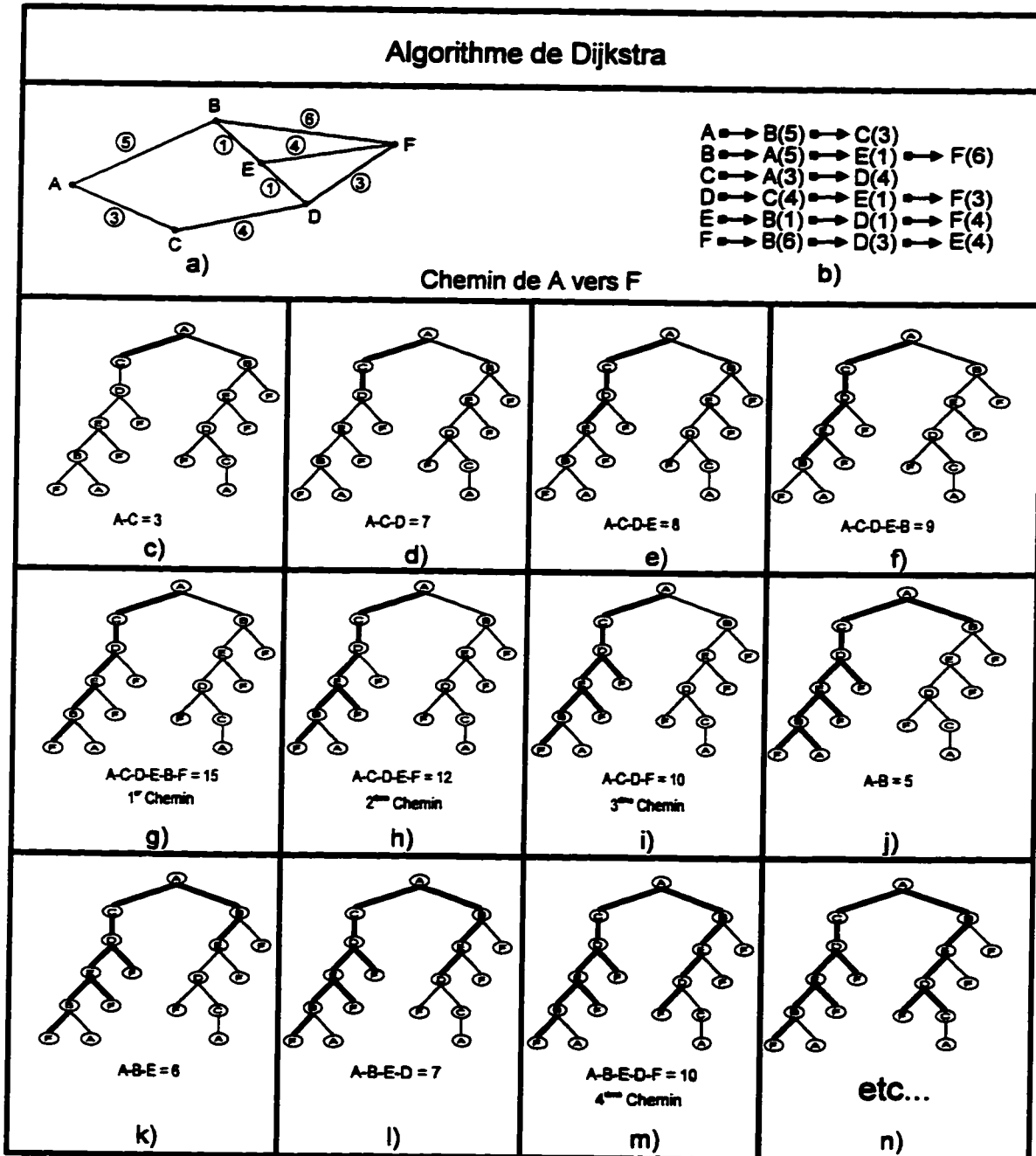


Figure 3.12: Arbres de parcours et algorithme de Dijkstra

D'une situation de départ suivante:

- Réseau orienté, mixte ou bidirectionnel (figure 3.12a)
 - Liste d'adjacence contenant la longueur de chaque arête $d(i,j)$ (figure 3.12b)
1. On initialise à l'infini (∞) la valeur de D_r qui servira à stocker la longueur du chemin le plus court entre le sommet de départ (A) et le sommet d'arrivée (F).
 2. On balaye la liste d'adjacence pour sélectionner un sommet adjacent au sommet "A" de façon à minimiser la distance à parcourir entre le sommet "A" et le prochain sommet: $\text{MIN}(D_r + d(i,j))$.
 3. Le sommet "C" ($A-C=3$) est retenu ($A-C=3 < A-B=5$) (figure 3.12b) et un premier embranchement est visité dans l'arbre de parcours (figure 3.12c)
 4. On balaye la liste d'adjacence à nouveau pour sélectionner un sommet adjacent au sommet "C" (sommet non-visité) de façon à minimiser la distance à parcourir (étape 2)
 5. Le sommet "D" ($A-C-D=7$) est retenu. Une deuxième branche de l'arbre est identifiée comme visitée (figure 3.12d).
 6. On procède selon les étapes 2, 3 et 4 pour la suite du parcours (figure 3.12e et 3.12f).
 7. A l'étape 3.12g, un premier chemin entre "A" et "F" est découvert. On stocke en mémoire la séquence de sommets qui composent le parcours (A-C-D-E-B-F) et la distance totale (15).
 8. On revient à l'embranchement supérieur immédiat dans l'arbre et on examine successivement les sommets non-visités.
 9. A l'étape 3.12h, un deuxième chemin entre "A" et "F" est trouvé (A-C-D-E-F). Sa longueur étant de 12 comparativement à 14 pour le chemin précédent, celui-ci devient le "nouveau chemin le plus court" entre A et F. On élimine le chemin précédent et on stocke en mémoire la séquence de sommets du nouveau chemin ainsi que sa longueur.
 10. On répète ce processus pour l'ensemble de l'arbre de parcours (figures 3.12i,,j,k,l,m,n) en conservant le chemin le plus court chaque fois qu'un nouveau chemin est découvert.

En procédant de la sorte et en parcourant la totalité de l'arbre, il est garanti que le chemin le plus court, s'il existe, sera découvert. Cependant, divers facteurs peuvent avoir un impact sur le déroulement de la méthode des arbres de parcours. Un premier facteur est l'ordre dans lequel la séquence des sommets est emmagasinée dans la liste d'adjacence. Cet ordre détermine la

structure de l'arbre, donc la séquence de recherche du chemin le plus court. Un deuxième facteur est la stratégie adoptée pour traverser l'arbre.

Dans la figure 3.12, la stratégie consistait à sélectionner une branche de l'arbre (par exemple la branche A-C, figure 3.12c) puis à traverser cette branche jusqu'à ce que le niveau inférieur de l'arbre soit atteint (branche B-F, figure 3.12g) puis à remonter d'un niveau et répéter la démarche. Dans ce type d'approche, on parle d'un *parcours en profondeur* (Depth-first), c'est-à-dire que l'on explore dès le début les niveaux inférieurs de l'arbre. L'avantage de cette méthode est que l'on est assuré d'avoir examiné toutes les solutions possibles avant de revenir à un niveau supérieur de l'arbre. Par contre, on peut parfois perdre beaucoup de temps à examiner des itinéraires longs et inutiles qui sont composés de très nombreuses arêtes.

La deuxième méthode de parcours des arbres est le *parcours en largeur* (Breadth-first). Contrairement au parcours en profondeur, le parcours en largeur visite tous les sommets d'un niveau avant de passer au niveau suivant de l'arbre (figure 3.13).

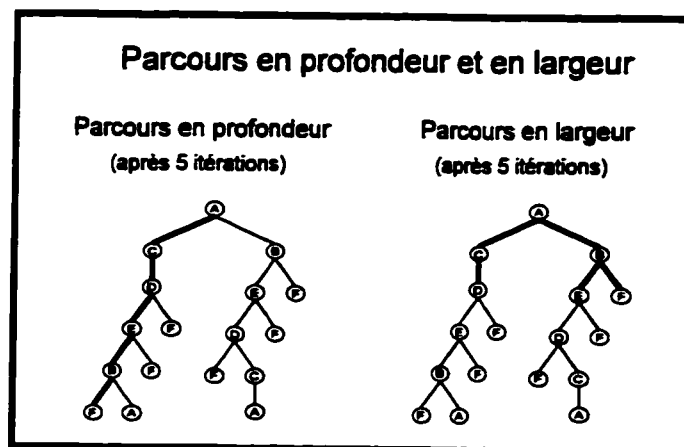


Figure 3.13: Parcours en profondeur et en largeur

Du sommet de départ, tous les sommets qui lui sont adjacents sont visités, puis tous les sommets adjacents aux sommets déjà visités sont examinés et ainsi de suite. Ce type de recherche s'étale en éventail, du sommet de départ, vers tous les autres sommets du réseau. L'intérêt de cette approche est que les chemins les plus courts (ou avec le moins d'arêtes) sont plus susceptibles d'être trouvés rapidement puisque le nombre d'arêtes parcourues est minimisé. En pratique toutefois, il n'est pas garanti que le chemin renfermant le moins d'arêtes soit vraiment le chemin le plus court. De plus, dans le parcours en largeur, le nombre de chemins potentiels s'accroît très rapidement en fonction du degré d'embranchement de l'arbre et peut devenir, dans le cas des réseaux complexes, très difficile à stocker en mémoire.

Que l'on adopte le parcours en largeur ou en profondeur, les arbres de parcours sont nettement supérieurs en terme d'efficacité et de rapidité à l'approche matricielle. Alors que l'algorithme de la cascade exige N^3 opérations pour traverser un réseau, les arbres de parcours et plus particulièrement l'algorithme de Dijkstra sont en mesure de la faire en N^2 opérations (ARLINGHAUS *et al*, 1990, p. 352). Pour un réseau de 1000 sommets, ce sont 1000 fois moins d'opérations à effectuer, ce qui peut se traduire par un temps de calcul de quelques minutes plutôt que de quelques... heures!

Pour augmenter d'avantage l'efficacité de l'approche par arbre de parcours, diverses tentatives ont été faites, notamment l'emploi du parcours bidirectionnel. De façon générale, et tel que démontré à la figure 3.12, la recherche du plus court chemin consiste à partir du sommet de départ et à progresser vers le sommet d'arrivée par itérations successives. Ce cheminement d'un état initial (départ) vers un état final (arrivée) correspond au *chainage avant*. A l'inverse,

lorsque l'on progresse à reculons, du sommet d'arrivée au sommet de départ, on parle de *chainage arrière*. L'idée principale du parcours bidirectionnel est que plutôt que d'effectuer la recherche d'un seul bout du chemin, on peut accélérer le processus en balayant le réseau des deux extrémités à la fois (SHAPIRO, 1987, p. 996).

A chaque itération, deux embranchements sont visités à la fois et si un plus court chemin existe, la présence d'un sommet commun aux deux parcours signifiera que les routes se sont rejointes quelque part au milieu du réseau. Pour que le chainage avant-arrière simultané fonctionne, il faut cependant opter pour le parcours en largeur aux deux extrémités de façon à balayer l'ensemble de l'arbre de parcours sinon les deux itinéraires risquent de ne jamais se rencontrer.

Bien que cette méthode ne réduise pas le nombre d'additions-comparaisons comparativement aux approches unidirectionnelles, elle permet de réduire le nombre d'itérations nécessaires et de gagner du temps de calcul. Pour les systèmes de guidage en temps réel qui exigent un temps de réponse rapide ou pour les réseaux complexes, ce gain de temps peut-être crucial. Ce procédé comporte cependant un inconvénient en terme de besoin de stockage. Étant donné que deux arbres sont traversés simultanément, toutes les variables et les tableaux doivent apparaître en double en mémoire de l'ordinateur.

L'ensemble des méthodes abordées jusqu'à maintenant ont toutes en commun une faiblesse: la recherche "à l'aveugle" des sommets pouvant potentiellement faire partie du plus court chemin. Toutes ces méthodes sélectionnent les sommets à visiter selon leur ordre d'apparition

dans la liste d'adjacence ou encore en favorisant l'arête la plus courte parmi celles disponibles. Aucun de ces algorithmes ne fait recours à une stratégie de sélection des sommets à visiter lui permettant de réduire les opérations de calcul et de maximiser la recherche du chemin optimum. Pour retrouver ce genre de démarche, on doit se tourner vers les algorithmes qui utilisent la recherche heuristique.

3.6.3 Plus court chemin: recherche heuristique

Les méthodes de recherche heuristique ont pour but d'améliorer la performance des modèles de calcul d'itinéraire dans les cas où un trop grand nombre de chemins alternatifs peuvent potentiellement conduire au plus court chemin entre deux sommets. Dans les réseaux renfermant plusieurs milliers de sommets, une recherche exhaustive de toutes les combinaisons peut très rapidement dépasser les capacités des ordinateurs usuels ou du moins exiger plusieurs jours de calcul⁶. Bien que cette approche systématique conduise inévitablement à une solution, si elle existe, des stratégies ont été développées pour réduire l'ampleur de la tâche. En fait, les méthodes par recherche heuristique ne diffèrent d'avec les arbres de parcours que par le mode de sélection des sommets à visiter.

Un des algorithmes les plus simple de recherche heuristique à avoir été développé est l'algorithme de l'escalade (Hill-climbing) (FISHER, 1990, p. 756). Cet algorithme fait référence, par analogie, à l'escalade d'une montagne, d'où son nom. En effet, lorsque l'on gravit une montagne avec pour objectif le sommet, on sait que l'on est dans la bonne direction si l'on

⁶Il est à noter que les applications d'analyse de réseaux se prêtent particulièrement bien à la technologie des ordinateurs parallèles dont le principe est de distribuer (paralléliser) une tâche ardue à plusieurs processeurs qui effectueront les opérations simultanément. Cette technologie est cependant relativement peu accessible à l'heure actuelle.

progresses constamment dans la direction de la pente la plus forte. En se fixant ainsi un azimut, on espère atteindre le plus rapidement possible l'objectif final. De la même manière, si l'on recherche prioritairement les sommets situés dans la direction du sommet d'arrivée, on augmente les chances de l'atteindre rapidement. Connaissant préalablement la position X,Y de chaque sommet, on pourra par exemple limiter la recherche des sommets à la partie du réseau située entre le sommet de départ et celui d'arrivée (figure 3.14).

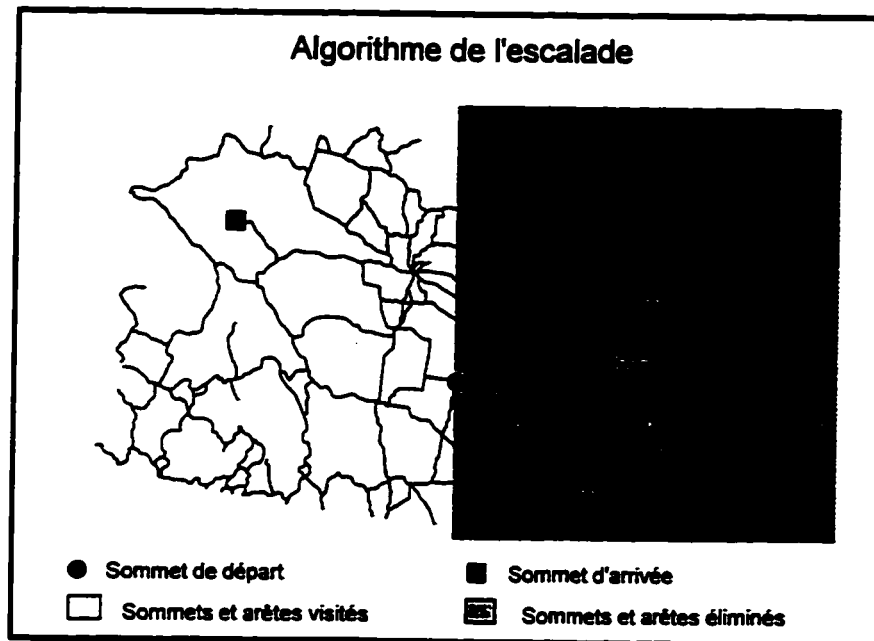


Figure 3.14: Algorithme de l'escalade

Pour la plupart des réseaux "normaux" ce type de démarche donne des résultats supérieurs, mais dans certaines circonstances, peut conduire à l'échec puisqu'un sommet éliminé pourrait conduire au chemin le plus court malgré sa position en retrait. Il arrive aussi qu'aucun chemin ne passe par les sommets retenus. On doit alors reprendre les calculs avec l'approche standard des arbres de parcours. Une autre particularité de l'algorithme de l'escalade est que contrairement aux autres méthodes, il ne garantit pas que le chemin découvert soit *vraiment le chemin le plus court* puisqu'il élimine dès le départ des sommets qui ne seront jamais visités.

Au mieux, on obtient un chemin qui est vraisemblablement le plus court compte tenu de la stratégie employée. Ce gain d'efficacité se fait donc au détriment de la garantie absolue du résultat optimum.

Pour remédier aux problèmes engendrés par l'algorithme de l'escalade, différents procédés heuristiques ont été développés dont l'algorithme A* (SHAPIRO, 1987, p. 997). Tout comme les algorithmes précédemment mentionnés, l'algorithme A* utilise la méthode des arbres de parcours pour traverser un réseau, mais par contre sélectionne les sommets selon la règle heuristique suivante:

$$f(i) = g(d,i) + h(i,a)$$

(D'après SHAPIRO, 1987, p. 997)

où: $f(i)$ = Estimé de la distance la plus courte entre le sommet de départ et celui d'arrivée
 $g(d,i)$ = La distance parcourue depuis le sommet de départ "d" jusqu'au sommet "i"
 $h(i,a)$ = Estimé de la distance la plus courte du sommet "i" jusqu'à l'arrivée "a".

Fait intéressant, l'algorithme A* trouvera toujours le chemin optimum entre deux sommets si la fonction heuristique $h(i,a)$ ne surestime jamais la distance actuelle entre le sommet "i" et le sommet d'arrivée "a". Dans le cas des réseaux de transport, la fonction heuristique $h(i,a)$ peut-être avantageusement remplacée par l'équation de calcul de la distance euclidienne entre deux points. Comme cette équation calcule la distance en ligne droite, elle ne peut en aucun cas surestimer $h(i,a)$. La fonction heuristique devient alors:

$$f(i) = g(d,i) + \sqrt{[(X_i - X_a)^2 + (Y_i - Y_a)^2]}$$

Les seuls inconvénients de l'algorithme A* sont tout d'abord la somme additionnelle de calculs que nécessitent la fonction heuristique mais surtout la quantité de mémoire nécessaire à son fonctionnement. A cet égard, il comporte les mêmes limites que la méthode des arbres de parcours en largeur (section 3.6.2).

Chapitre IV

Systemes d'information géographique et systemes de support à la décision à référence spatiale

“Tout ce qu'on pourrait attendre d'un système, c'est que par la vérité qu'il contient, il nous dédommageât de celle qu'il exclut.”

- Jean Rostand, Carnet d'un biologiste

Les chapitres précédents ont démontré que la recherche d'un itinéraire optimum entre deux points d'un réseau est un processus décisionnel qui implique des choix et des compromis. Pour mener à bien une recherche d'itinéraire optimum il faut, en plus d'un algorithme du plus court chemin, le plus grand nombre de données possibles desquelles on soutirera l'information qui permettra d'étayer ce choix d'itinéraire. Parmi les données nécessaires, plusieurs revêtent un caractère spatial comme le réseau de transport, l'utilisation du sol et les dénombrements de population. A ces entités spatiales, il faudra aussi relier des attributs qui permettront, entre autres, de qualifier les tronçons de routes du réseau et de faire la sélection de l'itinéraire le plus approprié.

Travailler avec autant de données issues de sources différentes, sous des formats différents, avec une précision différente et bien souvent sous des projections cartographiques différentes nécessite l'emploi d'outils capables d'intégrer et de ramener à un dénominateur commun toutes ces données. C'est à ce niveau qu'un *système d'information géographique (SIG)* est en mesure de jouer un rôle clef grâce à ses capacités de manipulation et d'analyse des données spatiales.

Cependant, la manipulation et l'analyse spatiale des données à l'aide d'un SIG n'est pas suffisante en soi pour permettre de prendre des décisions complexes, notamment si plusieurs alternatives doivent être évaluées. Qu'il s'agisse de déterminer l'itinéraire le plus sécuritaire possible pour le transport des matières dangereuses ou d'établir le plan de zonage d'une ville pour les quinze prochaines années, un schéma décisionnel doit être appliqué. Ce schéma comprendra généralement des *règles* à respecter, des *connaissances* empiriques sur lesquelles s'appuyer ou encore l'avis "*d'experts* en la matière" qui permettront d'en arriver à un choix final le plus judicieux possible. Un système d'information géographique capable d'offrir ce type d'approche correspond à un *système de support à la décision à référence spatiale (SSDRS)* et nécessite l'intégration de concepts d'intelligence artificielle aux SIG traditionnels.

Le présent chapitre abordera donc, dans un premier temps, le domaine des systèmes d'information géographique en s'attardant plus particulièrement à leur développement, leur mode de fonctionnement ainsi qu'à leur potentiel d'application en transport et dans la gestion des matières dangereuses. Dans un deuxième temps, nous verrons comment sont structurés les systèmes de support à la décision à référence spatiale et comment ils intègrent certains concepts d'intelligence artificielle et plus particulièrement de système expert. Cette deuxième

partie ne traitera que de l'apport des systèmes experts aux système d'information géographique puisque le prochain chapitre sera consacré entièrement aux systèmes experts et à la théorie de prise de décision.

4.1 Origine et développement des systèmes d'information géographique

Comme nous l'avons souligné au premier chapitre, le développement des méthodes quantitatives et de l'analyse spatiale a été tributaire des progrès de l'informatique. Le développement des systèmes d'information géographique peut lui aussi être relié à l'essor de l'informatique au cours des années cinquante et soixante. L'idée d'utiliser l'ordinateur pour manipuler et afficher les données à caractère spatial s'est manifestée dès que les périphériques d'affichage et de stockage de données sont apparus. SYMAP, le premier logiciel de cartographie assistée par ordinateur développé par Howard Fisher à l'Université Harvard, a fait son entrée dans le monde de la géographie en 1966, mais c'est au Canada que le concept de système d'information géographique a pris forme et que le tout premier système fut implanté (TOMLINSON, 1984).

Au cours des années soixante, le Canada s'est attaqué à une tâche qui représentait à l'époque un véritable défi: l'inventaire de l'utilisation du sol et la cartographie des potentiels forestiers, agricoles, fauniques et récréatifs pour tout le pays. Dès le départ, l'ampleur du projet, soit plusieurs milliers de cartes à produire, a mis en relief l'insuffisance des techniques manuelles et la nécessité d'automatiser le processus de production cartographique. Le stockage des données indispensables à cet inventaire posait aussi un problème tant par l'espace requis

que par la gestion et l'accès aux données. La technologie informatique représentait donc une solution potentielle malgré les limitations évidentes¹. Le maître d'oeuvre de ce qui allait être le Canadian Geographic Information System (CGIS) fut Roger Tomlinson, un géographe, et à la fin des années soixante le système était opérationnel.

Il est évident que les objectifs de développement des premiers systèmes d'information géographique étaient avant tout la manipulation et la gestion des données et la production cartographique. Développés par les grandes agences et pour leurs besoins particuliers, ces systèmes ne présentaient aucune capacité d'analyse spatiale et leur coût prohibitif à l'époque excluait la quasi-totalité du monde académique. Au cours des années soixante-dix, plusieurs grandes firmes américaines dont INTERGRAPH et SYNERCOM ont adapté leurs systèmes de conception assistée par ordinateur (CAO-CAD) pour répondre aux besoins de ce marché. Ces systèmes de CAO étaient toutefois mal adaptés aux besoins de la cartographie et c'est à ce moment que des firmes comme ESRI (Environmental System Research Institute) ont débuté le développement de ce qui allait être les SIG que l'on connaît aujourd'hui.

La baisse marquée du prix d'achat des ordinateurs et l'augmentation fulgurante de leur puissance qui s'est engagée vers la fin des années soixante-dix et poursuivie tout au long des années quatre-vingt ont favorisé l'accès à cette technologie. Les universités ainsi que les agences et les firmes à moyen et à petit budget pouvaient dorénavant se payer ce type d'outil et les géographes se sont tout naturellement tournés vers les SIG. C'est à ce moment que la faiblesse marquée des SIG en terme d'analyse spatiale s'est le plus fait sentir, notamment chez

¹A cette époque, le plus puissant ordinateur disponible (un IBM 1401) possédait 16k de mémoire, traitait 1000 instructions à la seconde, pesait 2 tonnes et coûtait 600 000\$ (TOMLINSON, 1984, p. 32).

la communauté académique qui y voyait un outil de recherche prometteur. Développés avant tout pour répondre aux besoins des grandes agences, les SIG excellaient en terme d'intégration, de stockage et de manipulation des données spatiales, mais étaient plutôt décevants en terme d'analyse. Depuis le milieu des années quatre-vingt, c'est sur l'amélioration des capacités d'analyse que se sont concentrés les efforts de développement.

Dans la multitude de systèmes d'information géographique (plusieurs centaines) que l'on retrouve aujourd'hui sur le marché, certains présentent des capacités analytiques plus avancées (comme ARC/INFO d'ESRI ou SPANS d'INTERA-TYDAC) alors que d'autres ne demeurent que des systèmes de gestion et d'interrogation de bases de données spatiales sans plus. Malgré cette hétérogénéité des SIG, il est possible d'en dégager les caractéristiques de bases et d'en définir les principales fonctionnalités.

4.2 Structure et composantes d'un système d'information géographique

Les systèmes d'information géographique ont été un sujet de prédilection dans la littérature géographique au cours des dix dernières années. Bon nombre d'auteurs ont abordé le sujet et parmi les plus marquants auxquels le lecteur peut se référer notons BURROUGH (1986), ARONOFF (1989), PEUQUET et MARBLE (1990), MAGUIRE *et al* (1991), GOODCHILD (1992) ainsi que LAURINI et THOMPSON (1992) pour ne mentionner que ceux-là.

Il n'est pas aisé de proposer une définition universellement acceptée de ce qu'est un SIG en raison de la diversité du domaine et de son manque de maturité. Diversité d'une part du point

de vue des concepteurs et des utilisateurs qui en tant que groupe hétérogène reflète les multiples facettes de ces systèmes; diversité d'autre part du côté des applications des SIG qui tentent de réunir en un tout des idées issues de champs de recherche parfois forts différents. Toute la rhétorique des "apôtres des SIG" directement liée aux aspects commerciaux et lucratifs, qui représentent annuellement plusieurs centaines de millions de dollars en Amérique du Nord, a aussi contribué à créer un fort bruit de fond sur les aspects théoriques qui sont de bien piètres arguments de vente.

MAGUIRE *et al* (1991) a dégagé trois grandes tendances concernant les différentes définitions des SIG dans la littérature (MAGUIRE *et al*, 1991, p. 13). Un premier groupe tend à percevoir les SIG du point de vue cartographique, ceux-ci étant avant tout un système de traitement et de production cartographique. Une deuxième tendance considère plutôt les SIG comme un système sophistiqué de gestion des bases de données spatiales, le SIG devenant alors le pourvoyeur de données pour les géographes. Finalement, un troisième noyau met l'emphasis sur l'importance de l'analyse spatiale et de la modélisation dans les SIG et tente de faire reconnaître une science de l'information spatiale. Bien que la majeure partie des systèmes soient encore limités en terme d'analyse spatiale, cette troisième voie semble être la plus acceptée dans le milieu des SIG et aussi la plus prometteuse en terme de développement.

Outre ces trois grandes tendances, ce qui se dégage des publications sur les systèmes d'information géographique est leur approche "boite à outils" (toolbox). En effet, les SIG ne se présentent pas comme des systèmes monolithiques mais plutôt comme une collection de modules intégrés en un ensemble de fonctions. CARTER (1989) les définit comme étant des

systemes integres permettant la saisie, le stockage, la gestion, l'analyse et la visualisation d'informations de nature géographique. D'autres auteurs tels que CRACKNELL (1986, p. 721) et ARONOFF (1989, p. 42) sont encore plus spécifiques quant aux composantes d'un logiciel de SIG et les décrivent comme un système composé de quatre modules de base 1) Module d'entrée et de codage des données 2) Module de stockage et de gestion 3) Module de manipulation et d'analyse 4) Module de sortie des résultats (figure 4.1).

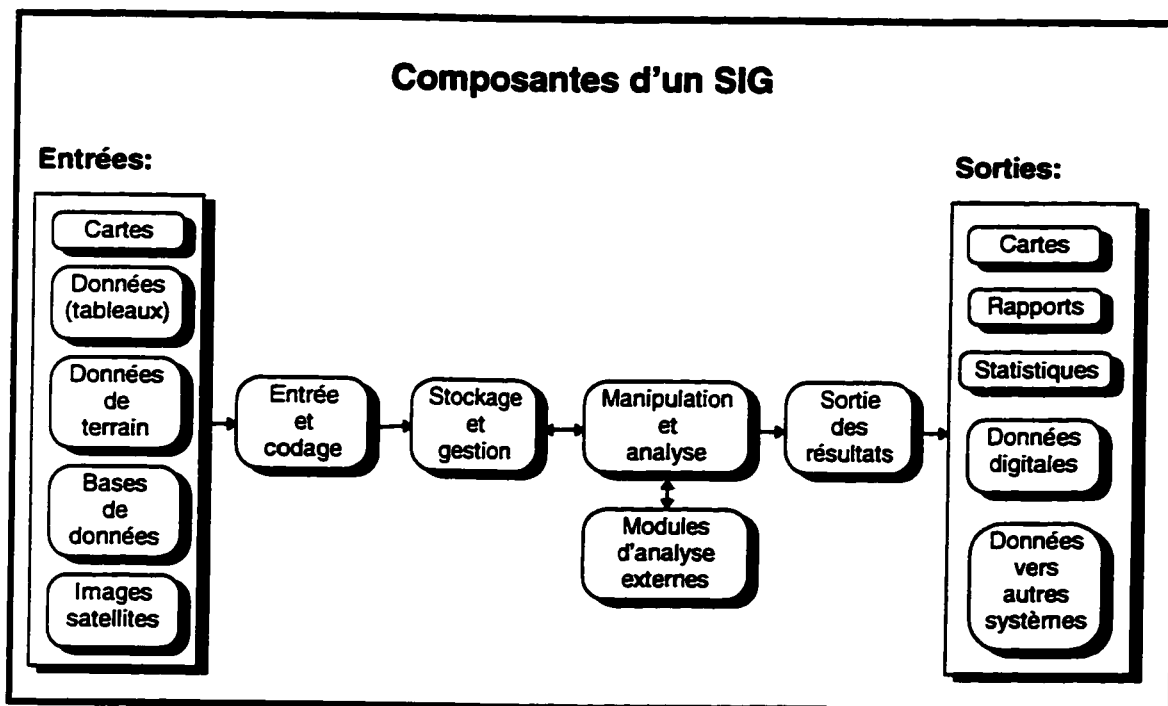


Figure 4.1: Composantes d'un système d'information géographique

A ces modules, il faut également ajouter des capacités d'intégration de données multisources ainsi que la possibilité de greffer des modules d'analyse externes développés par l'utilisateur. Cette ouverture est essentielle puisque c'est par elle qu'il devient possible d'adapter ou de créer de nouveaux modèles d'analyse et de les intégrer au système.

A chacun de ces modules correspond une série de fonctions que l'on retrouve couramment, avec plus ou moins d'emphase dans les SIG. Certains systèmes seront très performants en terme de manipulation et de gestion de données et plutôt faibles en analyse alors que pour d'autres, l'analyse et la sortie des résultats constituent un atout majeur. Sans faire office de liste exhaustive, la figure 4.2 présente les principales fonctions associées à chaque module.

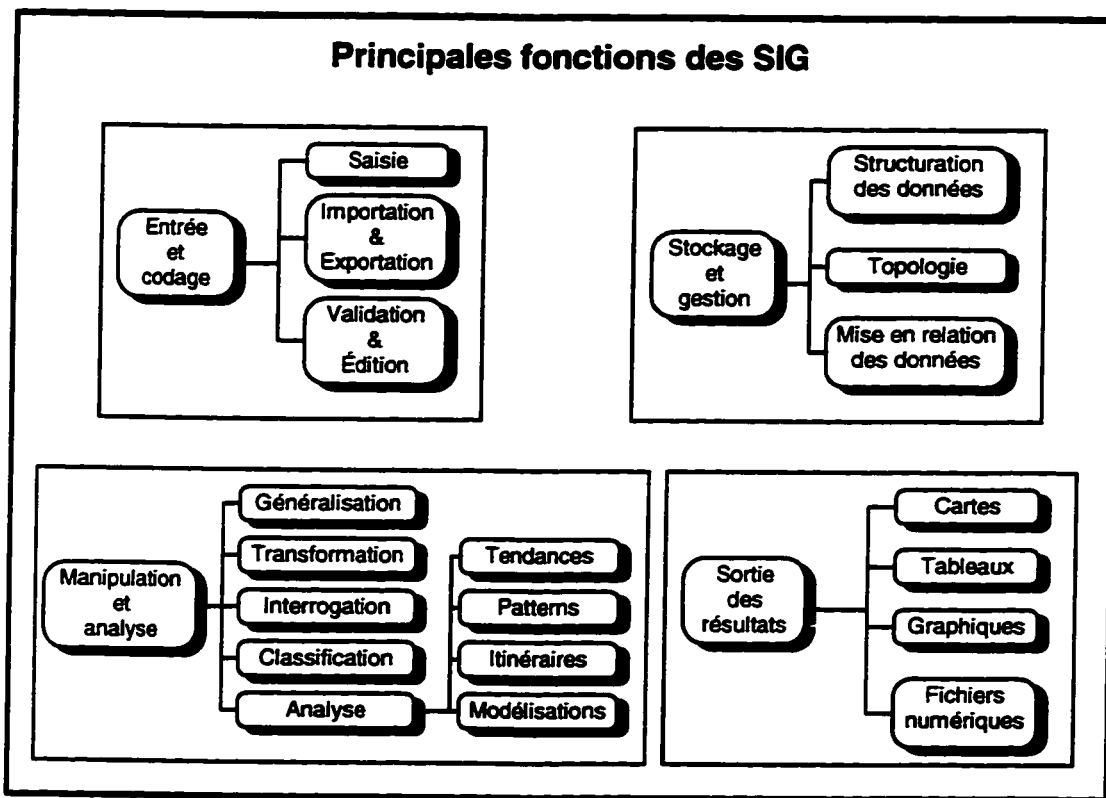


Figure 4.2: Principales fonctions des systèmes d'information géographique

Le module d'entrée et de codage des données permet la saisie par numérisation, balayage optique, imagerie satellite ou par système de positionnement global des données graphiques (spatiales) nécessaires au SIG. Il permet également la saisie des données d'attributs (données alphanumériques) à partir de fichiers informatiques ou à partir du clavier. Outre la saisie de données, ce module peut aussi valider les données pour en éliminer les erreurs de codage ou

les éditer à des fins de correction ou de mise-à-jour. La disponibilité sans cesse croissante de données déjà sous forme informatique implique également que le module d'entrée et de codage offre des fonctions d'importation et d'exportation des données graphiques et d'attributs vers d'autres systèmes.

Le module de stockage et de gestion est un module qui présente peu d'interactions avec l'utilisateur bien que sa présence soit essentielle au bon fonctionnement du SIG. C'est ce module qui opère la structuration des données, c'est-à-dire la construction de la topologie dans le cas des systèmes vectoriels ainsi que la mise en relation des données d'attributs avec les entités graphiques. Une fois les données introduites dans le SIG, c'est le module de stockage et de gestion qui s'occupe de l'intégrité des données.

Le module de manipulation et d'analyse est certainement celui sur lequel il se met actuellement le plus d'emphase au niveau du développement. Il s'agit certainement du module qui varie le plus en terme de fonctionnalités d'un système à l'autre. On y retrouve entre autres la généralisation cartographique, des fonctions de transformations affines ou de projections cartographiques, l'interrogation booléenne des bases de données, la classification des données d'attributs ou d'images satellites ainsi que des fonctions d'analyse spatiale telles que la détermination de tendances (surfaces de tendance) ou de patterns, de calculs d'itinéraires ainsi que de modélisation.

Un système d'information géographique "ouvert" permettra habituellement de greffer à ce module d'autres modèles d'analyse, modèles développés soit par l'utilisateur lui-même ou

encore provenant d'autres systèmes. C'est à ce niveau qu'il serait par exemple possible de fusionner un système expert à un système d'information géographique pour en faire un outil de prise de décision.

Quant au module de sortie des résultats, c'est à partir de lui que l'utilisateur peut mettre en forme le produit final qui servira à communiquer ses résultats. Ce module est capable, entre autres, de produire des cartes thématiques, des tableaux ou rapports, des graphiques ainsi que des fichiers numériques. Dans ce dernier cas par exemple, une carte produite sous forme numérique par le logiciel ARC/INFO pourra être récupérée par un autre système, tel qu'ARCVIEW, pour procéder à d'autres traitements ou pour faciliter la manipulation.

Pour que cette approche modulaire propre aux SIG soit fonctionnelle et qu'il soit possible à l'utilisateur de développer de nouvelles fonctions au système, il faut la présence d'un lien commun entre les différents modules. Ce lien doit permettre le transfert de données entre chaque programme et assurer que l'intégrité des données ne sera pas compromise suite à un traitement par un module spécifique. Ce lien qui assure aussi la communication entre chaque module du SIG correspond à la *structure de données* du système, c'est-à-dire à son mode de représentation des données spatiales et d'attributs.

4.3 Représentation des données dans un système d'information géographique

Bien que la structure de données dans un système d'information géographique joue un rôle primordial dans l'intégration des différents modules, son importance transcende les aspects

techniques et reflète le mode de représentation de la réalité de terrain à l'intérieur même du système. Le monde réel peut être reproduit de deux façons dans un SIG (figure 4.3). On peut d'une part représenter l'espace à partir d'une description exhaustive de la réalité à l'aide d'entités spatiales ou géométriques, entités auxquelles on associera des attributs et des relations. La représentation de l'espace géographique dans le SIG se fera alors à partir d'une liste (base de données) renfermant la totalité des entités de cet espace. Un SIG que l'on utiliserait pour un inventaire complet de l'utilisation du sol d'un territoire représenterait l'espace géographique sous cet angle puisqu'une description complète de la réalité est codée dans le système. En analyse de réseau, cette approche est habituellement employée puisque l'on connaît la totalité du réseau dès le départ.

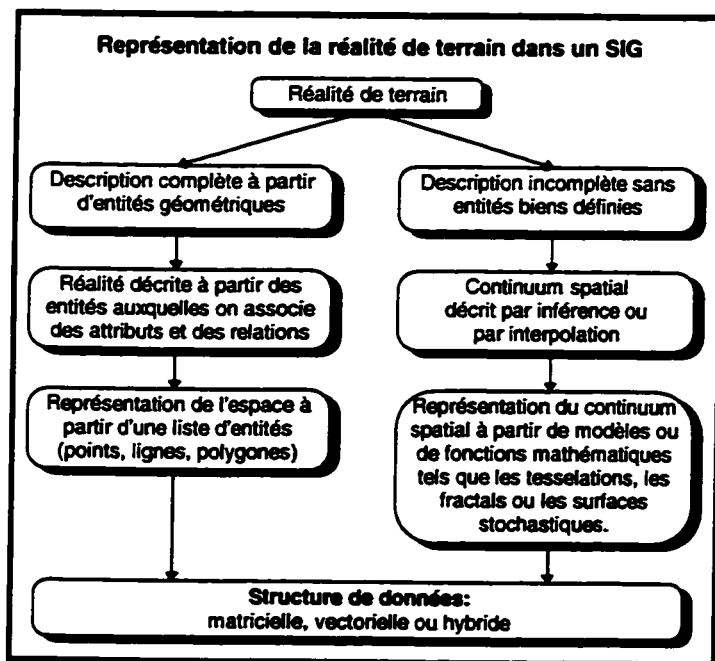


Figure 4.3: Représentation de la réalité de terrain dans un SIG

D'autre part, il est également possible de représenter l'espace à partir d'une description incomplète, sans connaître la totalité des entités qui le compose. L'espace géographique

devient alors un continuum spatial que l'on décrit par inférence ou par interpolation. Ce continuum spatial est généralement obtenu à partir de modèles mathématiques ou de fonctions permettant de combler l'absence d'information de la réalité de terrain. Ce type d'approche se retrouve par exemple lorsque l'on veut représenter le potentiel agricole des sols pour une région donnée. Des échantillons de sol seront prélevés ponctuellement sur le territoire et un modèle mathématique sera appliqué à ces échantillons pour interpoler à l'ensemble du territoire cette réalité de terrain incomplète.

Que l'on opte pour l'une ou l'autre de ces approches, la représentation de la réalité de terrain dans un SIG impose des compromis. Dans le cas de la description complète, on fixe la réalité de terrain lors de la saisie de données, ce qui enlève de la souplesse aux diverses fonctions de modélisation et de simulation des SIG. Par contre, en abordant la réalité sur une base incomplète, toute la représentation du continuum spatial repose sur des équations mathématiques dont on ne peut jamais vraiment garantir qu'elles reproduisent fidèlement la réalité. Ce gain de souplesse dans la représentation s'accompagne donc d'une approximation de la réalité de terrain véhiculée par le SIG. Des situations encore plus problématiques peuvent se produire lorsque les deux approches se retrouvent dans le même système et que des données issues d'une description exhaustive sont confrontées à des données provenant d'un modèle mathématique. Dans ces situations, les résultats obtenus dépendent pour beaucoup de la manière dont la réalité a été codée dans le SIG.

Sans égard à l'une ou l'autre des approches de représentation de la réalité de terrain, les données spatiales peuvent être codées de trois grandes façons dans un SIG. On peut opter

pour une structure de données matricielle, vectorielle ou hybride, choix qui repose souvent sur le type de problème que l'on désire solutionner à l'aide du SIG.

4.3.1 La structure de données matricielle

La structure de données matricielle est sans doute le moyen le plus simple de représenter l'information spatiale dans un système d'information géographique. Elle consiste à appliquer une matrice bidimensionnelle où chaque cellule (ou pixel) renferme une valeur d'attribut et peut être localisée par ses indices ligne-colonne (figure 4.4).

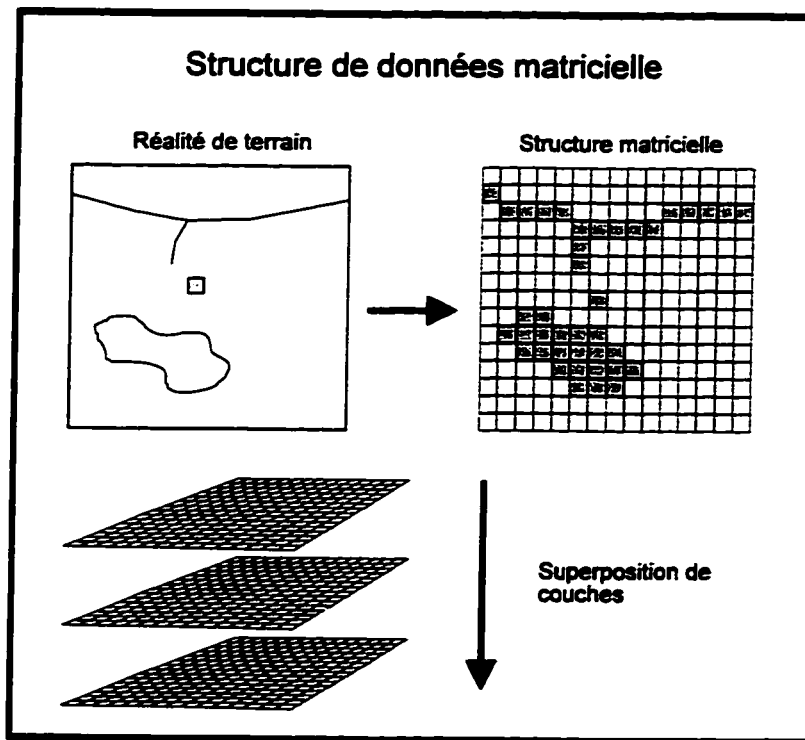


Figure 4.4: Structure de données matricielle

Parmi les avantages de cette structure notons la possibilité d'établir directement des relations spatiales, c'est-à-dire que pour un pixel donné on peut connaître directement la valeur de ses voisins immédiats. La structure matricielle permet également d'intégrer facilement les données

satellites et de superposer plusieurs couches d'information dans un SIG. Le pixel étant l'unité de base de cette structure, sa taille détermine le degré de généralisation cartographique, la précision des données et dans une bonne mesure la précision des résultats. Une grille fine représentera plus adéquatement la réalité de terrain mais nécessitera un plus grand espace de stockage, ce qui implique, malgré les techniques de compression de données, un compromis espace de stockage/précision.

Bien qu'il soit techniquement possible d'utiliser la structure matricielle pour les analyses de réseau, son emploi ne présente aucun avantage concret dans la manipulation des éléments linéaires puisque la structure arc-sommet se prête mal à l'unité spatiale discrète qu'est le pixel. La précision spatiale limitée à la taille du pixel et l'absence de topologie du mode matriciel sont aussi des inconvénients majeurs. Le seul rôle joué par la structure matricielle dans les applications en transport et en analyse de réseau est dans l'intégration des données matricielles (images satellites par exemple) de source externe qui seront par la suite converties en structure vectorielle avant d'être utilisées dans les analyses de réseau.

4.3.2 La structure de données vectorielle

La structure de données vectorielle est issue de la cartographie traditionnelle où les éléments cartographiques sont représentés par des points, des lignes ou des polygones localisés dans l'espace cartésien à partir de coordonnées X,Y. On y retrouve la structure vectorielle non-topologique (structure spaghetti) utilisée par les logiciels de conception assistée par ordinateur (CAO) et où les éléments cartographiques ne sont qu'une série de lignes sans relations topologiques. Cette structure vectorielle inefficace en analyse spatiale est de moins en moins

employée dans les SIG d'aujourd'hui. La structure de données vectorielle topologique, très populaire dans les SIG, renferme par contre une topologie explicite avec pour éléments de base le *noeud* et la *chaîne* (figure 4.5). Le noeud représente soit une entité ponctuelle ou encore l'intersection ou les extrémités d'une chaîne. Cette dernière, en plus des noeuds peut contenir une série de *points intermédiaires* qui déterminent son profil. Finalement, une ou plusieurs chaînes peuvent former un *polygone* dont la principale caractéristique est de symboliser une figure géométrique fermée.

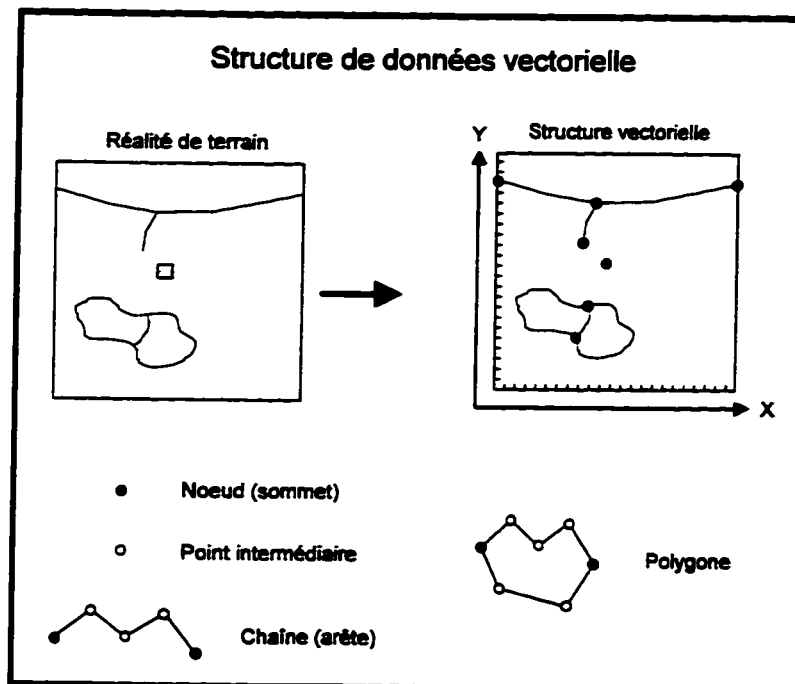


Figure 4.5: Structure de données vectorielle

L'ordre hiérarchique **noeud** → **chaîne** → **polygone** de la structure vectorielle présente plusieurs analogies avec la théorie des graphes. Les noeuds correspondent aux sommets des graphes alors que les chaînes sont l'équivalent des arêtes. L'encodage explicite de la topologie dans la base de données permet aisément de rattacher aux éléments noeuds, arêtes et polygones des attributs qui pourront être utilisés ultérieurement dans les analyses.

Contrairement à la structure matricielle, la structure vectorielle permet de maintenir une haute précision du produit cartographique, précision qui n'est pas dépendante de la taille fixe d'un pixel mais de la précision numérique des coordonnées X,Y. L'espace requis pour stocker les bases de données est également inférieur au mode matriciel puisque seules les coordonnées des vecteurs sont emmagasinées dans les fichiers, ce qui signifie une manipulation plus facile et plus rapide des données.

En géographie des transports et en analyse de réseaux, la structure vectorielle topologique est un choix *de facto* puisque aucune autre structure de données ne peut reproduire aussi fidèlement les caractéristiques d'un réseau tout en offrant autant de souplesse. En milieu urbain, il est commun qu'un même réseau de transport, représenté sous forme vectorielle, soit perçu par les différents utilisateurs à la fois comme un système unidimensionnel, bidimensionnel ou tridimensionnel.

Par exemple, le responsable du trafic routier modélisera chaque tronçon du réseau comme un élément unidimensionnel en ne tenant compte que de la direction du trafic. Le responsable des travaux de voirie transformera le même tronçon en un élément bidimensionnel à l'aide d'une zone tampon de part et d'autre de la ligne médiane pour représenter la surface de pavage à réparer. Enfin, l'ingénieur civil responsable des infrastructures urbaines associera les tronçons de rue à des éléments tridimensionnels puisque l'on y retrouve simultanément des égouts enfouis sous terre, des poteaux de signalisation au niveau du sol et des fils suspendus en hauteur. Seule la structure vectorielle est en mesure d'offrir cette vision multiple d'un même élément cartographique.

L'inconvénient le plus notable de la structure vectorielle réside dans la nécessité de décrire de façon explicite les liens **noeud** → **chaîne** → **polygone** dans la base de données et de gérer de façon cohérente ces liens lors des mises à jour. De façon concrète, la structure vectorielle implique le maintien de trois fichiers reliés entre eux par des pointeurs. Un premier fichier, celui des noeuds, renfermera la totalité des noeuds présents dans la base de données, un deuxième fichiers, l'ensemble des chaînes et un troisième tous les polygones. Chaque élément de ces trois fichiers est relié par pointeur aux éléments qui lui sont reliés spatialement.

Ainsi, pour un polygone donné, une série de pointeurs indiqueront dans le fichier des chaînes celles qui forment le polygone. De la même manière, chaque chaîne pointerait (dans le fichier noeuds) vers les noeuds qui définissent ses extrémités ou une intersection avec d'autres chaînes. En sélectionnant un polygone, on peut donc savoir immédiatement quels sont les éléments qui le composent et les attributs qui sont associés à chacun de ces éléments.

Bien que la structure vectorielle soit le choix évident en analyse de réseau, d'autres structures de données sont apparues au cours des dernières années. Il s'agit essentiellement de structures hybrides qui ont pour but de combler certaines lacunes des modes vectoriel, mais surtout matriciel.

4.3.3 Les structures de données hybrides

Parmi ces structures hybrides notons les tessellations qui sont une structure matricielle à résolution variable et dont la forme des pixels peut prendre la configuration d'hexagones ou de polygones de Thiessen. Le quadtree est probablement la structure hybride la plus populaire et

emprunte certains concepts aux arbres de parcours définis au chapitre 3. Elle consiste à subdiviser successivement l'espace en quadrats jusqu'à l'obtention de l'homogénéité des quadrats. Cette subdivision successive peut être représentée par un arbre de parcours à quatre embranchements. Cet arbre décrit en fait les liens topologiques dans la base de données puisqu'en le parcourant on peut connaître directement quels sont les éléments adjacents sur la carte. Bien que cette structure en arbre comporte certaines analogies avec la théorie des graphes, le quadtree est peu utilisé dans les applications en transport. Un certain nombre d'autres structures ont également été développées au cours des dernières années (octree, voxel, TIN, etc.). On peut cependant affirmer, qu'à l'heure actuelle, aucune structure de données autre que la structure vectorielle topologique ne peut répondre aussi adéquatement aux besoins des applications en transport et en analyse de réseaux dans les SIG.

Qu'ils utilisent une structure vectorielle ou matricielle, les SIG ne seraient que de simples outils de dessin s'ils n'étaient pas en mesure de rattacher aux éléments graphiques (pixel, point, ligne ou polygone) des attributs pouvant être utilisés dans les analyses. La mise en relation des données d'attributs alphanumériques aux éléments cartographiques constitue une fonction majeure des systèmes d'information géographique. Dans les applications en transport ou en analyse de réseau, c'est à partir de cette fonction qu'il est possible d'associer à chaque tronçon de route tous les paramètres permettant de modéliser les déplacements et d'effectuer les choix dans la recherche d'un chemin optimum.

4.3.4 Les données d'attributs

Dans un système d'information géographique, la tâche de mettre en relation les données d'attributs alphanumériques aux éléments cartographiques revient au module de stockage et gestion. Pour accomplir cette fonction, la plupart des SIG utilisent le modèle de base de données relationnel, c'est-à-dire que pour chaque élément cartographique (pixel, noeud, chaîne, polygone), un identificateur unique est généré lors de l'entrée et du codage des données. Cet identificateur, présent à la fois dans la base de données cartographique et la base de données alphanumériques permet de mettre en relation les fichiers à l'aide de pointeurs (figure 4.6).

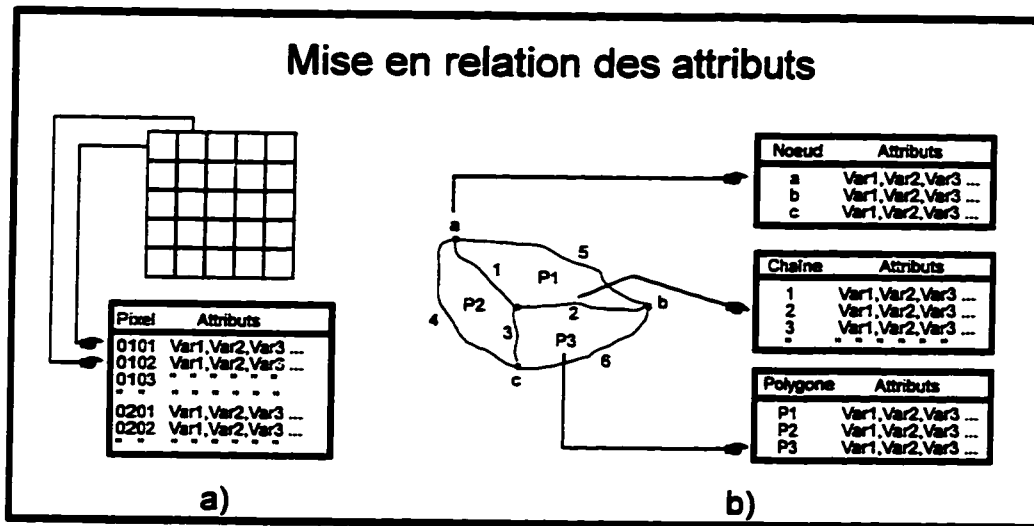


Figure 4.6: Mise en relation des attributs

En mode matriciel, l'identificateur unique de chaque pixel pourra être sa position ligne-colonne et sa présence dans le fichier alphanumérique permettra de connaître directement toutes les variables associées à un pixel donné (figure 4.6a). En mode vectoriel, les fichiers de noeuds, de chaînes et de polygones peuvent être traités de la même manière. Le maintien, en trois listes distinctes, des éléments cartographiques signifie cependant que trois séries

d'identificateurs uniques seront manipulées par le système de gestion de bases de données (figure 4.6b).

Les avantages du modèle relationnel sont nombreux. Tout d'abord, ce type de modèle offre un accès direct aux données d'attributs à partir d'une "clef géographique" car c'est l'identificateur unique, généré à partir de la base de données cartographique, qui constitue le moyen d'accès aux données. De plus, le modèle relationnel facilite le géocodage des données d'attributs, c'est-à-dire le positionnement géographique des données à partir des adresses civiques. Par exemple, si l'on possède les adresses civiques pour chaque tronçon de rue d'une ville, on pourra localiser spatialement et rattacher à un tronçon particulier toutes les bases de données ayant l'adresse civique parmi ses champs d'information. Il devient alors possible de connaître tous les propriétaires fonciers pour une rue donnée, tous les abonnés du téléphone pour un tronçon de rue, etc.

Cette possibilité de joindre des bases de données ayant au moins un champ commun peut s'appliquer à n'importe quel type de données, qu'elles soient spatiales ou non. Par exemple, à partir du code postal de l'adresse civique, il est possible pour les entreprises de marketing de localiser spatialement les profils socio-économiques établis à partir de sondages et ainsi de cibler des régions ou des clients potentiels. La plupart des applications des SIG tirent profit de cette capacité de mise en relation de données. En transport cette fonction permet, entre autres, d'intégrer au réseau routier les données provenant des recensements de Statistiques Canada pour effectuer de dénombrements de population par tronçon ou encore déterminer les caractéristiques de l'utilisation du sol à partir des secteurs de dénombrement. L'intérêt d'utiliser

les SIG en transport est donc accru par cette capacité d'intégrer au réseau des données qui ne pourraient l'être autrement. L'apport des SIG en transport ne se limite cependant pas à cet aspect puisque plusieurs applications ont été réalisées au cours des dernières années, applications qui démontrent leur potentiel au niveau de la gestion des infrastructures et de la modélisation des réseaux.

4.4 Les systèmes d'information géographique en transport

L'utilisation des systèmes d'information géographique pour le transport des matières dangereuses relève d'un cadre plus général, à savoir l'emploi des SIG en transport (gestion des infrastructures, calcul d'itinéraires, guidage de véhicules) ainsi que dans le domaine de la protection civile (impact des accidents, planification des mesures d'urgence, vulnérabilité aux risques). A l'origine, la plupart des modèles développés par les ingénieurs en transport étaient des boîtes noires (black box), c'est-à-dire des modèles dont les utilisateurs ne pouvaient modifier les paramètres internes et dont les entrées et les sorties consistaient en de longues séries de chiffres sans résultats visuels graphiques.

De plus, les efforts importants nécessaires pour "nourrir" ces systèmes en données de terrain, le caractère unique (application specific) de chaque système, la redondance et les coûts prohibitifs de la mise en forme des bases de données avaient pour effet de ralentir considérablement leur développement. Chaque système et chaque application faisait l'objet d'une collecte de données spécifique, dans un format spécifique, à une échelle spécifique et avec une structure de données spécifique. L'intégration de données et les transferts d'un modèle à l'autre se révélaient la plupart du temps impossibles.

Ces multiples carences devaient assez tôt générer des frustrations, notamment en ce qui concerne la visualisation des résultats, et dès 1975 une équipe de Princeton a travaillé à l'ajout d'une dimension (carto)graphique aux modèles de transport (KORNHAUSER, 1984, p. 53). Exception faite de l'affichage graphique, les diverses tentatives de l'époque n'offraient cependant que peu de fonctionnalités en regard des besoins réels de la modélisation en transport, plus particulièrement en ce qui a trait à l'intégration des données.

La popularité croissante des SIG dans les domaines autres qu'en transport au cours des années quatre-vingt a peu à peu fait prendre conscience aux spécialistes en transport (incluant les géographes) que presque toutes les données qu'ils utilisent peuvent être localisées dans l'espace. L'idée d'utiliser la référence spatiale comme clef unique d'accès aux données a ouvert la voie des SIG dans le domaine. En associant des coordonnées géographiques à chaque donnée disponible, l'intégration des données et des résultats des modèles devenait possible de même que la coopération entre agences et le partage des ressources. Cependant, l'emploi des SIG en transport a très rapidement débordé l'intégration de données pour s'insinuer dans les aspects de gestion et de planification des réseaux de transport.

4.4.1 Les systèmes d'information géographique en planification et en gestion des réseaux de transport.

Bien que les systèmes d'information géographique soient utilisés dans les domaines du transport aérien, ferroviaire et maritime, c'est au niveau de la planification et de la gestion des réseaux routiers que leur utilisation est la plus répandue. Plusieurs états américains (Californie, Wisconsin, Texas, Arizona) et quelques provinces canadiennes (Saskatchewan, Alberta) ont

développé des systèmes d'information routière (Road Information Systems) basés sur les SIG.

Ces systèmes visent entre autres la gestion et la planification des:

- Travaux de pavage
- Inspection et entretien des ponts
- Enquêtes sur la circulation
- Inventaires des panneaux de signalisation
- Enquêtes sur les accidents (points noirs)
- Travaux de construction et de maintenance
- Flux de trafic et embouteillage
- Transport en commun
- Inspection des traverses à niveau
- etc,

Cette liste est loin d'être exhaustive et l'avantage de la composante spatiale est de permettre de connaître facilement et rapidement les travaux en cours ou à venir pour un tronçon de route donné. La coordination des équipes de travail devient donc beaucoup plus efficace. D'ailleurs, ce type de système d'information n'est pas conçu comme une boîte fermée mais comme un système central rendant disponibles toutes les données nécessaires aux différents services publics (figure 4.7).

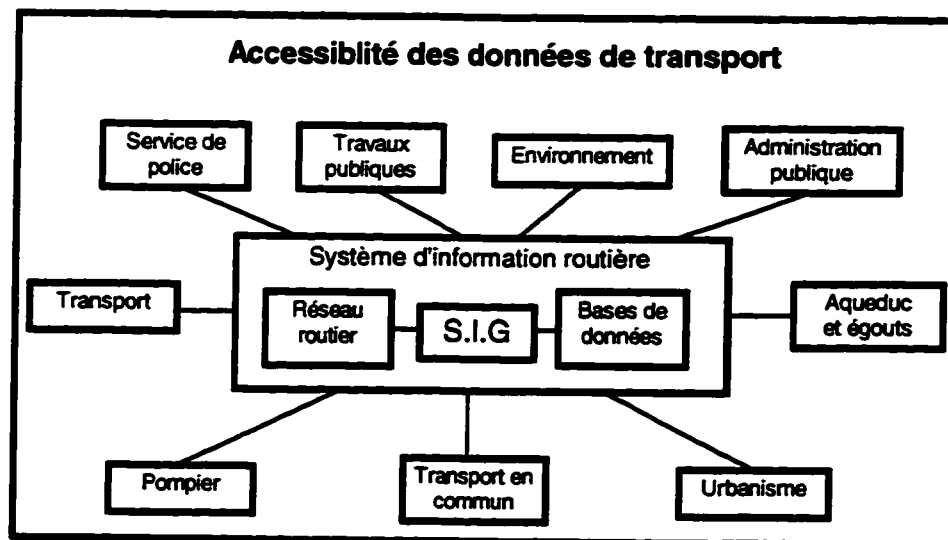


Figure 4.7: Accessibilité des données de transport

Chaque service est ainsi en mesure de connaître les interventions planifiées par les autres divisions ou encore d'obtenir des données relatives au réseau de transport. Ce type d'application implique cependant un protocole d'accès aux bases de données bien défini pour éviter que des données erronées ne viennent compromettre le fonctionnement de l'ensemble de l'organisation.

En plus de fournir informations et données sur les réseaux, les systèmes d'information routière offrent également diverses fonctions reliées au domaine des transports (VONDEROHE *et al*, 1993). Notons entre autres **a**) les mesures de base (distance, superficie) **b**) les analyses de proximité (tampons, adjacence, voisinage) **c**) la génération de modèles numériques de terrain (altitude, pente, volume) **d**) l'analyse de réseaux (segmentation dynamique, plus court chemin, itinéraires, affectation de ressources) **e**) les superpositions (de couches, de réseaux) **f**) les liens vers d'autres modèles (modèles de flux, simulation, origine-destination). Cette dernière fonction est des plus prometteuses puisqu'elle ouvre la porte vers les modèles de simulation et d'analyse spécialisés en transport tel que EMME/2 (KATZ et CHEVALIER, 1994) ou TRANSCAD (SIMKOWITZ, 1990).

Outre leur application en gestion et en planification, les SIG présentent également un potentiel pour le guidage et la répartition des flottes de véhicules. Plusieurs grandes organisations privilégient l'emploi des SIG pour la planification des itinéraires de livraison (Postes Canada) ou encore pour le suivi en temps réel (voir section 2.3.5). Dans le cas des véhicules d'urgence, les bénéfices se traduisent par une diminution du temps de réponse (BADILLO, 1993, p. 163) tandis que la planification des itinéraires permet d'optimiser les parcours, de maximiser le

travail du personnel et de réduire la consommation de carburant (LANDRY, 1993, BADILLO, 1993, GRUDE, 1992). Le guidage des véhicules automobiles privés est également à l'essai sur une base expérimentale en Californie (BADILLO, 1993, p. 173). Bien qu'à grande échelle le guidage automobile puisse permettre de réduire les embouteillages, la consommation d'essence et la pollution urbaine par une optimisation des déplacements, la complexité et le coût des équipements en freine l'application à l'heure actuelle (voir section 2.3.5). Quant à l'emploi des SIG pour le transport des matières dangereuses, c'est dans le cadre plus général de la protection civile que l'on peut le situer.

4.4.2 Les systèmes d'information géographique et la protection civile.

Comme son nom l'indique, la protection civile vise à minimiser l'impact sur la population des catastrophes technologiques ou naturelles. Historiquement, les intervenants en protection civile ont mis l'emphase sur la remédiation des effets négatifs des catastrophes en se limitant à "réparer les dégâts". Aujourd'hui, on tend à percevoir le rôle de la protection civile comme un cycle avec plusieurs phases: analyse de risques, analyse de vulnérabilité, réduction, prévention, alerte et remédiation. Les SIG peuvent donc être utilisés en amont de l'accident, pour en prévenir l'occurrence, et en aval, pour en mitiger les conséquences puisque la plupart des tâches comportent un caractère spatial.

Un des moyens les plus efficaces de réduire les risques associés à une catastrophe technologique est d'anticiper les conséquences d'une telle catastrophe à l'aide de scénarios ou de modèles de simulation. Ce type d'approche se déroule en 3 étapes:

- 1) Identifier les matières dangereuses présentes sur le territoire et définir les zones de risques associées à chacun des sites
- 2) Identifier la population vivant dans chaque zone de risque ainsi que ses caractéristiques économiques et sociales.
- 3) Identifier les ressources nécessaires pour une éventuelle évacuation.

L'identification des matières et la délimitation des zones de risque passent nécessairement par un inventaire sur le terrain. La détermination des zones de risques implique que l'on tienne compte de données climatiques (direction des vents) et topographiques (effet du relief) dans le cas des échappements dans l'atmosphère. Dans le cas des matières liquides, des données géomorphologiques sont requises pour évaluer l'impact sur la nappe phréatique. Pour les produits dangereux susceptibles de provoquer une explosion, la zone de déflagration doit le plus souvent être calculée à partir d'un modèle de simulation propre à chaque type de substance. Quant à l'identification de la population située dans les zones à risque, elle peut se faire aisément à l'aide d'un SIG puisqu'elle consiste à superposer les cartes de population aux zones de risque.

Enfin, l'identification des ressources nécessaires à l'évacuation est probablement l'étape la plus complexe puisqu'il faut tenter de prévoir le comportement de la population en situation de crise et tâcher d'utiliser à sa capacité maximale le réseau de transport pour l'évacuation. De plus, il faut généralement tenir compte, lors d'une évacuation, de la trajectoire du panache de dispersion pour éviter une seconde évacuation ainsi que de voir à garder l'accès au personnel d'intervention sur le site de l'accident.

Un autre usage des SIG en amont des catastrophes impliquant des matières dangereuses est dans le processus de sélection des sites de stockage. La nécessité de prendre en compte un grand nombre de facteurs environnementaux et sociaux de même que le syndrome NIMBY (Not In My Back Yard) ont grandement complexifié le processus de sélection des sites. La capacité d'intégration des données des SIG est en mesure d'alléger cette tâche et de présenter cartographiquement les nombreux facteurs ayant favorisé la sélection d'un site. A titre d'exemple, JENSEN et CHRISTENSEN (1989) présentent une application renfermant 18 couches d'information sur lesquelles toute une série d'opérations booléennes ont été effectuées.

L'emploi des SIG en aval des catastrophes peut prendre un caractère hautement dynamique (intervention pendant la catastrophe) ou à long terme (remédiation après la catastrophe). L'intervention lors d'un accident implique généralement la coordination de plusieurs équipes de spécialistes et parfois l'évacuation de la population, deux tâches nécessitant une logistique bien élaborée. Les systèmes globaux de positionnement (GPS) et les SIG en temps réel sont susceptibles de jouer un rôle majeurs dans ce domaine. Bien qu'il ne comporte actuellement aucun outil de positionnement en temps réel, le système IGLOU développé pour la région de Québec est un premier pas dans ce domaine (GUERRIER et GODON, 1993). Le système CAMEO développé par l'agence américaine NOAA (National Oceanic and Atmospheric Administration) offre aussi ce type de support à l'intervention.

La remédiation des impacts après une catastrophe est aussi un processus à long terme où les SIG jouent un rôle. Dans le cas des catastrophes technologiques, le suivi du processus de

décontamination des sites revêt un caractère spatial indéniable. La catastrophe de l'Exxon Valdez et la décontamination des plages des résidus d'hydrocarbures en est un exemple. En ce qui concerne les catastrophes naturelles, la récupération des ressources peut aussi nécessiter l'emploi des SIG. Par exemple, suite au passage de l'ouragan HUGO en Floride en septembre 1989, de vastes régions forestières ont été affectées. La logistique et la cartographie nécessaire à la récupération des arbres détruits a été réalisé à l'aide d'un SIG de façon à maximiser les ressources forestières limitées de la Floride (HALE, 1992,p. 77).

4.4.3 Les systèmes d'information géographique en transport des matières dangereuses

Selon LEPOFSKY et ABKOWITZ, pour les États-Unis seulement, 1.5 milliard de tonnes de matières dangereuses sont transportées annuellement (LEPOFSKY et ABKOWITZ, 1993, p. 239). Il n'est donc pas surprenant que plus en plus de gens exigent des mesures de réduction des risques liés à cette activité. De façon générale, il est possible de regrouper en trois grandes classes les intervenants en transport des matières dangereuses: **1) le public 2) les agences gouvernementales (transport, environnement, protection civile) 3) les industries (producteurs, transporteurs)**. Il va sans dire que ces groupes ont parfois des intérêts divergeant quant au transport de ces matières et que le recours à des outils facilitant la compréhension du problème du point de vue spatial est nécessaire.

L'approche la plus commune des SIG dans ce domaine consiste à effectuer une évaluation systématique des risques associés au transport d'un produit donné. Cette recherche d'itinéraire optimum est fondée sur la minimisation du risque relatif sommé sur toutes les sections de route

qui composent l'itinéraire, de l'origine à la destination. Un exemple typique d'application est celui de LASSARRE *et al* (1993) en Haute-Normandie. Leur approche comporte cinq étapes:

- 1) Le réseau est divisé en arcs (sections)
- 2) Un *taux d'accidents* est attribué à chaque section en fonction de ses caractéristiques. Ce taux est égal à la fréquence d'occurrence d'un accident impliquant des matières dangereuses multiplié par la longueur de l'arc.²
- 3) Dépendamment du type de substance impliquée, une zone d'impact de part et d'autre de la route définit la population potentiellement exposée.
- 4) Le taux d'accidents multiplié par le nombre de personnes exposées le long de la route donne la *valeur de risque*. Cette valeur sert de critère pour la recherche du chemin le plus sécuritaire.
- 5) L'algorithme du plus court chemin de DIJKSTRA est utilisé pour déterminer le chemin le plus sécuritaire.

L'inconvénient de cette approche est qu'elle nécessite une définition quantitative complète du problème ainsi que des données pour chaque arc du réseau. LEPOFSKY et ABKOWITZ (1993, p. 247) présentent une approche similaire pour le transport du carburant à fusée de l'Arizona vers la base de Vandenberg en Californie. Dans leur étude, ils ont pondéré la distance, le temps et les coûts de parcours avec la probabilité d'accidents, la population exposée et les zones écologiquement sensibles. D'autres exemples plus "simplistes" (VINCENT, 1989) ne tiennent compte que de la population totale adjacente à l'itinéraire sans égard à ses caractéristiques. Il faut toutefois noter que peu importe l'approche adoptée, ces applications

² Dans la pratique, ce taux est difficile à évaluer en raison de l'occurrence très faible d'accidents par tronçon.

comportent toutes la même lacune: seule la valeur finale de risque est utilisée comme impédance pour le calcul de l'itinéraire optimum (voir chapitre 2).

Comme les pages précédentes l'ont démontré, le champ d'intervention des SIG en protection civile et en transport des matières dangereuses s'inscrit dans le *processus décisionnel* face aux événements. Le SIG a pour rôle de seconder l'expert dans ses prises de décisions en fournissant des informations susceptibles de l'éclairer sur le problème. Vue de façon globale, le SIG devient un outil de support à la décision ou *système de support à la décision* et puisqu'un SIG traite les données spatialement, on est alors en présence d'un *système de support à la décision à référence spatiale (SSDRS)*.

4.5 Les systèmes de support à la décision à référence spatiale

Les systèmes de support à la décision (SSD) ont pour origine les systèmes d'information de gestion (Management Information System) dont les premiers développements ont eu lieu au cours des années 50 (WATERS, 1989b, p. 741). A cette époque, les systèmes d'information de gestion étaient développés dans le but de gérer l'information sans cesse croissante nécessaire aux entreprises et faciliter le processus de prise de décisions d'affaires. Le postulat était simple: en rendant disponible une grande quantité de données de qualités aux décideurs, la qualité des décisions prises devait s'accroître. Avec l'expérience, il est apparu que de noyer les décideurs sous de grandes quantités d'informations n'aidait en rien le processus décisionnel et qu'il fallait plutôt se tourner vers des systèmes capables, après analyse, de générer un nombre limité de solutions alternatives pouvant guider les décideurs.

De nombreux auteurs ont contribué à définir les systèmes de support à la décision; THERIAUF nous en donne la définition suivante (THERIAUF, 1988, p. 15):

“Decision support systems are systems designed to support the decisions of managers and some operating personnel [...]. Such systems assist organization personnel in reaching effective decision that contain elements of subjectivity and objectivity, typical of most business problems. The capability of combining subjectivity (individual judgment) with objectivity (the computer's output) permits a more thorough exploration of the problem.”

Les SSD n'ont donc pas pour tâche de remplacer les décideurs mais de les assister dans leur fonction, d'accroître leur efficacité sans remplacer leur capacité de jugement. Alors que les systèmes traditionnels sont essentiellement des outils de production et de gestion de l'information, les SSD intègrent le jugement humain aux modèles mathématiques pour la prise de décision (WATERS, 1989b, p. 742). Cette intégration des modèles quantitatifs et du jugement humain leur permet de résoudre des problèmes de tous genres, structurés, semi-structurés ou non structurés (THERIAUF, 1988, p. 50). Les modèles quantitatifs pourront être des modèles de programmation mathématiques (allocation des ressources, calcul d'itinéraires), des modèles probabilistes (simulation), des modèles statistiques (corrélation, régression) des modèles de séquence optimum ou encore des modèles reposant sur la programmation dynamique. Quant à l'intégration du jugement humain, elle s'opère par une approche basée sur l'inférence. Le système de support à la décision fait alors référence à une base de connaissances (les faits) ainsi qu'à un système de traitement des faits (moteur d'inférence) qui appliquera des règles de décision préétablies par des "experts". Il s'agit donc, globalement, d'intégrer aux modèles quantitatifs traditionnels une approche déductive propre aux systèmes experts.

Un peu bizarrement, les développements qui ont conduit aux systèmes de support à la décision à référence spatiale (SSDRS) ne sont pas issus des SSD mais ont évolué en parallèle depuis les années 60. Tout comme les systèmes d'information de gestion, les SIG ont été le point de départ d'une approche orientée vers la gestion des données. Puis graduellement, les besoins en analyses et en outils de prise de décision ont favorisé l'apparition des SSD et des SSDRS, pour les données à caractère spatial (figure 4.8).

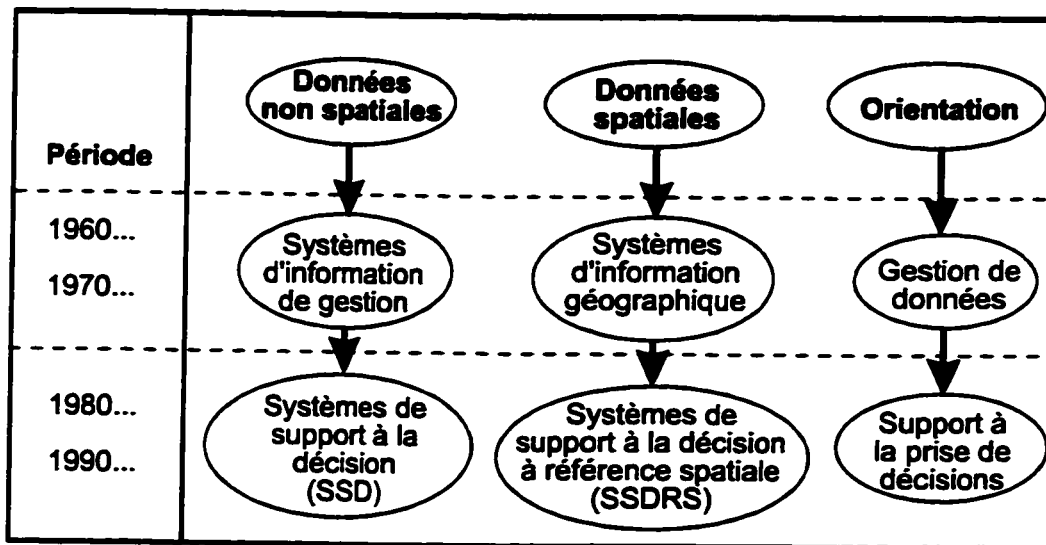


Figure 4.8: Gestion de données et prise de décisions

Comme cela est souvent le cas dans bien des domaines, les deux démarches sont demeurées relativement isolées tout au long de leur développement et selon DENSHAM (1991, p. 407), les SSDRS seraient de 10 à 15 ans en retard sur les SSD. L'intégration d'une approche déductive aux méthodes quantitatives traditionnelles des SIG est donc un domaine encore peu exploré.

L'organisation interne des SSDRS repose sur une structure modulaire reflétant le couplage SIG-système expert, ce dernier étant le moyen par lequel "le jugement humain" est intégré au SIG. Les modules internes de gestion, d'analyse, de génération d'affichage et de rapports font référence aux fonctions de base des SIG, alors que le coquille de système expert (expert system shell) englobe les aspects "non quantitatifs" (figure 4.9).

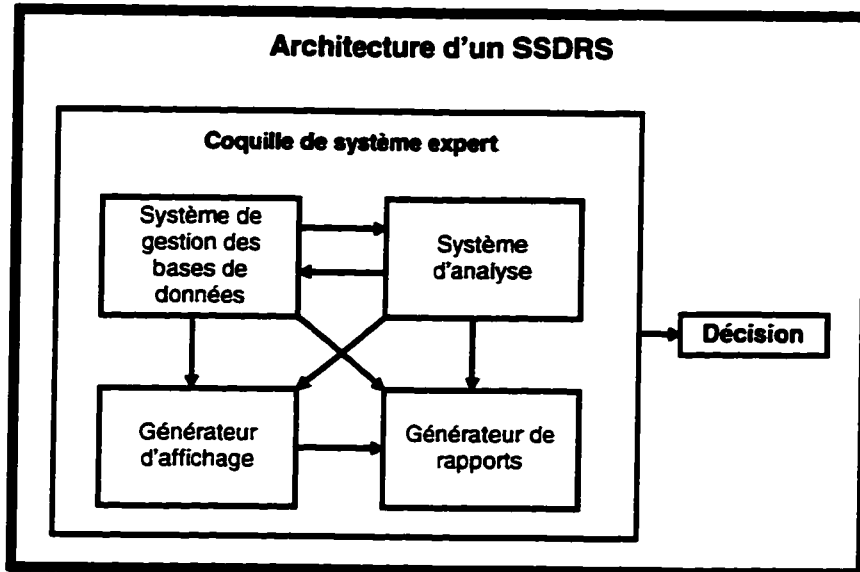


Figure 4.9: Architecture d'un SSDRS

En pratique, l'importance de chacun des modules peut varier selon les systèmes et les applications, l'approche "expert" pouvant prendre le pas sur les relations mathématiques ou l'inverse. L'intégration de cette approche requiert cependant que l'on possède une bonne compréhension de la structure et du fonctionnement d'un système expert ainsi que des mécanismes de prise de décision en présence d'incertitude. L'application au domaine du transport des matières dangereuses nécessite également que l'on définisse la notion de risque. Le chapitre 5 a pour objectif d'explorer les principaux aspects de la théorie de la décision et des systèmes experts alors que la notion de risque sera abordée au chapitre 6.

Chapitre V

Systèmes experts et théorie de la décision

"J'avais l'habitude d'être indécis, mais maintenant je n'en suis plus très sûr."

- Boscoe Pertwee

Vivre, c'est prendre des décisions. Qu'il s'agisse de décisions triviales du quotidien ou de décisions dont dépend la vie de milliers de gens, le mécanisme menant à un choix raisonné demeure le même. On parlera d'un "choix éclairé", d'une "sage décision" ou encore d'une décision arbitraire ou stupide pour qualifier le processus ayant conduit à un choix donné. A des degrés divers, toute prise de décision nécessite des informations ou des connaissances auxquelles seront appliquées des règles de décision. Ces règles pourront être purement mathématiques ou encore basées sur l'intuition et l'expérience. Dans les meilleures conditions, la totalité de l'information nécessaire à la prise de décision sera disponible, dans les pires situations, l'incertitude sera telle que seules l'intuition et l'expérience ou encore l'arbitraire pourront mener à une décision.

Au chapitre 4, nous avons pu constater que les systèmes d'information géographique, en raison de leur capacité à manipuler l'information spatiale et de leur potentiel de couplage à un système expert, sont en mesure de devenir de véritables systèmes de support à la décision à référence spatiale (SSDRS). Le présent chapitre abordera donc le domaine de la théorie de la décision et l'un de ses champs d'application particulier: celui des systèmes experts. Nous verrons, dans un premier temps, le processus de prise de décision puis, dans un deuxième temps, les composantes internes des systèmes experts ainsi que leur fonctionnement.

5.1 Théorie de la décision

La nécessité de prendre des décisions de façon rationnelle est probablement à l'origine de la théorie de la décision. La contribution la plus significative de cette théorie ayant été de mettre de l'avant des principes sur lesquels des critères de sélection seront construits et des solutions seront proposées. En plus d'offrir aux décideurs les moyens d'analyser leurs problèmes, cette théorie permet également de justifier les solutions proposées.

Bien que l'on considère le mathématicien Daniel Bernoulli comme le père de la théorie de la décision moderne, cette dernière se trouve à l'intersection de nombreuses disciplines et plus particulièrement des statistiques et des mathématiques (KAST, 1993, p. 6). La théorie des jeux de hasards et les probabilités y occupent une place importante et de nombreux scientifiques tels que POISSON (probabilité), LAPLACE-GAUSS (loi normale) et VON NEUMANN (utilité espérée) y ont contribué. En fait, trois aspects de la théorie de la décision méritent que l'on s'y attarde. En premier lieu, le processus de formalisation du problème, deuxièmement

l'etablissement des regles quantitatives ou qualitatives de decision et, finalement, l'introduction de la notion d'incertitude dans le processus de prise de decision.

5.1.1 Formalisation du probleme de decision

La formalisation du probleme de decision est un prealable a toute analyse de prise de decision. C'est souvent a partir de graphiques, d'arbres de decision ou de tableaux que le probleme sera formalise et que les methodes et criteres permettant de calculer et de justifier les meilleurs choix seront obtenus (KAST, 1993, p. 20). On peut distinguer quatre elements dans la formalisation d'un probleme de decision:

- 1) **La determination de l'ensemble des decisions** - Il s'agit de determiner les valeurs du vecteur renfermant l'ensemble des choix possibles. Dans le cas du transport des matieres dangereuses, il pourrait s'agir, a une intersection donnee, des troncons de routes pouvant etre empruntes.
- 2) **La determination de l'ensemble des evenements elementaires** - Il s'agit de determiner l'ensemble des etats des elements non controles par le decideur. L'utilisation du sol ou les differentes populations a risque le long de chaque troncon de route composent ce vecteur. On pourra, par exemple, regrouper sous differentes categories de vulnerabilite chaque troncon de route.
- 3) **La determination des consequences possibles** - Cette etape consiste a evaluer les consequences possibles (impacts) d'un choix donne. C'est a partir de cet ensemble que le decideur sera en mesure de comparer les differentes consequences et, de ces comparaisons, effectuer un choix raisonne. Dans le cas du transport des matieres dangereuses, les consequences d'un accident pourront etre classees selon le type de matiere impliquee et la gravite de l'accident. Parmi les consequences on retrouvera par exemple: accident sans impact - accident avec evacuation - accident avec evacuation et blesses - accident avec evacuation et decès, etc.

4) La détermination de la relation entre décisions, événements élémentaires et conséquences - Cette relation correspond à la "logique décisionnelle" qui est fonction des décisions possibles, des événements élémentaires et des conséquences possibles. Cette relation renferme habituellement un certain niveau d'incertitude qui peut se traduire par un degré de confiance associé au résultat final. Dans l'exemple du transport des matières dangereuses, cette relation pourrait prendre la forme suivante: "Si j'emprunte le tronçon T1, je serai en présence des événements élémentaires {E1, E2, E3} qui pourraient avoir comme conséquence possible {C1, C2, C3, C4}". De plus, à cette relation pourrait être associé un degré de confiance (exemple 0.8) indiquant la "force" de cette relation.

Outre leur caractère fondamental dans la théorie de la décision, ces quatre éléments de formalisation sont présents dans la démarche d'élaboration de tout système expert puisque ces derniers sont en fait des "systèmes de prise de décision" basés sur des critères rationnels définis par la théorie de la décision.

5.1.2 Arbre de décision

Une manière courante de décrire la formalisation d'un problème de décision consiste à la représenter sous forme d'arbre de décision, c'est-à-dire un graphe dont les sommets sont soit les décisions (sommets de décision), les événements (sommets d'événement) ou les conséquences (sommets de conséquence) et les arêtes la relation entre chaque composante (SHAPIRO, 1987, p. 229). La figure 5.1 illustre la structure que pourrait prendre un arbre de décision hypothétique dans le cas du transport des matières dangereuses. Dans ce cas, le choix d'un tronçon (tronçon 1 ou 2) avec un type de matière dangereuse particulier (acide, explosif ou gaz), selon que l'on est le jour ou la nuit, nous mettra selon le cas en présence d'une population X, Y, X1 ou Y1 avec pour conséquence potentielle A, B, A1 ou B1 (figure 5.1).

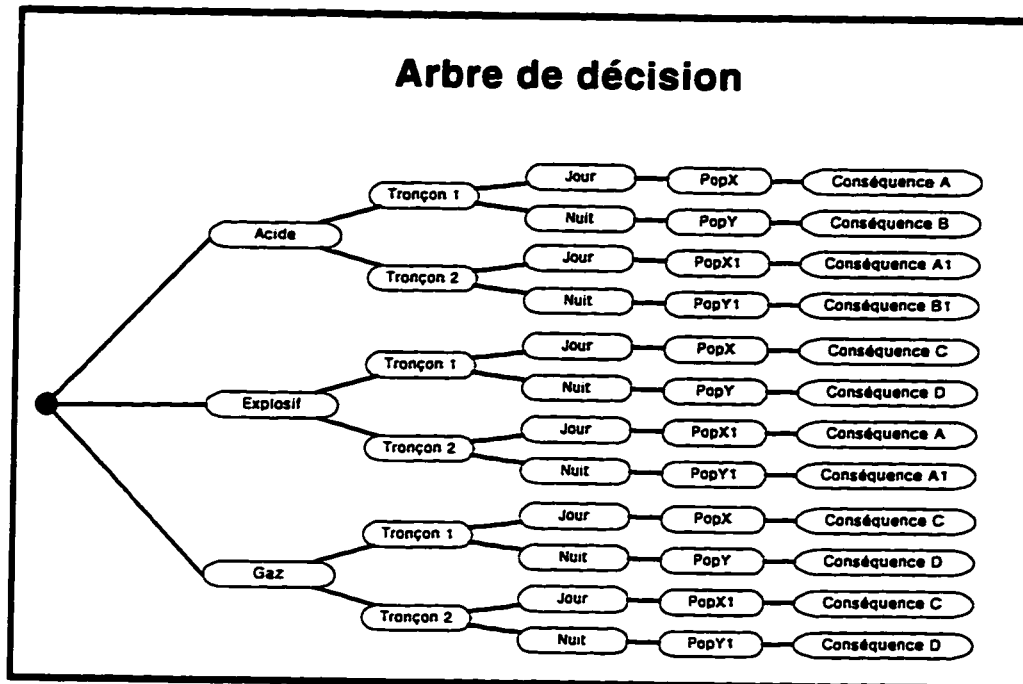


Figure 5.1: Arbre de décision

L'avantage de représenter un problème de décision à partir d'un arbre réside dans le fait que celui-ci permet de décrire le problème tel qu'il se présente dans la réalité, en faisant apparaître progressivement les différents éléments (KAST, 1993, p. 28).

Évidemment, chaque embranchement de l'arbre mène à une situation potentiellement différente et une décision implique donc d'évaluer les alternatives à chaque sommet de l'arbre. Cette évaluation des alternatives pourra se faire de façon quantitative et utiliser l'approche probabiliste ou encore de façon qualitative et opter pour les méthodes de raisonnement qualitatif. En plus d'illustrer les différentes alternatives, l'arbre de décision permet également de formaliser l'incertitude, phénomène courant dans le processus de décision. Sans s'attarder aux détails de la notion d'incertitude qui sera traitée ultérieurement dans ce chapitre, il faut toutefois

préciser que selon la théorie de la décision, l'incertitude sera formalisée à l'aide des trois éléments suivants (KAST, 1993, p. 30):

- Un ensemble de résultats possibles
- Un ensemble d'événements qui affectent les conséquences des décisions. Un événement étant caractérisé par les résultats qui le vérifient.
- Une pondération de ces événements.

Le poids affecté à chaque événement peut provenir de la fréquence à laquelle il a été observé ou encore d'une évaluation subjective de la vraisemblance de son apparition. Une fois l'arbre de décision construit et l'incertitude formalisée, il faut procéder à la prise de décision à l'aide d'outils d'analyse. L'approche la plus souvent retenue est celle de la formulation mathématique à l'aide de la théorie des probabilités.

5.2 Théorie de la décision: l'approche probabiliste

Historiquement, le développement de la théorie des probabilités est intimement lié à celui de la théorie de la décision. C'est au 17^{ième} siècle que les premières formalisations de cette théorie ont fait leur apparition dans le contexte des jeux de hasard. Le chevalier de Méré, philosophe mais surtout joueur invétéré, fut l'un des premiers à poser le problème en termes scientifiques (KAST, 1993, p. 16). Mais c'est généralement à POISSON et à LAPLACE, au 19^{ième} siècle, à qui l'on attribue les travaux de pionniers. LAPLACE fut d'ailleurs le premier à publier une synthèse sur le sujet. L'avantage majeur incitant à l'emploi des probabilités en prise de décision réside, d'une part, sur les bases mathématiques solides de cette théorie et, d'autre part, dans la capacité d'introduire le concept d'incertitude dans la résolution de problème.

5.2.1 De l'événement indépendant aux probabilités conditionnelles

La probabilité qu'un événement "E" se produise $P(E)$ représente la quantification de la vraisemblance de l'occurrence ou de la non-occurrence de cet événement. Une valeur de $P(E) = 1$ indiquera l'occurrence certaine de cet événement alors qu'une valeur de $P(E) = 0$ assurera l'impossibilité de l'occurrence de l'événement. Lorsqu'une expérience donnée pourra conduire à plusieurs événements possibles, une probabilité sera associée à chaque événement et correspondra au rapport du nombre de cas favorables à la réalisation de l'événement "E" par rapport au nombre total d'événements "N" pouvant prendre place. On parlera alors de *probabilité classique* ou *a priori*. Mathématiquement, cette relation s'exprime:

$$P = E/N$$

Par exemple, en lançant un dé, on sait que la probabilité d'obtenir un nombre particulier $P(2)$ est de une sur six et que la somme des probabilités des événements possibles $P(1)...P(6)$ est égale à 1. Dans un processus décisionnel, les événements uniques n'ont généralement que peu d'intérêt et, dans la mesure où l'arbre décisionnel comporte plusieurs sommets de décisions, d'événements et de conséquences, il faut avoir recours à d'autres outils probabilistes. La *probabilité composée* permet de trouver la probabilité d'un événement qui résulte de l'occurrence successive de deux ou plusieurs événements de probabilité connue (BOURSIN, 1986, p. 49). La probabilité que deux événements E_1 et E_2 se produisent correspond au produit des probabilités individuelles de E_1 et E_2 (ROLSTON, 1988, p. 98):

$$P(E_1 E_2) = P(E_1) * P(E_2)$$

Evidemment, cette approche ne peut être utilisée que dans les cas où les événements considérés sont indépendants, c'est-à-dire que l'occurrence de $P(E_1)$ ne modifie pas la probabilité d'occurrence de $P(E_2)$, donc $P(E_1/E_2) = P(E_2/E_1)$. En transport des matières dangereuses, la probabilité composée peut être utile pour calculer l'occurrence d'un accident "A" pour une matière dangereuse donnée "M" $P(AM)$ en considérant que le type de matière transportée et la probabilité d'un accident sont indépendants.

Toutefois, cette condition d'indépendance s'accorde plutôt mal aux processus décisionnels où les conséquences sont généralement dépendantes des événements et où il n'est pas rare qu'un événement E_2 se produise seulement si les conditions d'un événement E_1 sont déjà en place. Par exemple, les probabilités d'un déversement de matières dangereuses dans l'environnement sont conditionnelles à ce qu'il y ait un accident. Dans ce cas particulier, on utilisera l'approche des *probabilités conditionnelles* ou encore celle des *probabilités bayésiennes*. Dans le cas des probabilités conditionnelles, la solution est formulée à l'aide de l'équation suivante (GIARRATANO et RILEY, 1994, p. 200):

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

où:

$P(B|A)$ = Probabilité d'occurrence de l'événement "B" si l'événement "A" s'est produit.

La formule de BAYES, bien que similaire, offre encore plus de flexibilité que la formule des probabilités conditionnelles puisqu'elle permet également de calculer la probabilité inverse.

Selon ce théorème la probabilité qu'un événement "B" se produise si l'événement "A" s'est déjà produit $P(B|A)$ est égale à (ROLSTON, 1988, p. 99):

$$P(B|A) = \frac{P(A|B) * P(B)}{\sum_{i=1}^K P(A|B_i) * P(B_i)}$$

ou:

- $P(B)$ = La probabilité d'occurrence de l'événement "B" sans l'occurrence de l'événement "A".
- $P(A|B)$ = La probabilité d'occurrence de l'événement "A" si l'événement "B" s'est produit.
- $P(B|A)$ = La probabilité conditionnelle que l'événement "B" se produise étant donnée l'occurrence de l'événement "A". Il s'agit de la probabilité "révisée" de $P(B)$ en tenant compte de l'information issue de l'occurrence de "A".
- K = Nombre total d'événements "B" possibles.

L'avantage de l'approche bayésienne est qu'elle permet de parcourir l'arbre de décision en calculant la probabilité conditionnelle à chaque sommet de l'arbre. A partir d'une situation connue à un sommet donné (connaissance à priori), on détermine la probabilité de chaque conséquence possible (a postériori). Il faut toutefois préciser que, lorsque l'on associe cette approche à un modèle d'optimisation tel que l'algorithme du plus court chemin, la quantité de calculs nécessaires pour en arriver à une solution devient rapidement astronomique.

Cependant, le problème majeur de l'approche probabiliste réside dans sa nature même: l'usage des fréquences relatives. Il faut connaître la probabilité que chaque événement de l'arbre se réalise pour connaître la probabilité finale de chaque conséquence. Or en transport des matières dangereuses, il est plutôt rare (pour ne pas dire impossible) que l'on ait en main l'ensemble des données nécessaires (nombre total de véhicules, nombre d'accidents, nombre

de deversements, etc.) au calcul des probabilités pour chaque tronçon de route. Les données manquantes ou hétérogènes sont trop nombreuses.

Clairement, une approche quantitative nécessitant une définition complète du problème est inappropriée, d'autant plus que la rareté des événements considérés (faible probabilité d'occurrence) enlève tout sens au concept de fréquence relative. En effet, bien qu'un accident soit possible, les valeurs très faibles des probabilités conditionnelles calculées enlèvent une grande partie du pouvoir discriminant entre chaque tronçon de route.

5.2.2 Probabilités subjectives

Certains moyens ont été mis de l'avant pour pallier aux problèmes de l'approche inférentielle basée sur la fréquence d'occurrence d'un événement. Un de ces moyens est l'utilisation des *probabilités subjectives* par opposition aux *probabilités objectives* issues de l'inférence statistique. La probabilité subjective d'un événement représente le degré de conviction qu'un observateur bien informé aura sur les chances de voir un événement particulier se réaliser (FRENCH, 1988, p. 222). Il s'agit évidemment d'une notion intuitive qui peut s'appuyer sur des connaissances plus ou moins fondées mais qui, par contre, évite la collecte exhaustive des données nécessaires à la compréhension objective d'un système quelconque.

Cette approche est intéressante puisqu'elle permet, dans le cas du transport des matières dangereuses, de contourner le problème des données manquantes ou incomplètes. Par exemple, un expert du domaine pourrait calculer la probabilité d'occurrence d'un accident pour chaque tronçon de route et, au besoin, faire appel à son intuition dans les cas où les données

sont insuffisantes. La somme de travail qu'exigerait cette approche dans le cas de réseaux complexes avec des milliers de segments de routes devant être analysés manuellement rend toutefois impossible en pratique cette méthode. De plus, chaque expert pourrait, selon son interprétation, en arriver à des probabilités d'accident différentes pour un même tronçon. Devant une telle situation, les connaissances et le raisonnement de chaque expert deviendraient potentiellement une "boîte noire" entre les données de bases (les caractéristiques du réseau routier) et le résultat final (l'itinéraire déterminé).

C'est à ce niveau que le couplage système d'information géographique-système expert peut jouer un rôle important puisqu'il est possible d'introduire dans la base de connaissances du système expert les connaissances et les règles d'inférence utilisées par les experts humains pour solutionner le problème. Ainsi, on utiliserait des règles d'inférences et des connaissances bien définies et connues de l'utilisateur; règles d'inférence qui seraient appliquées uniformément sur l'ensemble du réseau. Du point de vue de l'approche strictement probabiliste, les systèmes experts sont donc en mesure de rendre plus flexible la détermination d'itinéraires optimum. Leur contribution au domaine peut cependant être bien plus importante si l'on considère l'intégration de données qualitatives ainsi qu'un raisonnement qualitatif dans la détermination d'un itinéraire le plus sécuritaire possible pour le transport des matières dangereuses.

5.3 Théorie de la décision: l'approche qualitative

De nombreuses recherches en intelligence artificielle portant sur le processus de prise de décision chez l'humain ont démontré que ces derniers font appel à trois grands types

d'habiletés dans la prise de décision: l'expérience, l'intuition et la déduction logique (DEWHURST et GWINNETT, 1990, p. 694). Alors que la déduction logique peut aisément se traduire sous forme de modèles mathématiques ou de propositions logiques, il en va autrement de l'expérience et de l'intuition qui reposent sur une approche plus floue.

L'expérience, qui permet dans bien des cas de prendre des décisions éclairées (et qui n'exclut pas la déduction logique) se bâtit au fil des ans chez l'expert humain qui développe une compréhension holistique de son domaine. L'expert peut alors se comparer à une "banque de connaissances vivante", connaissances qui pourront s'étendre du modèle mathématique le plus rigide aux relations de causalité plus ou moins floues. Bien qu'une décision prise par un expert et basée sur l'expérience ne puisse pas toujours être traduite en termes mathématiques, il s'agit habituellement d'une décision qui peut être justifiée logiquement. Par contre, les décisions basées sur l'intuition relèvent essentiellement d'un "sentiment intérieur" qui dicte la décision à prendre. Contrairement à la décision basée sur l'expérience, la décision basée sur l'intuition ne peut se justifier en termes mathématiques ou logiques.

En plus de s'exprimer sur un spectre allant de l'intuition à la déduction logique stricte, la prise de décision implique également la sélection des informations significatives à la décision. *"Le monde est infiniment complexe et notre connaissance de celui-ci est finie et par conséquent incomplète"* (KUIPERS, 1994, p. 1). L'habileté de l'expert humain, c'est d'être en mesure de raisonner et de décider même si sa connaissance du système en cause n'est pas parfaite et

absolue.¹ Cette approche de la theorie de la decision qui consiste a faire appel, au besoin, aux modeles mathematiques, a la connaissance qualitative des experts et au traitement des donnees incompletes ou incertaines est le *raisonnement qualitatif*.

5.3.1 Raisonement qualitatif: l'arithmetique d'intervalle

Une foule de decisions sont prises quotidiennement sur la base d'un raisonnement qualitatif (ou semi-quantitatif). Raisonner de facon qualitative, c'est prendre une decision sans connaitre la valeur numerique exacte des parametres du systeme. Toucher un cable electrique pour determiner s'il surchauffe, remplir un reservoir "aux trois quarts"; diagnostiquer une maladie ou prospecter de l'or impliquent dans tous les cas une part de raisonnement qualitatif. Alors qu'une approche strictement quantitative exigerait une connaissance precise (temperature, volume, poids, etc.) de la part de l'expert, dans une approche qualitative ce dernier saura, avec une marge d'erreur reduite, comment evaluer qualitativement (ou semi-quantitativement) les differentes variables du systeme. Des qualificatifs comme "froid", "tiède", "chaud" ou encore "petit", "moyen", "grand" pourront faire partie du modele de prise de decision.

Une des manieres d'aborder le domaine du raisonnement qualitatif est via l'arithmetique d'intervalle. Alors que dans l'approche mathematique conventionnelle une variable aura une valeur precise determinee par une fonction, dans l'arithmetique d'intervalle, les limites inferieure et superieure permettront de determiner l'etendue possible que pourra prendre cette variable (RAGHUNATHAN, 1994, p. 79). Ainsi, aux qualificatifs "froid", "tiède", "chaud", "brulant" sera

¹Par exemple, un electricien mis en presence d'un cable electrique qui surchauffe s'aura, bien que ne connaissant pas a fond toutes les lois de la physique electrique et de la conductivite des materiaux, que la quantite d'energie transmise par le cable depasse la capacite reelle et que cela constitue un danger.

associe un intervalle de temperature froid(-15,5), tiède(6,30), chaud(31,70), brulant(71,100) qui permettra de connaître de façon approximative l'état du système et de prendre une décision. En transport des matières dangereuses, une telle méthode permettrait d'associer à chaque tronçon de route non pas une probabilité précise d'occurrence d'accident mais plutôt un "degré de danger" ayant une étendue plus ou moins considérable selon les informations disponibles. Une plus ou moins grande incertitude dans les données s'exprimerait alors par un intervalle arithmétique plus ou moins grand.

L'arithmétique d'intervalle permet également de traiter les variables à partir de règles de structure (structure rules) obéissant aux lois de bases de l'arithmétique. Le solutionnement d'un problème passe donc par l'inférence comme pour les mathématiques conventionnelles. Parmi les nombreuses règles sur lesquelles s'appuie cette mécanique d'inférence sur les intervalles on retrouve (RAGHUNATAN, 1994, p. 80):

1)Règle d'intervalle:

Quand l'arithmétique d'intervalle génère plus d'un intervalle pour une variable ou une fonction, le plus contraignant des intervalles est retenu;

$$A = B + C$$

$$\hat{A} = D + E$$

$$6 \leq B \leq 30, 4 \leq C \leq 10, 3 \leq D \leq 10, 2 \leq E \leq 20$$

En utilisant l'arithmétique de base on peut démontrer que "A" se voit assigner les intervalles (10,40) et (5,30). Donc dans le processus de raisonnement qualitatif, l'intervalle (10,30) sera retenu comme valeur de "A". Dans le cas ou 2 intervalles n'ont pas de région commune, il y aura infaisabilité.

2)Règle de transitivité:

La règle de transitivité permet la dérivation de nouvelles contraintes à partir de contraintes connues. Par exemple, si le modèle de raisonnement indique que:

$$A = B \text{ et } B \geq C$$

On peut donc en conclure que l'intervalle $A \geq C$

3) Règle d'infaisabilité:

Un modèle est infaisable si deux intervalles ou plus n'ayant pas de région commune sont dérivés d'une fonction ou assignés à une variable.

$$0 \leq X \leq 1 \quad X(0,1)$$

$$0 \leq Y \leq 2 \quad Y(0,2)$$

$$0 \leq Z \leq 3 \quad Z(0,3)$$

$X + (Y \cdot Z) \geq 10 \dots$ Ce modèle est infaisable puisqu'égal à (0,7) donc < 10 .

4) Intervalle illimité:

Un modèle est illimité (unbounded) si une variable ou une fonction prend une valeur (∞, ∞) ou $(-\infty, -\infty)$.

$$\text{Maximiser } X + Y - 2Z$$

sujet à:

$$2 \leq X \leq 10$$

$$0 \leq Z \leq 5$$

$$Y \geq 0$$

Dans cet exemple, "X" a un intervalle (2,10), "Z" un intervalle (0,5) et "Y" (0, ∞). Selon l'arithmétique d'intervalle, l'intervalle de cette fonction est donc (2, ∞). Comme l'objectif du modèle est de maximiser (2, ∞) la fonction est illimitée.

Plusieurs autres règles analogues à celles-ci ont été développées pour mettre en pratique les principes du raisonnement qualitatif (KUIPERS, 1994 et WERTHNER, 1994). Le bénéfice le plus évident de cette approche est de conserver le pouvoir d'inférence des mathématiques tout en offrant plus de flexibilité dans la manipulation des données incertaines ou manquantes.

C'est d'ailleurs au traitement de l'incertitude et des données floues ou manquantes que les systèmes experts et le raisonnement qualitatif sont le plus en mesure de contribuer.

5.3.2 Raisonnement qualitatif: incertitude et ensembles flous

L'incertitude peut se définir comme un manque d'information adéquate pour prendre une décision. Son traitement est donc essentiel dans tout système de prise de décision. Il n'y a pas de consensus clair sur la meilleure méthode de traitement de l'incertitude et parmi ces méthodes, on retrouve les probabilités classiques et bayésiennes abordées précédemment ainsi que la théorie mathématique de l'évidence de Dempster-Shaffer et la théorie de Shannon très populaire en physique, en télécommunication et en géographie.

Dans l'approche par arithmétique d'intervalle, les situations d'informations manquantes ou incertaines sont traduites sous forme d'intervalles plus ou moins grands. Si une valeur n'est pas connue précisément, l'intervalle sera accru jusqu'à ce que le symbole qualitatif (exemple: "fort") représente l'ensemble des valeurs numériques possibles de ce qui est "fort". Les limites d'intervalle forment des frontières clairement définies et toute valeur située à l'extérieur immédiat des limites en est exclue. Dans cette approche, la limite entre ce qui est "fort" et "très fort" est précisément défini et il n'y a aucune gradation possible à l'intérieur même d'un intervalle. Pour un intervalle $fort(0, 10)$, des valeurs numériques de 1 ou de 9 auront toutes deux le qualificatif "fort" sans égard au degré de la force. L'arithmétique d'intervalle se prête également plutôt mal à la situation inverse où une valeur numérique connue peut être représentée par deux valeurs qualitatives. C'est en partie pour répondre à ces lacunes que la théorie des *ensembles flous* et de la *logique floue* a été développée (ZADEH, 1965).

Un ensemble flou peut se définir comme une classe d'éléments ayant des limites d'intervalles flottantes (ROLSTON, 1988, p. 107). Chaque élément d'un ensemble flou est identifié par un degré d'appartenance (grade of membership), variant entre 0 et 1, indiquant dans quelle mesure il est membre de l'ensemble. Un degré d'appartenance de 1 indique que l'élément est assurément un membre de l'ensemble alors qu'un degré de 0 indique qu'il n'en est pas membre. La transition entre être membre de l'ensemble et ne pas l'être est graduelle plutôt qu'abrupte comme dans le cas de l'arithmétique d'intervalle. Par exemple, la fonction d'appartenance à l'ensemble "jeune homme" pourrait prendre l'allure de la figure 5.2a où l'on constate que l'appartenance la plus élevée (et la plus plausible) à cette classe se situe aux alentours de 25 ans et que ce degré d'appartenance diminue dès que l'on s'éloigne de ce point central².

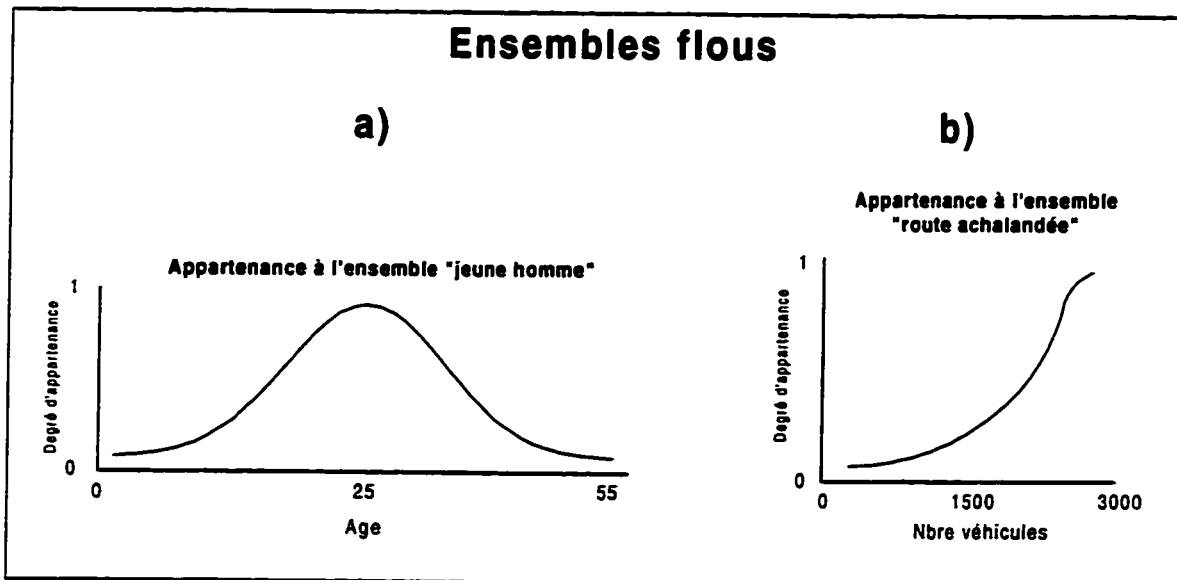


Figure 5.2: Ensembles flous

²A cinquante ans, on n'est définitivement plus un jeune homme, du moins, pas plus qu'à dix ans!

De même, en transport, on peut illustrer le degré d'appartenance à l'ensemble "route achalandée" par la fonction de la figure 5.2b.

Généralement, si aucune fonction mathématique n'est disponible, le degré d'appartenance à un ensemble est déterminé de façon subjective par un expert et correspond à la "crédibilité d'appartenance" d'un élément à une classe donnée. Quant à la *logique floue*, le seul aspect qui permet de la différencier de la logique conventionnelle est son application aux ensembles flous par opposition aux valeurs binaires usuelles. Ceci nous permet d'affirmer que, dans un processus de prise de décision à logique floue, seulement les éléments manipulés sont flous; les règles de logique, elles, sont bien définies (ROLSTON, 1988, p. 108).

Les pages précédentes ont établi que la théorie de la décision repose sur des approches quantitatives et qualitatives. Que ce soit par des méthodes conventionnelles ou des approches telles que l'arithmétique d'intervalle ou les ensembles flous, le traitement des informations nécessaires au transport des matières dangereuses passe par la capacité de manipuler les données manquantes ou incertaines et par la possibilité d'inclure les connaissances et le raisonnement d'experts dans le modèle de prise de décision. Les systèmes experts sont en mesure de répondre à ces besoins.

5.4 Les systèmes experts

Les systèmes experts sont nés du cadre de recherche plus vaste qu'est celui de l'intelligence artificielle. Les premières recherches en intelligence artificielle datent de la fin de la Seconde Guerre mondiale et les premières expériences ont porté sur le langage et plus particulièrement

la traduction du russe à l'anglais (GUIGO, 1995, p. 16). Comme le démontre la figure 5.3, les systemes experts ne sont que l'un des domaines de recherche de l'intelligence artificielle qui s'intéresse aussi à la reconnaissance de voix, la vision par ordinateur, la robotique, etc.

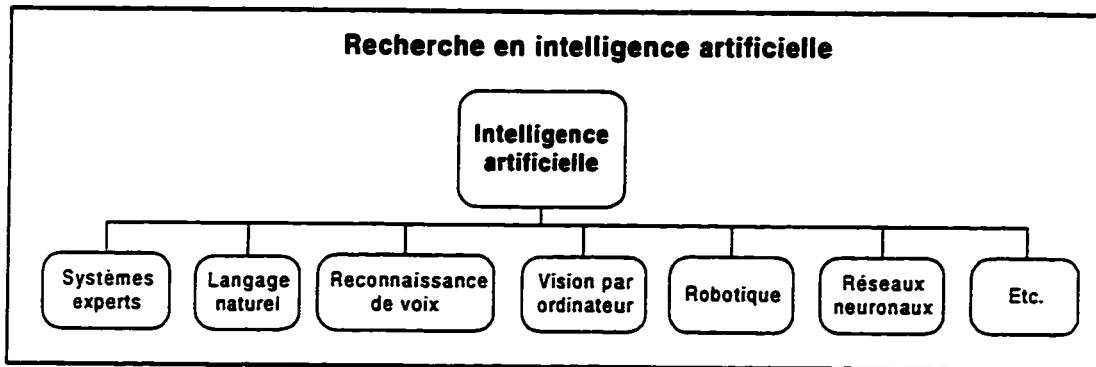


Figure 5.3: Recherche en intelligence artificielle

Le premier vrai systeme expert fut DENDRAL (Université Stanford) au milieu des années 60 dans le domaine de la chimie, puis vinrent MYCIN (Université Standford) pour les diagnostics médicaux et PROSPECTOR (Université Standford) pour la prospection de l'or (FRENZEL, 1987, p. 23). Ce n'est cependant qu'au début des années 80 que les premiers systemes experts ont été commercialisé sous forme de "coquille de systeme expert" (expert system shell); c'est-à-dire un systeme permettant à l'utilisateur de développer sa propre application.

5.4.1 Approche conventionnelle et approche par intelligence artificielle

La très grande majorité des logiciels développés pour solutionner un problème à partir d'un ordinateur utilisent l'approche dite conventionnelle dans laquelle un algorithme définit de façon rigide et précise les étapes et les actions devant être prises pour solutionner le problème. Les

données sont traitées par l'unité centrale selon une procédure bien définie et le comportement du système peut-être connu en tout temps (figure 5.4).

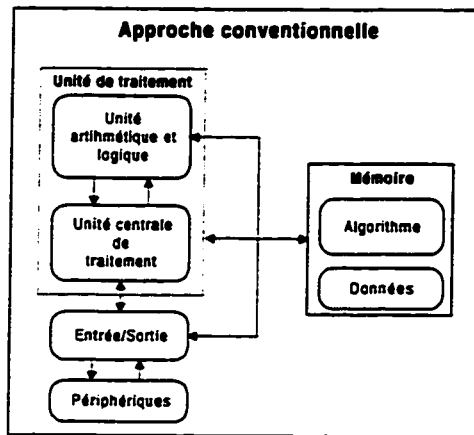


Figure 5.4: Approche conventionnelle de traitement

L'approche par intelligence artificielle ou par système expert est totalement différente puisqu'elle s'appuie sur les connaissances du domaine d'application. Ces connaissances peuvent être des faits, des règles, des relations, etc., représentés par des symboles dans la base de connaissances. Un moteur d'inférence traitera la base de connaissances pour en extraire la chaîne logique permettant de résoudre le problème (figure 5.5).

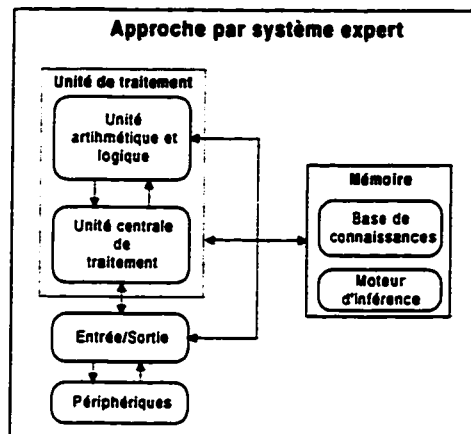


Figure 5.5: Approche par système expert du traitement

Systèmes experts et théorie de la décision

La méthode conventionnelle s'attarde à traiter les données pour produire un résultat précis tandis qu'un système expert traite les connaissances symboliques selon une série de règles logiques, de façon similaire au raisonnement humain lors de la prise d'une décision. Alors qu'un algorithme traditionnel pourra traiter tout genre de données, le système expert est intimement lié à la base de connaissances qui est spécifique à chaque domaine d'application. Les règles d'inférence menant à un diagnostic médical ne sont pas transposables en prospection minière! En d'autres termes, les systèmes experts sont limités au domaine spécifique de leur application (domain bounded).

Bien qu'ils permettent un raisonnement analogue à celui d'un être humain, les systèmes experts présentent certaines forces et certaines faiblesses qui en déterminent les capacités. EDMUNDS (1988, p. 41) donne quelques exemples de ce que peuvent et ne peuvent pas accomplir les systèmes experts:

Les systèmes experts peuvent:

- Manipuler un grand nombre de faits et de règles d'inférence
- Interroger l'utilisateur pour obtenir plus de précision
- Expliquer comment il en est arrivé à une conclusion
- Justifier son raisonnement
- Traiter les données incomplètes ou incertaines
- Traiter les informations symboliques
- Considérer plusieurs solutions alternatives
- Prioriser une ou des solutions alternatives

Les systèmes experts ne peuvent pas:

- Reasonner de façon vague sur un grand nombre de sujets
- Reasonner à partir d'axiomes
- Utiliser le "sens commun"
- Reasonner au-delà du contenu de la base de connaissances
- Accomplir un jugement biaisé
- Sauter prématurément à une conclusion

Cette simple énumération nous permet de croire que bon nombre d'applications pourraient trouver place en géographie, notamment en analyse spatiale et dans le domaine des systèmes d'information géographique. Quant au transport des matières dangereuses et au déplacement dans les réseaux en général, la possibilité de considérer plusieurs itinéraires alternatifs, de prioriser un itinéraire, de traiter les informations symboliques, incomplètes ou incertaines ainsi que de justifier le choix d'un itinéraire rend les systèmes experts, couplés à un système d'information géographique, la solution la plus adéquate actuellement.

5.4.2 Structure d'un système expert

Pour parvenir à appliquer des règles de déduction logique et à prendre des décisions, les systèmes experts ont besoin d'un certain nombre de fonctionnalités ou modules pour combiner règles et faits et tirer des conclusions. Selon FRENZEL (1987) et GUIGO (1995), un système expert typique comporte les cinq modules suivants (figure 5.6):

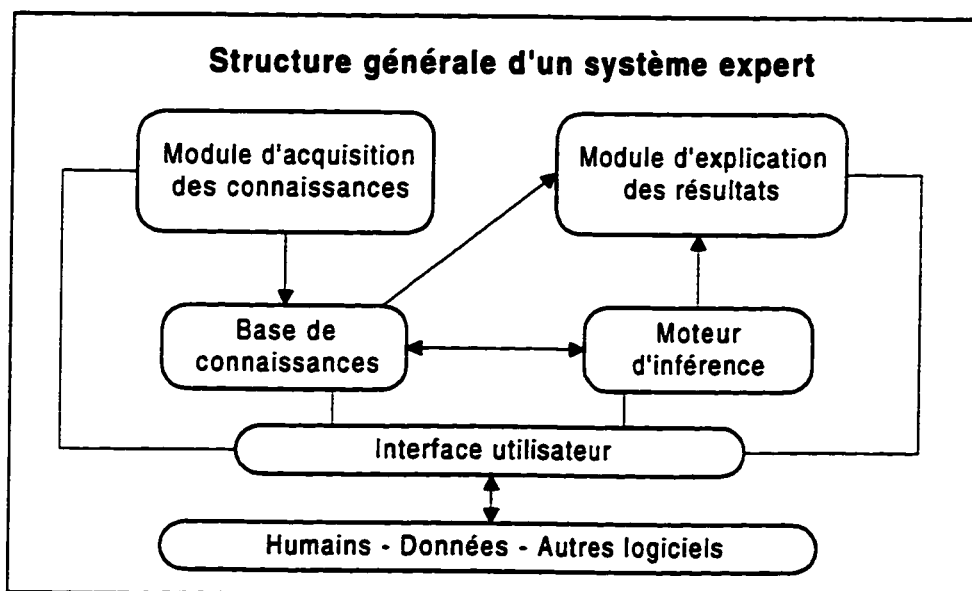


Figure 5.6: Structure générale d'un système expert

Une *base de connaissances* constituée de faits, de classes, d'objets, de règles représentant l'expertise du domaine d'application. Un *moteur d'inférence* dont la tâche est de contrôler le processus de raisonnement conduisant à une conclusion. Pour parvenir à une conclusion, le moteur d'inférence prend en compte la structure de la connaissance (base de connaissances) et applique les règles aux faits. Plusieurs systèmes experts comportent également un *module d'explication des résultats* qui présente à l'utilisateur la chaîne de raisonnement logique ayant conduit à une conclusion donnée. L'utilisateur peut alors vérifier la logique avant d'accepter la conclusion. Une *interface utilisateur* agit à titre d'intermédiaire de communication entre l'utilisateur humain et le système expert. Cette interface a aussi pour fonction de permettre la communication avec des bases de données externes ainsi que d'autres logiciels tels que des logiciels de statistique. Enfin, la force d'un système résidant dans son bassin de connaissances, un *module d'acquisition des connaissances* permet à l'utilisateur d'ajouter ou de modifier des connaissances et des règles au système.

Bien que chacun de ces cinq modules ait un rôle à jouer dans le fonctionnement d'un système expert, ce sont la base de connaissances et le moteur d'inférence qui constituent le noyau d'un tel système. La représentation des connaissances ainsi que leur interprétation ont un impact majeur sur la qualité d'un système expert.

5.4.3 Représentation des connaissances et inférence

Dans le domaine particulier des systèmes experts, les connaissances peuvent se définir comme la compréhension d'un champ spécifique, compréhension acquise par l'éducation et l'expérience sur le terrain. Alors que l'information correspond à des données brutes auxquelles

on intègre un sens particulier, les connaissances sont faites d'idées, de concepts, de faits, de théories et de relations entre ces divers éléments (FRENZEL, 1987, p. 4). Bref, les connaissances sont faites de tout ce qu'un expert aura besoin pour solutionner un problème.

On peut distinguer deux grands types de connaissances dans un système expert. Les *connaissances déclaratives* qui énoncent un fait sur un élément donné et qui permettent de déduire des relations, de classer des objets ou de décrire des situations. Les connaissances déclaratives définissent des connaissances statiques ou descriptives sur le domaine considéré (GUIGO, 1995, p. 30). Par exemple, les énoncés suivants constituent des connaissances déclaratives:

- Les routes sinueuses sont dangereuses
- Les virages à gauche augmentent les risques d'accident
- La neige est blanche
- etc.

Les *connaissances procédurales* forment le deuxième type de connaissances et représentent le savoir-faire d'un expert, son processus de raisonnement ou sa façon d'agir. Contrairement aux connaissances déclaratives qui décrivent une situation, les connaissances procédurales indiquent comment interpréter les faits et quelles actions prendre. Ces connaissances sont du genre:

- Si il y a présence d'un hôpital, éviter le tronçon routier
- Si route achalandée, éviter heure de pointe
- Si chargement de gaz toxique, éviter centre-ville
- etc.

Dans la base de connaissances d'un système expert, les connaissances procédurales se distinguent par leur représentation sous forme de *règles de production*. Les règles de production sont des énoncés conditionnels sous la forme "IF-THEN" et constituent les "unités

de raisonnement" du systeme. Ces regles se presentent en deux parties: une premisses (situation) et une conclusion (action) (GUIGO, 1995, p. 30).

SI «situation» ALORS «action»
IF «route sinueuse» THEN «eviter tronçon»

Chaque fois qu'une situation est connue, l'action doit être exécutée. Chaque règle de production est modulaire et exprime un morceau indépendant de connaissances. La situation correspond à l'hypothèse ou au membre de gauche de la règle et l'action à la conclusion, au membre de droite de la règle. La prémisses ou la conclusion d'une règle de production peuvent être composées de sous-ensembles pour définir une connaissance plus complexe:

IF «présence d'une école»
AND «vacances estivales»
AND «entre minuit et 6:00 AM»
THEN «école vide»
AND «tronçon sécuritaire»

Les petits systemes experts destinés aux applications simples contiennent généralement une dizaine ou une vingtaine de règles de production alors que les applications plus complexes requerront bien au-delà d'une centaine de règles (FRENZEL, 1987, p. 75). L'avantage des règles de production est leur grande modularisation des connaissances qui facilite la création et la mise-à-jour de la base de connaissances. Les connaissances y sont cependant représentées pêle-mêle, en unités élémentaires et sans lien logique entre elles, ce qui a pour effet de rendre plus ardue et plus lente la tâche de reconstruction du raisonnement logique menant à une conclusion. La présence de plusieurs dizaines de règles peut parfois conduire à la redondance ou encore à une certaine ambiguïté. Bien que les systemes experts à base de règles de production demeurent les plus répandus, d'autres modes de représentation des connaissances ont été développés en intelligence artificielle.

Les réseaux sémantiques (semantic network) qui consistent en un ensemble de noeuds reliés par des arcs sont une alternative aux règles de production (figure 5.7). Chaque noeud de ce réseau représente un objet ou un descripteur et chaque arc une relation entre l'objet et le descripteur (GILLIES, 1991, p. 27).

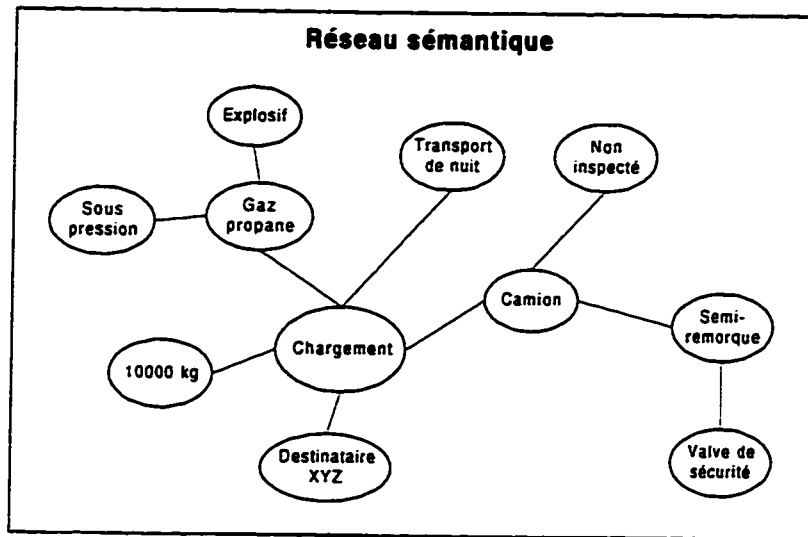


Figure 5.7: Représentation par réseau sémantique

Cette structuration graphique se prête bien à la représentation des connaissances déclaratives et permet d'accéder facilement à l'ensemble des connaissances pour un élément donné. Les réseaux sémantiques sont similaires à la mémoire humaine en ce sens qu'ils stockent les connaissances par associations d'idées et de concepts. Le problème des réseaux sémantiques réside dans l'addition et la mise à jour de la base de connaissances. Contrairement aux règles de production, les réseaux sémantiques ont une structure plutôt rigide qui se prête mal à l'édition; la structure du réseau impose également une "direction" dans le processus d'inférence en raison des liens qu'elle renferme.

L'approche orientée objet (object oriented) a aussi gagné en popularité depuis quelques années. Cette approche est similaire aux réseaux sémantiques en ce sens qu'elle consiste à structurer les connaissances autour des objets manipulés et non autour des tâches à accomplir (GUIGO, 1995, p. 31). L'unité de stockage et de structuration des connaissances est l'*objet structuré* (frame) qui se divise en *champs* (slots) regroupant les connaissances interreliées. On parle alors d'*encapsulation*. Les champs peuvent contenir une valeur unique ou un ensemble de valeurs. Ces valeurs peuvent être numériques, symboliques, logiques, graphiques ou textuelles (GILLIES, 1991, p. 29). Dans la représentation orientée-objet, les éléments du domaine de connaissances sont décomposés en classes, sous-classes, objets, sous-objets (figure 5.8).

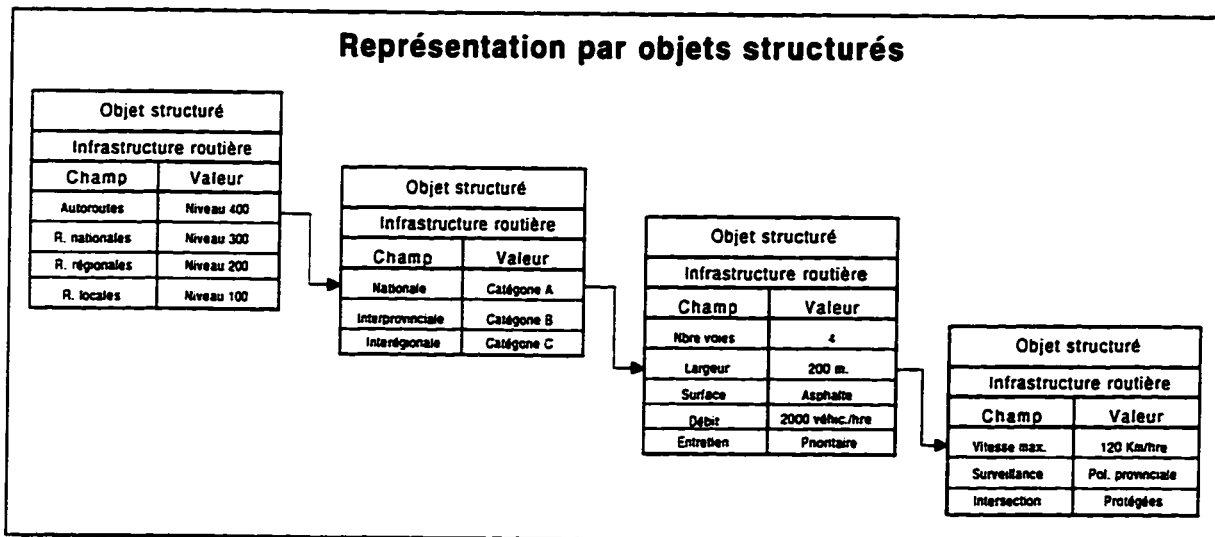


Figure 5.8: Représentation par objets structurés

Chaque classe est la description d'une famille d'objets ayant même structure et même comportement (Exemple: classe autoroutes, fig. 5.8). Chaque objet correspond à une unité élémentaire de description et est une instance d'une classe (Exemple: l'objet "nbre de voies"

est une instance de la catégorie "A"). Classes et objets peuvent se diviser en sous-classes et en sous-objets qui hériteront des propriétés des classes supérieures (inheritance). Un objet structuré "enfant" hérite de toutes les propriétés d'un objet structuré "parent" (GILLIES, 1991, p. 29). Ainsi, les routes à vitesse maximale de 120 km/hre héritent des propriétés: routes à 4 voies, de catégorie "A" et de niveau 400. L'intérêt de la représentation orientée-objet se situe dans sa division des connaissances en parties indépendantes regroupant les objets par affinités. On peut y mélanger connaissances déclaratives et connaissances procédurales sans distinction et la mise à jour des connaissances est plus facile que dans le cas des réseaux sémantiques.

D'autres méthodes de représentation et de traitement des connaissances ont également été développées en intelligence artificielle. C'est notamment le cas des *réseaux neuronaux* (neural nets) dont l'objectif est de reproduire les processus cognitifs du cerveau humain. Les réseaux neuronaux sont donc des systèmes capables "d'apprendre" de leurs expériences. Cependant, le fonctionnement de ces systèmes est fort complexe et nécessite une capacité de traitement que seul les ordinateurs à architecture parallèle (transputer) sont en mesure d'offrir. Intégrer une telle approche aux modèles d'optimisation itératifs du plus court chemin est donc au-delà des capacités des ordinateurs actuels.

C'est le moteur d'inférence du système expert qui a pour tâche d'exploiter les connaissances stockées dans la base de connaissances. A partir de cette dernière, il est en mesure d'effectuer des déductions logiques, de produire un raisonnement de façon analogue à un expert. En utilisant les règles de production, le moteur d'inférence enchaîne les règles les unes aux autres

de façon à simuler un raisonnement. Un moteur d'inférence peut produire trois grands types de raisonnement: le *chaînage avant*, le *chaînage arrière* et le *chaînage mixte*. Le chaînage avant consiste à déduire à partir des faits observés (la prémisse des règles de production) de nouveaux faits pour aboutir à une hypothèse terminale (conclusion). Le mécanisme d'enchaînement suit alors la logique: "Puisque cet ensemble de connaissances est vrai (prémisse), alors les conclusions que l'on en tire sont vraies". Le chaînage arrière consiste à partir d'une hypothèse déjà établie et à vérifier les conditions de cette hypothèse. La démarche logique est alors du type: "Si cette connaissance est vraie alors les connaissances qui l'impliquent doivent l'être aussi". Dans le cas du chaînage mixte, les deux types de chaînage (avant et arrière) peuvent être invoqués pour atteindre le but.

Chaque fois que le moteur d'inférence exécute une règle, il modifie l'état de la pile de connaissances, de nouveaux faits s'ajoutent et d'autres sont modifiés. Il y a alors accomplissement d'un *cycle d'inférence*. Si ces nouveaux faits correspondent à des conditions de la partie gauche d'une règle, un nouveau cycle d'inférence sera déclenché, ainsi de suite jusqu'à ce que l'ensemble des règles de production aient été vérifiées. Il apparaît donc évident à ce niveau que la structuration des connaissances dans la base de connaissances aura un impact direct sur la démarche logique du moteur d'inférence.

5.4.4 Traitement de l'incertitude dans un système expert

Nous avons pu constater précédemment que l'incertitude peut être abordée, du point de vue du raisonnement qualitatif, par l'arithmétique d'intervalle ou les ensembles flous. Dans ce dernier cas, le degré d'appartenance à un ensemble détermine la crédibilité ou la véracité d'un fait

(exemple: la véracité d'appartenir à l'ensemble "jeune homme"). La plupart des systèmes experts traitent l'incertitude de façon analogue à la théorie des ensembles flous, à l'aide d'un *facteur de certitude* (certainty factor).

Le concept de facteur de certitude a été développé pour le système expert MYCIN pour palier aux problèmes de l'approche probabiliste traditionnelle et bayésienne en ce qui concerne la quantification des probabilités à priori et conditionnelles (ROLSTON, 1988, p. 99). Un facteur de certitude est une valeur numérique qui exprime le degré auquel, basé sur un ensemble d'évidences, on doit croire en une conclusion, une règle de production ou une connaissance. Un facteur de certitude de 1 indique une parfaite croyance alors qu'un facteur de -1 indique un rejet total (ROLSTON, 1988, p. 100). Un facteur de certitude peut être établi à partir d'une fonction mathématique ou de l'inférence statistique si les informations nécessaires à la quantification sont disponibles, ou encore être établi de façon subjective à partir du jugement et de l'intuition d'un expert si aucune quantification n'est possible.

Le traitement des règles de production et des facteurs de certitude par le moteur d'inférence lors du chaînage aura pour effet de produire une *mesure de croyance* en la conclusion finale ainsi qu'une *mesure de rejet* de cette même conclusion. La mesure de croyance $MB[c,e]$ indiquera le degré de croyance en la conclusion "c" basée sur l'évidence "e". Par définition (ROLSTON, 1988, p. 100):

$$0 \leq MB[c, e] \leq +1$$

De facon similaire, la mesure de rejet $MD[c,e]$ indiquera le degre de rejet de la conclusion "c" basee sur l'evidence "e". Par definition:

$$-1 \leq MD[c, e] \leq 0$$

La base de connaissances d'un systeme expert se doit donc d'etre structuree de facon a ce que chaque regle de production ajoute soit au degre de croyance, soit au degre de rejet d'une conclusion. Un facteur de certitude de 0 ne contribue d'aucune maniere au processus d'inférence du systeme.

Généralement, lorsqu'un systeme expert en arrive a une conclusion, plusieurs regles de production auront été chainées, de facon a recréer un raisonnement logique. Etant donné que chacune de ces regles de production possède un facteur de certitude, il est possible de calculer un *facteur de certitude cumulée* indiquant le degre de croyance ou de rejet de l'ensemble du raisonnement. Ce facteur de certitude cumulée correspond a la combinaison de la *mesure de croyance cumulative* et de la *mesure de rejet cumulative* pour la conclusion (ROLSTON, 1988, p. 101):

$$FC[c, e_c] = MB[c, e_f] - MD[c, e_a]$$

où:

- $FC[c, e_c] =$ Le facteur de certitude cumulative pour la conclusion "c" étant données les évidences cumulées "e_c".
- $MB[c, e_f] =$ Mesure de croyance cumulative de la conclusion "c" étant donnée les évidences cumulées "e_f".
- $MD[c, e_a] =$ Mesure de rejet cumulé de la conclusion "c" étant données les évidences cumulées "e_a" **contre** la conclusion "c".

Cette definition implique la necessite de calculer les valeurs de MB et MD pour chacune des conclusions au probleme en cause. La chaine de raisonnement menant a la conclusion ayant un facteur de certitude cumule le plus eleve sera retenue comme solution au probleme. Par exemple, dans la situation suivante:

Règle de production "A": **IF «présence d'une école»**
 AND «vacances estivales»
 AND «entre minuit et 6:00 AM»
 THEN «tronçon sécuritaire»
 (FC = 0.75)

Règle de production "B": **IF «route sinueuse»**
 AND «peu de circulation»
 THEN «route sécuritaire»
 (FC = 0.3)

$$\begin{aligned} FC_{\text{cumulé}} &= FC(A) + FC(B) - FC(A) * FC(B) \\ &= 0.75 + 0.3 - 0.75 * 0.3 \\ &= 1.05 - 0.225 \\ FC_{\text{cumulé}} &= 0.825 \end{aligned}$$

Bien que ce tronçon soit sinueux et renferme une école, la logique veut qu'il soit sécuritaire avec un facteur de certitude cumulée de 0.825. Il est à noter dans cet exemple que la règle de production "B" présente un lien de causalité moins robuste (0.3) que la règle "A" (0.75).

Il apparaît donc que non seulement les systèmes experts sont en mesure d'appliquer un raisonnement logique analogue à un expert humain mais qu'ils sont également capables de traiter l'incertitude et les données qualitatives (raisonnement qualitatif). Les chapitres 7 et 8 ont pour but de démontrer comment le couplage système expert - système d'information géographique pourrait être bénéfique au problème particulier du transport des matières

Systèmes experts et théorie de la décision

dangereuses, tandis que le prochain chapitre traite du problème du transport des matières dangereuses sous l'angle du bilan accidentel et de la notion de risque.

Chapitre VI

Transport des matières dangereuses et risques

“L’âge que nous vivons est dangereux; comme il serait ennuyeux s’il ne l’était pas.”

- John Steinbeck

La protection de l’environnement occupe une place importante dans l’opinion publique contemporaine et plusieurs événements récents nous forcent à reconnaître que la préoccupation environnementale fait désormais partie du discours et de la pratique populaire. La médiatisation de certaines grandes catastrophes écologiques et technologiques des dernières années, le phénomène du “pas dans ma cour” (NIMBY) et l’intérêt croissant accordé au transport des matières dangereuses en sont des exemples. A ce titre, la globalisation et la spécialisation des marchés de même que la différenciation spatiale marquée entre sites de production et sites de consommation ont contribué à accroître le volume de marchandises transportées d’un site à l’autre.

Transport des matières dangereuses et risques

Globalement, les matières dangereuses représentent environ entre 5 et 15% du tonnage total du transport terrestre (OCDE, 1988, p. 11). Aux États-Unis, on estime que 1.5 milliard de tonnes de matières dangereuses sont transportées annuellement (LIST *et al*, 1991, p. 100). Généralement, les statistiques sur ce type de transport fluctuent au rythme de l'économie des régions, ce qui démontre bien le lien étroit qui existe entre le transport des matières dangereuses, l'industrie et la vigueur économique en général. Il y a cependant fort à parier que la popularité croissante du camionnage ainsi que le foisonnement des produits toxiques et dangereux laissent présager pour le futur des risques de plus en plus diversifiés pour les usagers de la route, le public et l'environnement.

Au chapitre 6 nous aborderons, dans un premier temps, le transport des matières dangereuses au Canada. Nous verrons brièvement le cadre législatif ainsi que les principales caractéristiques et dangers potentiels des matières dangereuses transportées au pays. Dans un deuxième temps, nous nous attarderons à la notion de risque, aux différentes formes qu'il peut prendre ainsi qu'aux moyens de le mitiger.

6.1 Les matières dangereuses

Le transport des matières dangereuses est dans une large mesure un trafic franchissant les frontières, donc un problème de portée internationale (OCDE, 1988, p. 12). De nombreuses organisations internationales préparent et mettent à jour des règlements et recommandations sur ce problème dans le but d'uniformiser l'intervention entre pays. Parmi ces organisations, mentionnons notamment les *Nations Unies* qui ont pour rôle d'établir le cadre général de réglementation; l'*Agence Internationale de l'Énergie Atomique* concernant les matières

Transport des matières dangereuses et risques

radioactives; l'*Organisation Maritime Internationale* ainsi que l'*Organisation de l'Aviation Civile Internationale* pour le trafic maritime et aérien. La responsabilité du contrôle des mouvements de matières dangereuses revient cependant à chaque pays, en accord avec les principes généraux établis au niveau international. Au Canada, une matière dangereuse se définit comme "une substance dotée de caractéristiques physiques, chimiques ou biologiques qui impose des manipulations particulières de nature à éviter tout risque pour la santé ou tout effet nocif pour l'environnement" (JACOB et MONGRAIN, 1988, p. 115).

Outre les matières dangereuses, une grande quantité de déchets dangereux sont également transportés au pays. Selon le BAPE (Bureau d'Audience Publique sur l'Environnement), un déchet potentiellement dangereux est en "raison de sa nature (types de produits, degré de concentration, degré de contamination) et de sa quantité, une substance qui ne peut être traitée, recyclée, ou éliminée directement sans danger pour l'humain ou l'environnement" (BAPE, 1990).

La distinction entre matière dangereuse et déchet dangereux réside, d'une part, dans l'origine et la destination du produit qui, étant donné le caractère non recyclable des déchets, devrait avoir comme point de chute final un site de neutralisation ou d'élimination. D'autre part, cette distinction se situe également au niveau du degré de contamination du déchet; une substance originellement non dangereuse peut être considérée comme déchet dangereux si elle a subi un degré de contamination suffisamment élevé. De façon générale, on peut donc dire qu'une substance sera considérée comme étant une matière dangereuse si elle possède des qualités

intrinsèques la rendant dangereuse alors qu'elle sera considérée comme étant un déchet dangereux s'il s'agit d'un rejet ou d'une contamination.

6.2 Le transport des matières dangereuses au Canada

En raison de sa grande étendue, il n'est pas étonnant de constater que le transport de marchandises représente un élément essentiel de l'activité économique canadienne. Dans un rapport de Transport Canada (1986), on estimait que les envois par camion, par rail, par mer ou par air contenaient respectivement 15%, 12%, 20% et 2% de matières dangereuses. Dans son rapport sur le camionnage au Canada en 1991, Statistique Canada affirme qu'en terme de poids, le transport routier accapare 46% de l'expédition des marchandises dangereuses alors que les expéditions ferroviaires et maritimes représentaient respectivement 31% et 23% (Statistique Canada, 1994, p. 150). Les liquides inflammables et combustibles sont les marchandises dangereuses les plus transportées avec 63% du poids total. Ils constituent 85% des expéditions maritimes, 78% des expéditions du camionnage et 24% des expéditions par chemin de fer (Statistique Canada, 1994, p. 154).

Comme dans les autres pays industrialisés, le transport routier accapare la majeure partie des expéditions de matières dangereuses à l'intérieur du Canada. Les expéditions à l'intérieur et entre les principaux centres urbains du pays représentaient 40% de la somme globale de marchandises transportées par camion en 1989 (Statistique Canada, 1994, p. 151). Le plus grand nombre de mouvement de marchandises dangereuses en 1989, soit 244 000, ont eu lieu dans la région métropolitaine de Toronto, suivie de près par Vancouver, deux des régions les

plus fortement peuplées du pays. A la lumière de ces données, il n'est donc pas étonnant de constater que les gouvernements aient eu à intervenir dans ce domaine.

6.2.1 Le cadre législatif

Dès 1972, une étude intergouvernementale sur la réglementation du transport routier avait démontré que les marchandises dangereuses n'étaient pas réglementées de façon adéquate au pays. Cependant, peu d'actions concrètes avaient été entreprises pour remédier à cette situation, probablement en raison de la complexité du domaine. A cette époque, 22 organismes fédéraux, provinciaux et territoriaux s'intéressaient de près ou de loin au transport des marchandises dangereuses (Transport Canada, 1986, p. 3).

C'est au moment où le gouvernement fédéral s'afférait à élaborer une réglementation plus sévère en matière transport des marchandises dangereuses qu'un événement catalyseur est venu accélérer le processus et alerter l'opinion publique. En 1979, un train transportant du propane et du chlore déraillait, entraînant l'évacuation de 250 000 personnes dans la région de Mississauga, en Ontario (The Toronto Area Rail Transportation of Dangerous Goods Task Force, 1988, p. 59). La même année, le centre d'intervention d'urgence CANUTEC (Centre Canadien d'Urgence Transport) débutait ses opérations d'information et de communication 24 heures par jour pour répondre aux situations d'urgence en transport. Ce centre a joué et joue encore un rôle clef dans la collecte et la constitution de bases de données relatives aux matières dangereuses. Sa base de données contient actuellement au-delà de 250 000 entrées sur lesquelles les usagers peuvent s'informer (Transport Canada, 1991).

Transport des matières dangereuses et risques

En 1980, suite aux événements de Mississauga, la *Loi sur le transport des marchandises dangereuses* était adoptée puis, en 1985, le *Règlement sur le transport des marchandises dangereuses* venait encadrer à l'échelle nationale le transport de ce type de marchandises. Les principales dispositions de la loi prévoient entre autres (Transport Canada, 1986, p. 4):

L'établissement d'une seule loi visant la manutention, la demande de transport et le transport des marchandises dangereuses par tous les moyens de transport au Canada.

L'application à toutes les activités de transport et à tous les modes.

La désignation par le ministre d'inspecteurs autorisés à pénétrer et à visiter des locaux ou des lieux où s'effectuent la manutention ou le transport des marchandises dangereuses.

La mise au point d'un ensemble de dispositions réglementaires portant sur les normes de sécurité, les règles de sécurité et les indications de danger, plus particulièrement en ce qui concerne l'identification et la classification des marchandises dangereuses.

Malgré le caractère national de cette loi et la désignation d'inspecteurs fédéraux agissant sous son pouvoir, l'intervention dans ce domaine demeure une responsabilité partagée. Plusieurs provinces possèdent leurs propres règlements sur le transport des matières dangereuses, règlements qui rejoignent à la base ceux du gouvernement fédéral. Au Canada cependant, la responsabilité des municipalités occupe une place centrale puisque ces dernières ont la charge d'établir les plans d'urgence et d'intervenir en cas de catastrophe. La désignation d'itinéraires pour le transport des matières dangereuses est donc un pouvoir municipal ou régional au même titre que l'élaboration des plans d'urgence. Les municipalités ne sont toutefois pas

Transport des matières dangereuses et risques

laissées à elles-mêmes dans cette tâche puisque *Protection Civile Canada* fournit formation et expertise (via son Collège d'Amprior) aux municipalités qui le désirent.

Toutefois, il faut admettre que cette concentration de responsabilités aux mains des municipalités, qui souvent n'ont pas les ressources nécessaires, rend très inégale d'un territoire à l'autre la planification des mesures d'urgence et du transport des matières dangereuses. Au Québec par exemple, dans un mémoire remis au BAPE en 1990, l'UMRCQ (Union des Municipalités Régionales de Comté du Québec) faisait état d'un sondage effectué auprès des municipalités du Québec indiquant que trois municipalités sur quatre ne connaissaient pas la situation des déchets dangereux sur leur propre territoire. Ce sondage indiquait également que 65% des répondants disent être au courant du va-et-vient des matières dangereuses sur leur territoire et que pour 73% des répondants, ces matières empruntent les artères les plus densément peuplées. La capacité des municipalités à réagir en cas de catastrophe est aussi limitée puisque seulement 46% d'entre elles ont affirmé avoir un plan d'urgence. Il y a fort à parier que cette situation n'est pas particulière au Québec et que d'autres provinces sont dans le même état de désorganisation. Il faut cependant préciser que la très grande majorité des grandes villes canadiennes possèdent un plan d'urgence, contrairement aux petites municipalités qui tardent à emboîter le pas.

6.2.2 La classification des matières dangereuses

Un des effets positifs de la *Loi sur le transport des marchandises dangereuses* a été la mise sur pieds d'organismes dont la tâche est l'inventaire et la classification des marchandises dangereuses. Avant de faire transporter une matière dangereuse, l'expéditeur doit vérifier si

Transport des matières dangereuses et risques

l'appellation individuelle de cette matière apparaît aux listes en annexe du Règlement sur le transport des marchandises dangereuses. Lorsque le produit y est répertorié, ces listes fournissent l'appellation réglementaire, le numéro d'identification, la classification, le type d'emballage ainsi que d'autres précisions concernant les exigences de sécurité propres à la matière transportée.

Classe 1	Explosifs (Matières comportant un danger d'explosion en masse ou en fragments ou d'explosion par effet de souffle)
Classe 2	Gaz comprimés (Gaz inflammables, ininflammables, toxiques ou corrosifs)
Classe 3	Liquides inflammables (Liquides comportant des dangers d'incendie, d'explosion, de toxicité, de corrosion et de contamination de l'eau. Exemple: essence, méthanol, kérosène, etc.)
Classe 4	Solides inflammables (Matières sujettes à l'inflammation spontanée ou qui au contact de l'eau ou de l'air dégagent des gaz inflammables)
Classe 5	Matières comburantes et peroxydes organiques (Matières sensibles à la chaleur, au choc et au frottement. Peuvent former des mélanges explosifs)
Classe 6	Matières toxiques et matières infectieuses (Solides ou liquides toxiques par inhalation, ingestion, ou par absorption cutanée de leurs vapeurs. Organismes ou toxines connus comme étant infectieux ou susceptibles de présenter un danger d'infection.)
Classe 7	Matières radioactives (Danger en raison de l'exposition au rayonnement ou en cas de fuite de matières radioactives)
Classe 8	Matières corrosives (Matières pouvant causer une nécrose de la peau et matières qui corrodent les métaux. Exemple: acide sulfurique).
Classe 9	Matières ou produits divers (Non-classifiées, toxiques pour les humains ou pour l'environnement. Déchets dangereux).

Tableau 6.1: Classification des matières dangereuses

La classification canadienne se conforme au système international de classification des marchandises dangereuses de l'Organisation des Nations-Unies (ONU) et mis à part son usage

Transport des matières dangereuses et risques

par les expéditeurs, son utilité est majeure pour plusieurs raisons. Tout d'abord, les milliers de substances répertoriées, dont 3500 sont considérées comme dangereuses, nécessitent qu'une mise en ordre logique soit appliquée pour garder un accès facile aux bases de données. En second lieu, pour être en mesure de dresser un portrait plus juste des déplacements de ces matières, l'utilisation de catégories s'impose. Enfin, en cas d'accident, il est beaucoup plus facile de mettre sur pieds des mesures d'urgence en fonction de grandes catégories de produits que d'établir une procédure spécifique pour chaque substance dangereuse. Au Canada, la classification comporte neuf grandes classes ainsi que leurs subdivisions (Transport Canada, 1990, pp. 4-12). Le tableau 6.1 présente le contenu général de cette classification.

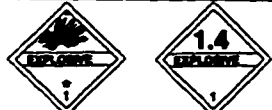





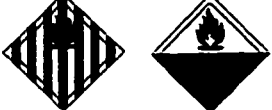


Exemples de pictogrammes					
Classe 1			Classe 6		
Classe 2			Classe 7		
Classe 3			Classe 8		
Classe 4			Classe 9		
Classe 5					

Figure 6.1: Pictogrammes d'identification des matières dangereuses

En plus de cette classification, toute marchandise dangereuse possède une plaque iconique en forme de losange permettant d'identifier le type de matière (selon les classes) ainsi que le type

de danger. Chaque plaque comporte un pictogramme ainsi qu'un numéro d'identification de 4 chiffres attribué par l'ONU pour chaque matière considérée comme dangereuse. Il est donc possible, à tout moment, de savoir de quel produit il est question et quels dangers il représente; ce qui facilite le travail des inspecteurs et des forces de l'ordre qui ont pour tâche de faire respecter les règlements.

6.3 Profil statistique des accidents

C'est à Transport Canada, à la *Direction générale du transport des marchandises dangereuses*, que revient la tâche d'inventorier les accidents impliquant des matières dangereuses. Depuis 1988, les accidents rapportés à la Direction sont intégrés à une base de données nationale, le "Dangerous Goods Accidents Information System" ou DGAIS. Bien que cette base de données ne contienne pas la totalité des cas d'accidents impliquant des matières dangereuses et que la plupart de ceux-ci n'aient pas fait l'objet d'une enquête particulière par Transport Canada, le DGAIS demeure une source de données intéressante pour l'étude de ce phénomène.

Pour la période 1988-1992, le DGAIS renferme plus de 90 variables décrivant divers aspects des accidents rapportés. Un premier groupe de variables s'attarde à décrire le type d'accident et le site où il s'est produit, le mode de transport, les causes probables de l'accident ainsi que les conditions météorologiques. Un deuxième groupe de variables a pour fonction de quantifier les conséquences de l'accident, c'est-à-dire le type et la zone de contamination, le nombre de blessés, de morts et de personnes évacuées. Le type de véhicule et les caractéristiques du chargement, le type de matière dangereuse ainsi que les quantités transportées et impliquées dans l'accident font partie du troisième groupe. Finalement, on retrouve également une

composante géographique, à savoir, la localisation de l'accident par province, par division et subdivision de recensement de même que l'origine et la destination du chargement. Il est à noter cependant, qu'en raison des nombreuses modifications qu'a subies cette base de données au cours de la période 1988-1992, très peu d'observations possèdent l'ensemble des 90 variables. Nous nous limiterons donc dans les pages qui suivent à une description générale à partir des variables recoupant le plus grand nombre d'observations.

6.3.1 Caractéristiques générales des accidents

Annuellement, le centre CANUTEC répond à environ 18 000 appels liés aux matières dangereuses. Évidemment, la majeure partie de ces appels concerne des demandes d'information et non des accidents. Pour la période 1988-1992, le nombre d'accidents impliquant des matières dangereuses, tous modes de transport confondus, a varié entre 702 et 1104 accidents pour un total de 4783 accidents (figure 6.2).

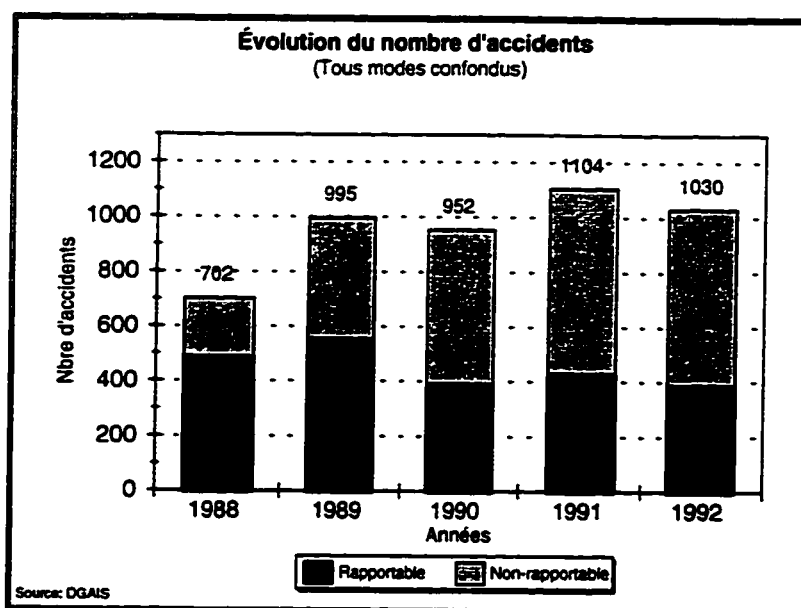


Figure 6.2: Évolution du nombre d'accidents.

Transport des matières dangereuses et risques

Si l'on fait exception de l'année 1988, année d'implantation de CANUTEC, le nombre annuel d'accidents a peu varié. La même situation se retrouve au niveau des accidents rapportables (rapportables selon la loi) qui eux ont varié entre 392 et 557 pour un total de 2269 accidents. Si l'on s'en tient uniquement aux accidents rapportables, cela représente une moyenne annuelle de 453 accidents, soit plus d'un accident par jour (1.25 accident/jour).

La répartition des accidents par mode de transport et par phase de transport révèle également des faits particuliers. En ce qui a trait aux phases de transport, la moitié des accidents (49.2%) ont lieu aux terminaux de transbordement, 18.2% dans les entrepôts, alors que 30.6% des accidents se produisent pendant la phase de transport des marchandises (figure 6.3).

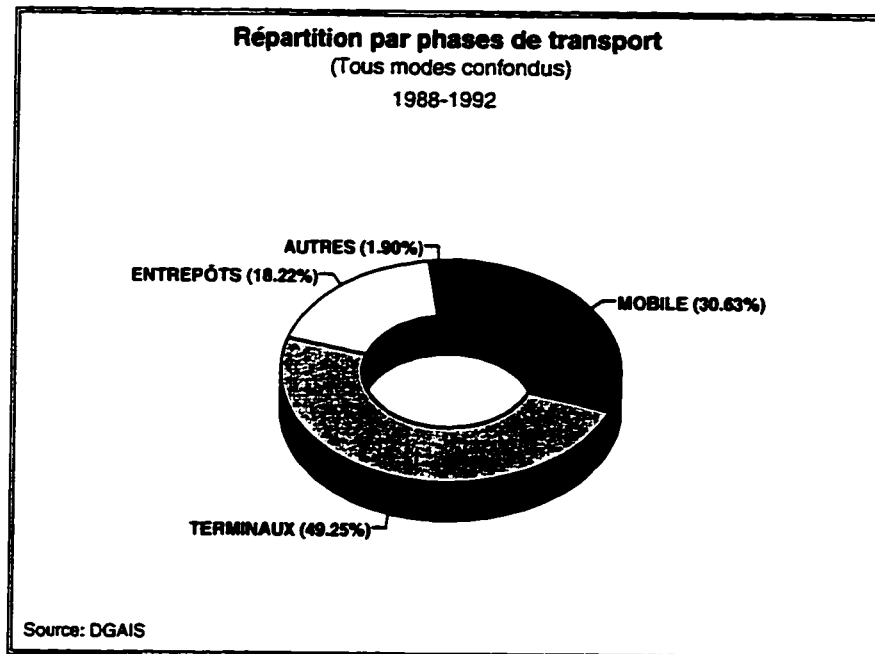


Figure 6.3: Répartition par phase de transport

Le taux élevé d'accidents aux terminaux de transbordement s'explique par le fait qu'il s'agit d'une phase cruciale de manutention où la possibilité d'un bris de matériel (valve, conteneur,

Transport des matières dangereuses et risques

etc) est plus élevée et par conséquent le nombre d'accidents aussi. C'est donc dire que, pour tous les modes de transport confondus, 30% des accidents se produisent lors du transport en tant que tel.

La répartition par mode de transport se fait aussi de façon inégale. Globalement, de 1988 à 1992, c'est le transport par rail qui arrive bon premier avec 44% des accidents suivi par le transport routier avec 32%. Le transport aérien et maritime ne représente que 2.1% et 1.6% respectivement (figure 6.4). La route et le rail s'accaparent donc de plus des trois quarts des accidents impliquant des matières dangereuses.

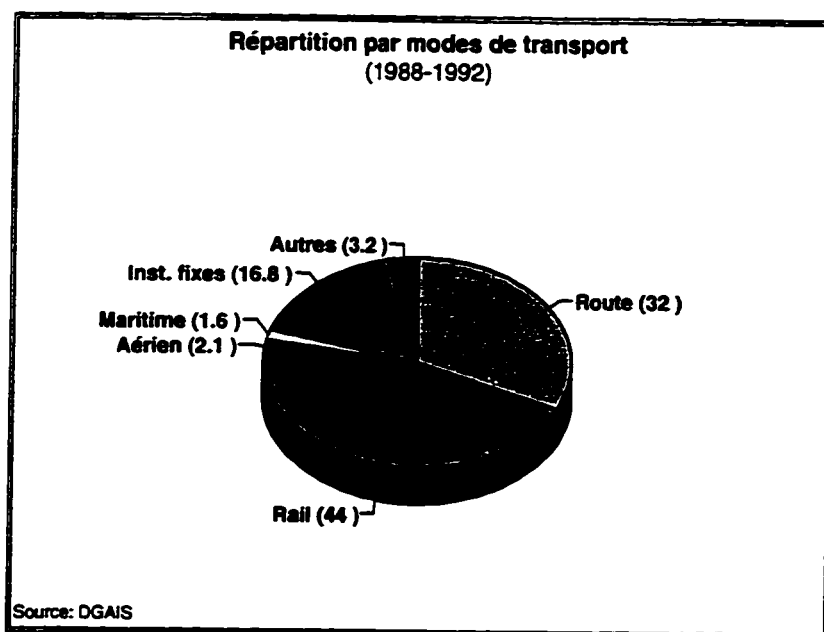


Figure 6.4: Répartition par modes de transport

Il faut toutefois préciser que cette répartition est basée sur le nombre d'accidents, sans tenir compte de la gravité ou du tonnage de matières impliquées. De plus, il est beaucoup plus facile pour les intervenants du transport routier de "passer sous silence" un accident ou un

Transport des matières dangereuses et risques

déversement accidentel alors que le contrôle est beaucoup plus stricte en ce qui concerne le rail ou le transport aérien. En ce sens, les données du DGAIS ne sont peut-être pas parfaitement représentatives de la situation réelle en sous-estimant la part du transport routier. L'activité économique liée à chaque matière dangereuse transportée ainsi que le volume total transporté annuellement ont un impact indirect sur le nombre d'accidents impliquant ces matières. Ainsi, la répartition des accidents par classes de matières dangereuses révèle que ce sont les liquides inflammables (classe 3) qui sont au premier rang avec 34% des accidents suivis des gaz comprimés (classe 2) avec 27% et des matières corrosives (classe 8) avec 19% (figure 6.5). C'est donc dire qu'à elles seules, ces 3 classes de matières comptent pour 80% des accidents.

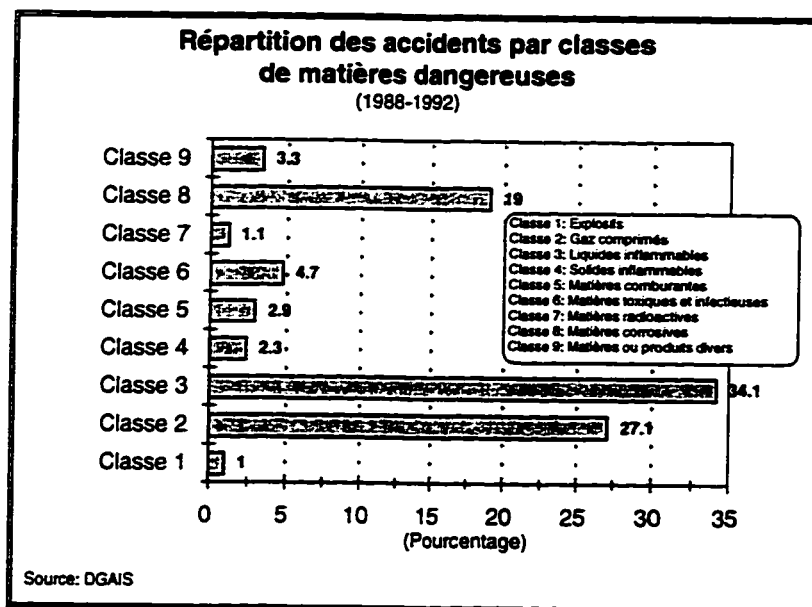


Figure 6.5: Répartition des accidents par classes de matières dangereuses

Comme il a été mentionné précédemment dans ce chapitre, le transport des marchandises dangereuses est fortement lié à l'activité économique des régions. Il n'est donc pas surprenant de constater de fortes variations dans la distribution spatiale des accidents (figure 6.6). On

Transport des matières dangereuses et risques

constate que les plus grandes concentrations d'accidents se retrouvent dans les deux grandes régions urbanisées du pays, c'est-à-dire l'axe Windsor-Montréal-Québec ainsi que Vancouver sur la côte ouest.

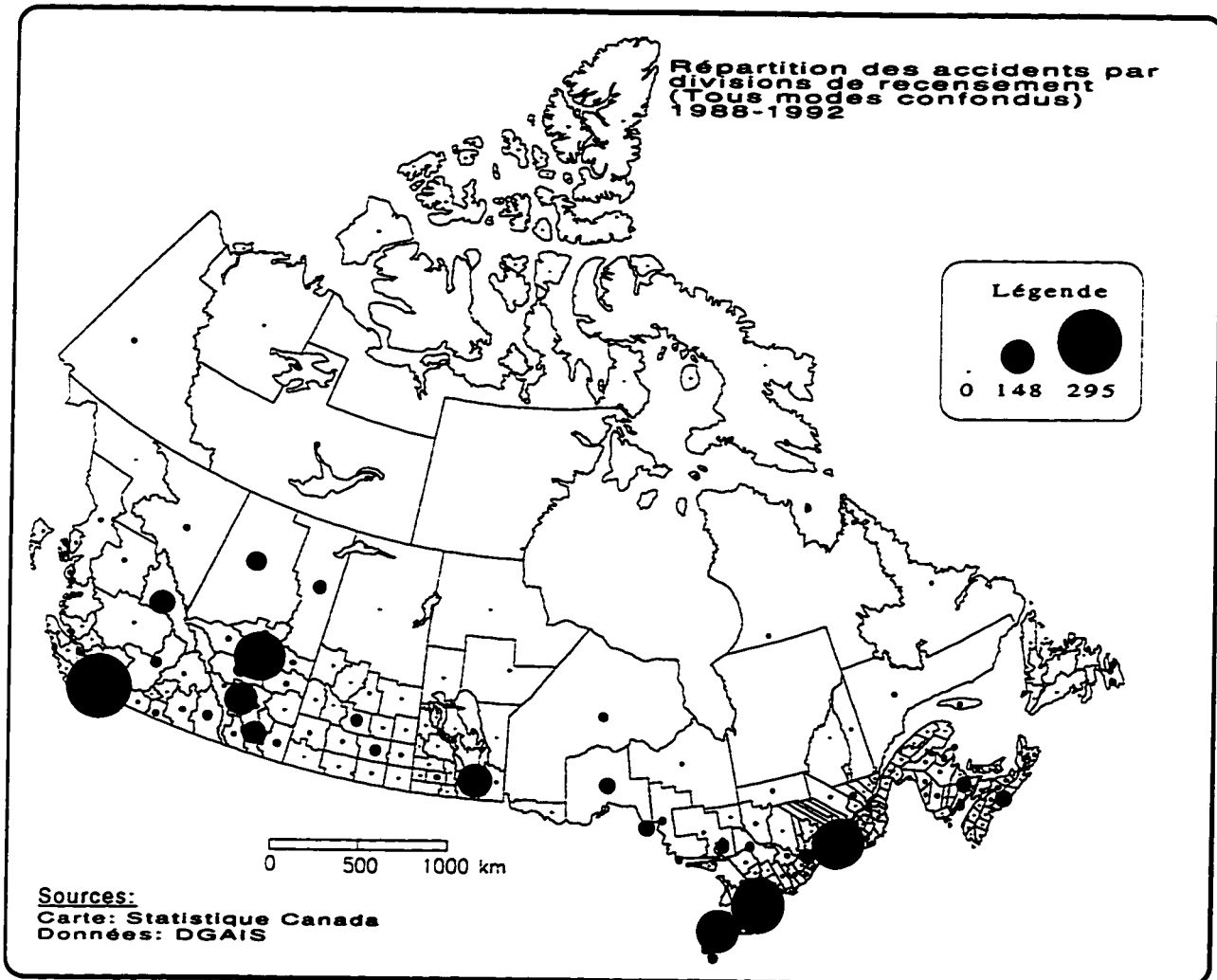


Figure 6.6: Répartition des accidents par division de recensement

Un foyer secondaire, mais de grande importance se situe en Alberta, dans la région d'Edmonton-Calgary-Lethbridge. Enfin, la région de Winnipeg constitue également un foyer secondaire mais de moindre importance. Diverses raisons permettent d'expliquer cette

Transport des matières dangereuses et risques

répartition inégale. Tout d'abord, les grandes régions urbaines de Vancouver et du corridor Windsor-Québec sont de grandes productrices et de grandes consommatrices de matières dangereuses en raison de leur forte activité économique, de leurs infrastructures portuaires et ferroviaires. D'autre part, l'activité de transport étant à son maximum dans ces régions, les risques d'accidents y sont logiquement plus élevés. Quant au foyer secondaire situé en Alberta, il est essentiellement dû à l'exploitation des ressources pétrolières de cette région. De cette simple observation, on peut en déduire que les mesures qui seront mises de l'avant pour réduire les risques devront s'adresser prioritairement aux problèmes rencontrés dans ces régions.

6.3.2 Le transport routier

De façon à mieux saisir la situation du transport routier, nous avons sélectionné parmi les 4783 accidents du DGAIS tous les accidents ayant eu lieu lors de la phase de transport des marchandises et impliquant le mode de transport routier. Pour la période 1988-1992, cela représente 825 accidents.

La répartition des accidents par classe de matières dangereuses (figure 6.7) démontre que ce sont les liquides inflammables (classe 3) qui occupent le premier rang en terme d'accidents, tout comme pour la répartition pour l'ensemble des modes de transport. Dans le cas du transport routier, la part des liquides inflammables atteint 44% contre 34% pour l'ensemble des modes. Cet écart s'explique en partie par la grande quantité de livraisons de pétrole qui se font quotidiennement au pays et surtout par l'activité pétrolière intense dans l'Ouest canadien. Au deuxième et troisième rang des accidents routiers on retrouve les matières corrosives (classe

Transport des matières dangereuses et risques

8) et les gaz comprimés (classe 2), ce qui correspond globalement à la distribution pour l'ensemble des modes de transport confondus.

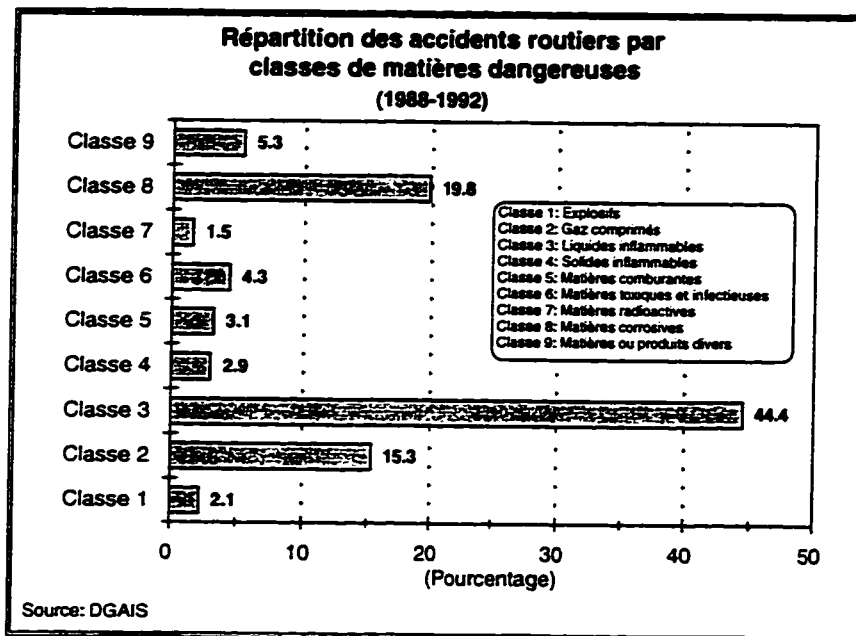


Figure 6.7: Répartition des accidents par classe de matières dangereuses

Dans 56% des cas, les accidents ont donné lieu à des déversements de matières, dans 18% des cas à des fuites alors que dans 15% des accidents, aucun effet suite à l'événement n'a été constaté (figure 6.8). Fait à noter, les explosions et les incendies ou la combinaison de l'ensemble des effets suite à l'accident, ne comptent que pour 10% des cas d'accidents.

La très grande majorité des accidents, soit 70%, ont eu lieu dans une zone rurale ou agricole alors que 20% se sont produits dans un secteur commercial ou industriel. Les accidents en zone urbaine résidentielle ne représentent que 5% du total (figure 6.9). Le pourcentage élevé d'accidents en zone rurale s'explique par le nombre élevé d'accidents liés à l'activité pétrolière albertaine, activité qui s'exerce principalement en milieu rural. A l'extérieur de cette province, on

peut supposer que bon nombre d'accidents se produisent dans la zone de contact entre le milieu urbain et le milieu rural.

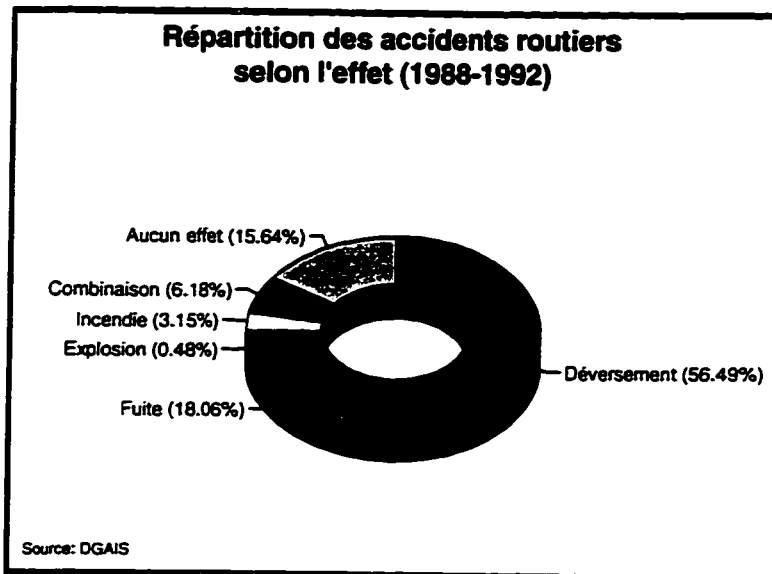


Figure 6.8: Répartition des accidents routiers selon l'effet.

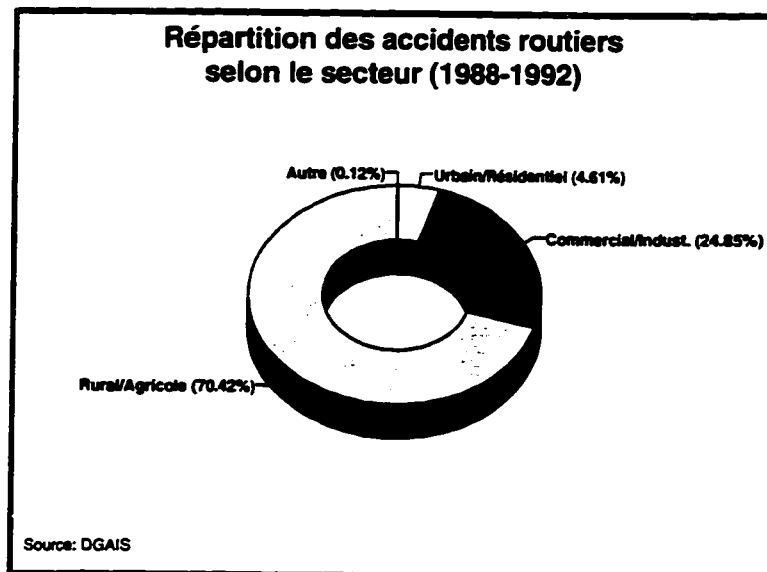


Figure 6.9: Répartition des accidents routiers selon le secteur

Quant à l'impact de la contamination, la base de données du DGAIS nous révèle que 40% des accidents ont résulté en une contamination de l'environnement; 14% en une contamination de la propriété et 13% une contamination humain/propriété/environnement (figure 6.10). La contamination sur les humains ne représente que 2% des cas. Il est à noter toutefois que 31% des observations sont des cas de "donnée non disponible" sur le type de contamination, rendant ainsi incertaine l'interprétation de cette variable.

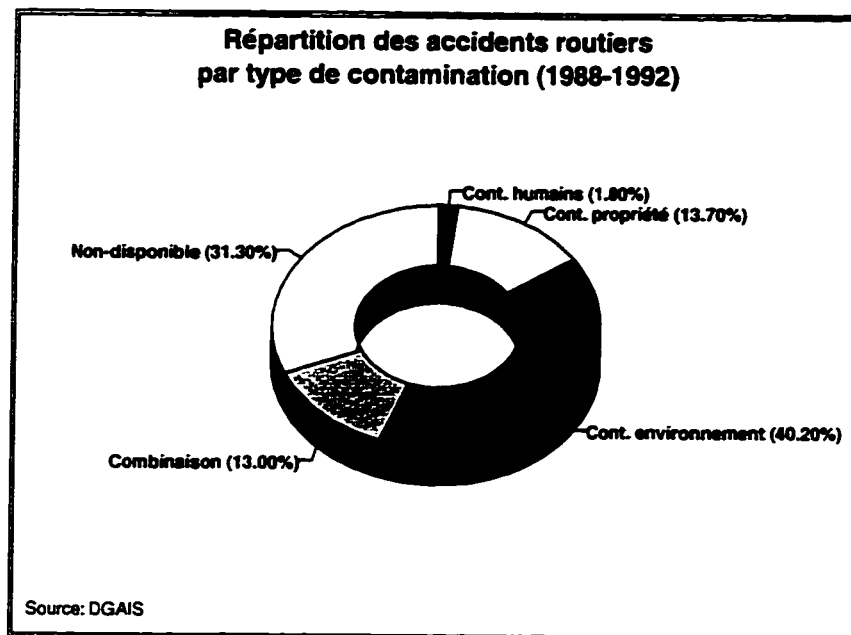


Figure 6.10: Répartition des accidents routiers selon la contamination

La répartition temporelle des accidents laisse également voir certaines particularités (figures 6.11 et 6.12). Sur une base quotidienne, on constate que 40% des accidents ont lieu le mercredi ou le jeudi, avec une diminution importante pendant la fin de semaine. En réalité, cette distribution ne fait que refléter le cycle hebdomadaire des activités de production. Cependant, la répartition horaire démontre plus de variations. Tout d'abord, on peut observer

que le profil horaire des accidents est sensiblement différent selon que l'on considère l'ensemble des modes de transport ou le mode routier seulement (figure 6.12a et 6.12b).

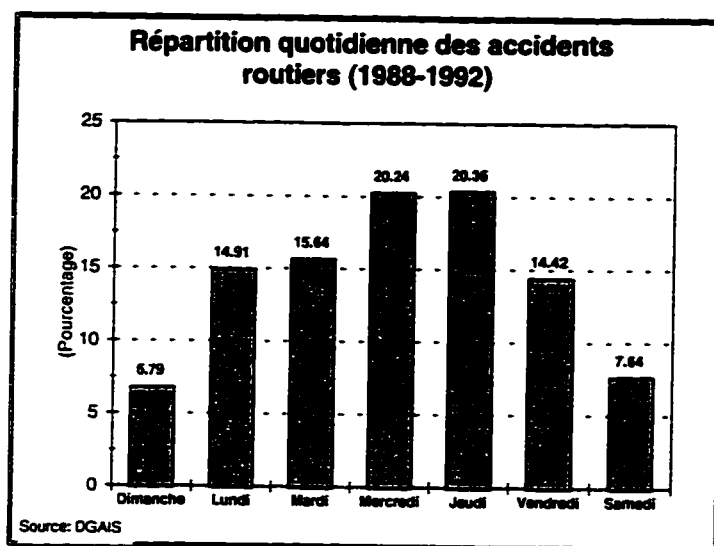


Figure 6.11: Répartition quotidienne des accidents routiers

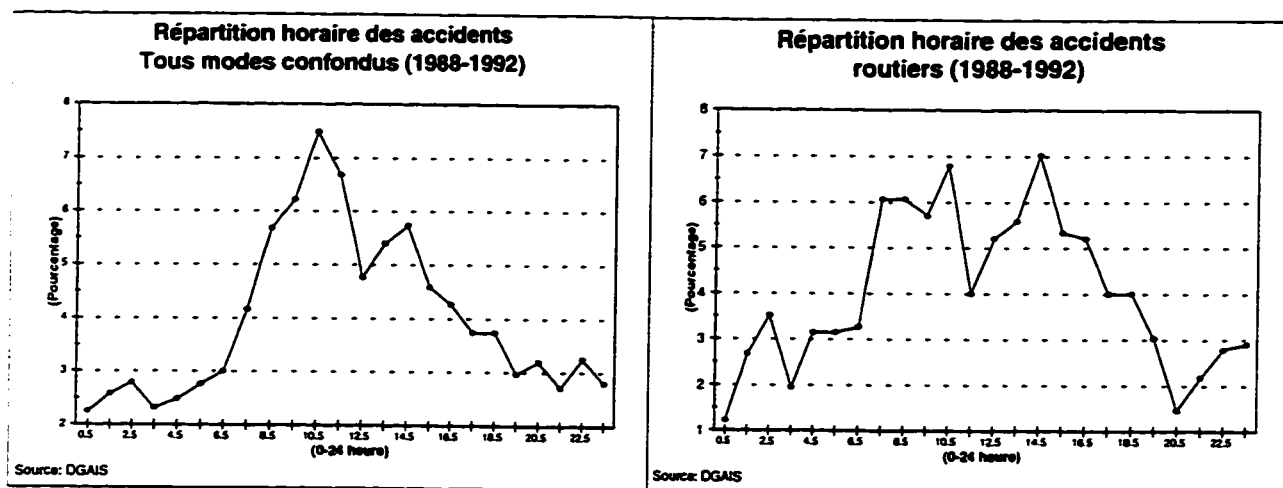


Figure 6.12a: Répartition horaire des accidents, tous modes confondus

Figure 6.12b: Répartition horaire des accidents routiers

Alors que pour tous les modes de transport on retrouve une concentration maximum des accidents entre 10 et 11 heure et des fréquences plus faibles pour le reste de la journée, dans

Transport des matières dangereuses et risques

le cas du transport routier, les fréquences maximum correspondent au cycle de livraison AM et PM avec deux sommets entre 10-11 heure et 14-15 heure. Deux périodes apparaissent donc plus propices aux accidents routiers, soit le milieu de l'avant-midi et le milieu de l'après-midi.

Les accidents routiers affichent aussi une distribution spatiale fort différente de l'ensemble des modes de transport. L'Alberta, à elle seule, s'accapare de 32% de tous les accidents routiers, suivie de l'Ontario (22%) et du Québec (13%) (figure 6.13).

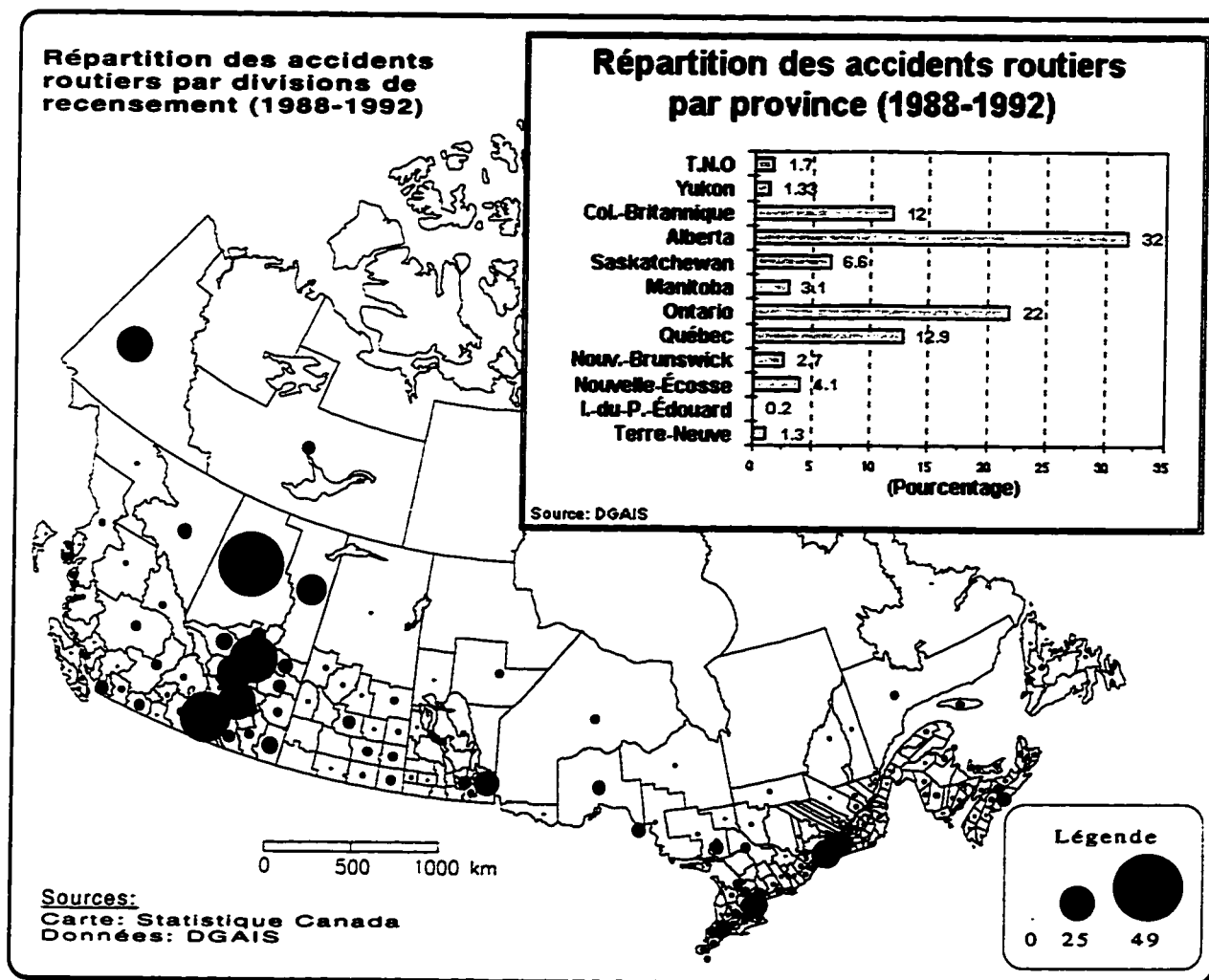


Figure 6.13: Répartition des accidents routiers par divisions de recensement

Transport des matières dangereuses et risques

En Colombie-Britannique, l'écart considérable entre le transport routier et l'ensemble des modes s'explique pour une large part par l'importance de l'activité portuaire de Vancouver et dans une moindre mesure de celle du rail, ce qui laisse peu de place au camionnage.

Quant au corridor Windsor-Québec, les fréquences relativement faibles d'accidents routiers confirment que le rail occupait encore en 1988-1992 une part importante du transport de marchandises. La popularité sans cesse croissante de l'industrie du camionnage nous porte cependant à croire qu'il y aura une augmentation du nombre d'accidents dans cette région au cours des prochaines années.

Pour confirmer l'importance de l'industrie pétrolière dans le lourd bilan d'accidents de l'Alberta, nous avons produit un tableau de contingence entre les variables "province" et "classe de matières dangereuses" (tableau 6.2). La valeur "*p-value*" du khi-carré confirme qu'il y a dépendance entre la classe de matières impliquées dans l'accident et la province où a eu lieu l'accident. Il faut toutefois interpréter ces résultats avec beaucoup de circonspection puisque plus de 20% des cellules du tableau ont une fréquence attendue inférieure à 5 et que le coefficient de contingence est plutôt faible avec 0.53. Ce croisement nous indique que 53.2% des accidents ayant eu lieu en Alberta (48) impliquaient des matières inflammables (classe 3). Ce pourcentage est aussi fort élevé pour la Saskatchewan (47) et le Manitoba (46) avec respectivement 60% et 57.7% des accidents. On constate par contre que les bilans ontariens et québécois sont plus diversifiés puisque les matières inflammables ne comptent que pour 30% et 35.5% des accidents et que les matières corrosives (classe 8) occupent une place aussi importante avec 31.7% et 29%.

Transport des matières dangereuses et risques

2-Way Contingency Table

CLASSE	PROVINCE													TOTAL
FREQUENCY	10	11	12	13	24	35	46	47	48	59	60	62	99	
TOT. PCT.														
ROW PCT.														
COL. PCT.														
1	0	0	0	0	3	3	1	0	8	3	0	0	0	18
	0.0	0.0	0.0	0.0	0.6	0.6	0.1	0.0	1.0	0.4	0.0	0.0	0.0	2.2
	0.0	0.0	0.0	0.0	16.7	16.7	5.6	0.0	44.4	16.7	0.0	0.0	0.0	
	0.0	0.0	0.0	0.0	2.8	1.7	3.8	0.0	3.0	3.1	0.0	0.0	0.0	
2	2	0	2	1	16	27	4	6	49	20	0	0	0	127
	0.2	0.0	0.2	0.1	1.9	3.1	0.5	0.7	5.9	2.4	0.0	0.0	0.0	15.4
	1.6	0.0	1.6	0.8	12.6	21.3	3.1	4.7	38.6	15.7	0.0	0.0	0.0	
	18.2	0.0	5.9	4.3	15.0	15.0	15.4	10.9	28.6	20.4	0.0	0.0	0.0	
3	5	0	15	9	38	54	15	33	140	43	4	11	0	167
	0.6	0.0	1.8	1.1	4.6	6.5	1.8	4.0	27.0	5.2	0.5	1.3	0.0	44.5
	1.4	0.0	4.1	2.5	10.4	14.7	4.1	9.0	38.1	11.7	1.1	3.0	0.0	
	45.5	0.0	44.1	19.1	35.5	30.0	57.7	60.0	53.2	43.9	16.4	78.6	0.0	
4	0	0	0	3	3	2	0	2	12	1	0	0	1	24
	0.0	0.0	0.0	0.4	0.4	0.2	0.0	0.2	1.5	0.2	0.0	0.0	0.1	2.9
	0.0	0.0	0.0	12.5	12.5	8.3	0.0	8.3	50.0	4.2	0.0	0.0	4.2	
	0.0	0.0	0.0	13.0	2.8	1.1	0.0	3.6	4.6	1.0	0.0	0.0	100.0	
5	0	0	2	1	3	5	0	1	6	8	0	0	0	26
	0.0	0.0	0.2	0.1	0.4	0.6	0.0	0.1	0.7	1.0	0.0	0.0	0.0	3.2
	0.0	0.0	7.7	3.8	11.5	19.2	0.0	3.8	23.1	30.8	0.0	0.0	0.0	
	0.0	0.0	5.9	4.3	2.8	2.8	0.0	1.8	2.3	8.2	0.0	0.0	0.0	
6	0	0	4	1	3	16	0	3	5	1	2	1	0	36
	0.0	0.0	0.5	0.1	0.4	1.9	0.0	0.4	0.6	0.1	0.2	0.2	0.0	4.4
	0.0	0.0	11.1	2.8	8.3	44.4	0.0	8.3	11.9	2.8	5.6	2.8	0.0	
	0.0	0.0	11.8	4.3	2.8	8.9	0.0	5.5	1.9	1.0	18.2	7.1	0.0	
7	0	2	0	0	2	3	0	2	4	0	0	0	0	13
	0.0	0.2	0.0	0.0	0.2	0.4	0.0	0.2	0.5	0.0	0.0	0.0	0.0	1.6
	0.0	15.4	0.0	0.0	15.4	23.2	0.0	15.4	30.8	0.0	0.0	0.0	0.0	
	0.0	100.0	0.0	0.0	1.9	1.7	0.0	3.6	1.5	0.0	0.0	0.0	0.0	
8	2	0	7	4	31	57	4	4	33	19	2	1	0	164
	0.2	0.0	0.8	0.5	3.8	6.9	0.5	0.5	4.0	2.3	0.2	0.2	0.0	19.9
	1.2	0.0	4.3	2.4	18.9	34.8	2.4	2.4	20.1	11.6	1.2	0.6	0.0	
	18.2	0.0	20.6	17.4	29.0	31.7	15.4	7.3	22.5	19.4	18.2	7.1	0.0	
9	2	0	3	4	7	11	2	4	4	3	3	1	0	44
	0.2	0.0	0.4	0.5	0.8	1.3	0.2	0.5	0.5	0.4	0.4	0.2	0.0	5.3
	4.5	0.0	6.8	9.1	15.9	25.0	4.5	9.1	9.1	6.8	6.8	2.3	0.0	
	18.2	0.0	8.8	17.4	6.5	6.1	7.7	7.3	1.5	3.2	27.3	7.1	0.0	
TOTAL	11	2	34	23	107	180	26	55	263	98	11	14	1	825
	1.3	0.2	4.1	2.8	13.0	22.0	3.2	6.7	31.9	11.9	1.3	1.7	0.1	100.0

WARNING - Some Expected values less than 5. Chi-Square may not be valid.

Statistic	DF	Value	p-value
Chi-Square	108	322.262	0.000
Phi Coefficient		0.625	
Cramer's V		0.208	
Contingency Coefficient		0.530	

Tableau 6.2: Tableau de contingence, province/classe de matières

La base de données du DGAIS renferme également des indications sur les causes de l'accident ainsi que sur le degré potentiel et actuel de sévérité des accidents. Il appert que plus de 68% des accidents sont causés par une erreur humaine, soit une erreur du conducteur ou d'un tiers (figure 6.14). Les mauvaises conditions routières ou la présence d'obstacles contribuent pour 19% des accidents alors que les bris mécaniques ne sont la cause que de 5% des accidents.

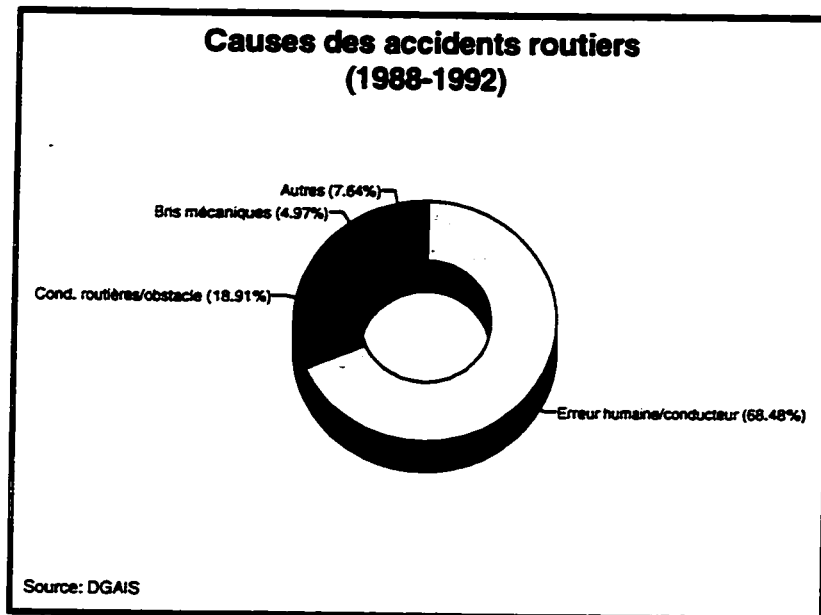


Figure 6.14: Facteur ayant provoqué l'accident routier

Dans la très grande majorité des cas (70%), le degré de sévérité des accidents a été jugé mineur par CANUTEC, c'est-à-dire ayant peu ou pas de conséquences, peu de danger de décès ou d'impact environnemental (figure 6.15).

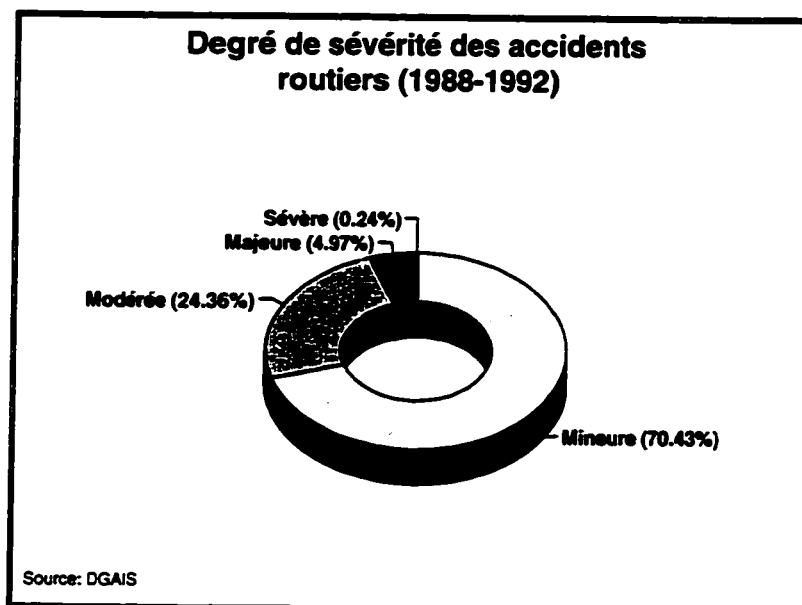


Figure 6.15: Degré de sévérité des accidents routiers

Transport des matières dangereuses et risques

25% des accidents se sont avérés d'une sévérité modérée, soit avec blessés nécessitant une hospitalisation et des dommages variant entre 10 000 et 2 millions de dollars. Enfin, 5% des accidents ont reçu la cote "majeure" ou "sévère", c'est-à-dire des accidents avec blessés et décès et des coûts variant entre 2 et 20 millions de dollars.

Le degré potentiel de sévérité, c'est-à-dire les conséquences potentielles d'un accident incontrôlé, est plus partagé: 43% des accidents ont été classés à potentiel mineur, 43% à potentiel modéré, 13% à potentiel majeur et 1% à potentiel majeur ou sévère (figure 6.16). Avec un taux de 14% d'accidents à potentiel majeur ou sévère, cela représente sur 5 ans environ 115 accidents dont les conséquences auraient pu être le décès d'une ou plusieurs personnes et des dommages variant entre 2 et 20 millions de dollars.

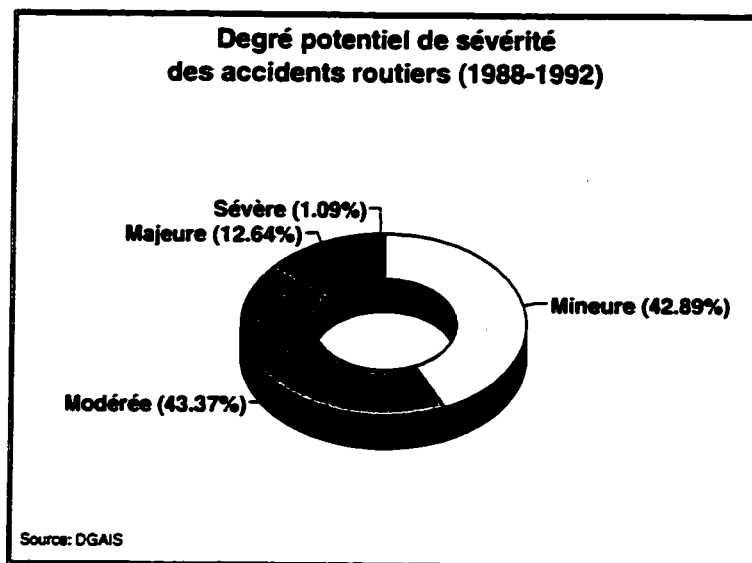


Figure 6.16: Degré potentiel de sévérité des accidents routiers

Pour vérifier l'existence d'un lien entre le degré de sévérité actuel de l'accident et le secteur où il s'est produit, nous avons eu recours à un tableau de contingence (tableau 6.3). La valeur du

Transport des matières dangereuses et risques

khi-carré confirme qu'il y a un lien entre la sévérité de l'accident et le secteur où il a eu lieu. Cependant, plus de 20% des cellules ont des fréquences attendues en deçà de 5, ce qui nous empêche de conclure de façon catégorique, d'autant plus que le coefficient de contingence n'est que de 0.194.

2-Way Contingency Table					
AREA	ACTUAL				
FREQUENCY	1	2	3	4	TOTAL
TOT. PCT					
ROW PCT					
COL. PCT					
1	26	22	0	0	38
	3.2	1.5	0.0	0.0	4.6
	68.4	31.6	0.0	0.0	
	4.5	6.0	0.0	0.0	
2	172	29	3	1	205
	20.8	3.5	0.4	0.1	24.8
	83.9	14.1	1.5	0.5	
	29.6	14.4	7.3	50.0	
5	383	159	38	1	581
	46.4	19.3	4.6	0.1	70.4
	65.9	27.4	6.5	0.2	
	65.9	79.1	92.7	50.0	
6	0	1	0	0	1
	0.0	0.1	0.0	0.0	0.2
	0.0	100.0	0.0	0.0	
	0.0	0.5	0.0	0.0	
TOTAL	581	201	41	2	825
	70.4	24.4	5.0	0.2	100.0

WARNING - Some Expected values less than 5. Chi-Square may not be valid.

Statistic	DF	Value	p-value
Chi-Square	9	32.398	0.000
Phi Coefficient		0.198	
Cramer's V		0.114	
Contingency Coefficient		0.194	

Tableau 6.3: Tableau de contingence, sévérité/secteur

Bien que 68.4% des accidents à s'être produit en zone urbaine ou résidentielle (1) aient été classifiés de mineurs, 31.6% d'entre eux ont reçu la côte modérée, c'est-à-dire un accident avec blessés et dommages de 10000 à 2 millions de dollars. De plus, ce même tableau indique que 92.7% des accidents jugés majeurs (décès, blessés, 2-10 millions en dommages) se sont produits en zone rurale ou agricole et que ce type d'accident est totalement absent des zones urbaines ou résidentielles pour la période 1988-1992. Un tableau de contingence similaire a été produit avec le degré potentiel de sévérité des accidents (tableau 6.4). Les mêmes contraintes qu'au tableau précédent, quant à la validité des résultats, s'appliquent à ce tableau. Ce dernier laisse voir que 77.8% des accidents qualifiés de potentiellement sévères (décès multiples,

Transport des matières dangereuses et risques

blessés, dommages 10-20 millions) étaient localisés dans des secteurs ruraux ou agricoles et que le reste des accidents potentiellement sévères (22.2%) se sont produits en zone commerciale ou industrielle.

2-Way Contingency Table					
AREA	POTENTIAL				
FREQUENCY	1	2	3	4	TOTAL
TOT PCT					
ROW PCT					
COL PCT					
1	14	23	1	0	38
	1.7	2.8	0.1	0.0	4.6
	36.8	60.5	2.6	0.0	
	4.0	6.4	1.0	0.0	
2	107	84	12	2	205
	13.0	10.2	1.5	0.2	24.8
	52.2	41.0	5.9	1.0	
	30.3	23.5	11.5	22.2	
5	232	249	91	7	581
	28.1	30.2	11.0	0.8	70.4
	39.9	42.9	15.7	1.2	
	65.7	69.7	27.5	77.8	
8	0	1	0	0	1
	0.0	0.1	0.0	0.0	0.1
	0.0	100.0	0.0	0.0	
	0.0	0.3	0.0	0.0	
TOTAL	353	357	104	9	825
	42.8	43.3	12.6	1.1	100.0

WARNING - Some Expected values less than 5. Chi-Square may not be valid.

Statistic	DF	Value	p-value
Chi-Square	12	25.926	0.013
Phi Coefficient		0.177	
Cramer's V		0.102	
Contingency Coefficient		0.175	

Tableau 6.4: Tableau de contingence, sévérité potentielle/secteur

Il faut donc en conclure que les zones urbaines ou résidentielles ont été épargnées pendant cette période, du moins en ce qui concerne les accidents pouvant entraîner de graves conséquences. La notion de risque déborde toutefois du cadre du simple bilan accidentel. Plusieurs accidents majeurs impliquant des matières dangereuses se sont produits dans le passé (Webb, Saskatchewan, 1980, 22 morts / Hinton, Alberta, 1986, 40 morts, etc.). Les risques potentiels demeurent malgré le caractère sporadique de ces événements. C'est à ce niveau que la notion de risque prend le pas sur la notion de probabilité issue de l'inférence statistique, d'autant plus que la sensibilité du public face aux risques s'est considérablement accrue depuis quelques décennies.

6.4 La notion de risque

Bien que l'approche probabiliste de l'évaluation des risques ait été abordée à la section 2.3.3 du chapitre 2, il nous apparaît important d'y consacrer ici quelques pages en regard de la manière dont la société évalue, gère et tolère les risques inhérents à toute activité humaine. En fait, le danger a toujours fait parti de l'environnement humain. Si, au tout début, l'humain était principalement confronté aux risques naturels, les premiers risques technologiques sont probablement apparus avec la découverte du feu et n'ont cessé de croître depuis.

Avant d'aller plus loin, il importe de distinguer entre deux concepts fortement dépendants mais représentant une réalité différente: le *risque* qui fait l'objet de la présente section et l'*incertitude* dont il a été question au chapitre précédent. "Jouer à la roulette russe est un risque si l'on sait que le pistolet est chargé d'une balle dans l'une des six chambres du barillet. Il y a alors une chance sur six d'être tué à chaque coup. Cependant, si l'on joue à la roulette russe et que l'on ne sait pas si le pistolet est chargé ou non, nous place en situation d'incertitude. Plus l'incertitude est présente dans une situation de prise de décision, plus elle aura pour effet de limiter l'application des modèles qui assument une connaissance complète de la situation.

Cet exemple correspond bien au contexte du transport des matières dangereuses en raison de l'incertitude liée au manque de données relatives à chaque tronçon routier. Cette conjoncture n'a toutefois pas empêché l'apparition de la discipline de l'évaluation des risques, discipline qui s'est attardée entre autres à définir les différents types de risques.

6.4.1 Les types de risques

L'identification du risque constitue la reconnaissance de l'existence d'un danger présentant des caractéristiques définissables. Le risque réel peut se définir comme la probabilité qu'un événement ou une situation se produise et crée un préjudice au milieu ou aux organismes. La probabilité d'occurrence pourra, à des degrés divers, être entachée d'incertitude alors que les préjudices pourront être des effets adverses, des dommages, ou des pertes humaines ou monétaires.

$$\text{Risque réel} = [\text{Probabilité d'occurrence}] \times [\text{Sévérité des préjudices}]$$

Au fil des années, les spécialistes se sont aperçus que l'on ne peut définir le risque de façon uniquement objective, sans tenir compte des valeurs et des priorités des sociétés. Il y a souvent différence entre l'évaluation scientifique des risques et la façon dont le public perçoit, accepte ou rejette ces mêmes risques. A ce titre il est indéniable, quoi qu'en dise la logique, qu'un accident impliquant des matières dangereuses attire plus l'attention du public que la multitude d'accidents banals qui font annuellement beaucoup plus de victimes¹ La magnitude de l'accident et les dangers potentiels jouent un rôle important dans la perception du risque par les populations. En fait, outre le risque réel, il faut distinguer plusieurs autres types de risques qui font tous plus ou moins référence à un "construit psychologique et social" (LEVESQUE, 1995, p. 19).

En premier lieu, l'idée que se forment les gens du risque dépend de son caractère volontaire ou involontaire. Un risque volontaire correspond à un risque qu'un individu ou un groupe

¹Par exemple, en France, les accidents de la route ont fait plus de victimes que les 2 grandes guerres mondiales réunies (RUBISE et GAUTIER, 1995, p. 118)

d'individus choisissent librement d'assumer sans qu'il ne soit requis à la conduite des activités quotidiennes (ROBERGE, 1984, p. 14). Escalader une montagne relève du risque volontaire. Généralement, la société est moins sensible à ce genre de risques puisqu'il repose sur le libre choix des individus. Par contre, le risque involontaire qui intervient dans les activités courantes reçoit plus d'attention de la part du public et des gouvernements car dans ce cas, la victime n'a pas choisi librement le risque, elle le subit.

C'est à ce niveau que le phénomène de perception du risque prend son importance. On parle alors de *risque perçu*. Le risque n'est pas un phénomène objectif perçu de la même manière par toutes les parties intéressées (gouvernements, scientifiques, groupes populaires, etc.). En fonction de son intérêt, chaque individu ou groupe donne une signification différente au risque, signification qui repose sur l'interaction entre trois composantes (WHYTE, 1983, p 121): 1) les caractéristiques de la personne (âge, sexe, éducation, etc.) 2) la situation (contexte social et économique) 3) les caractéristiques du risque ou du danger (figure 6.17).

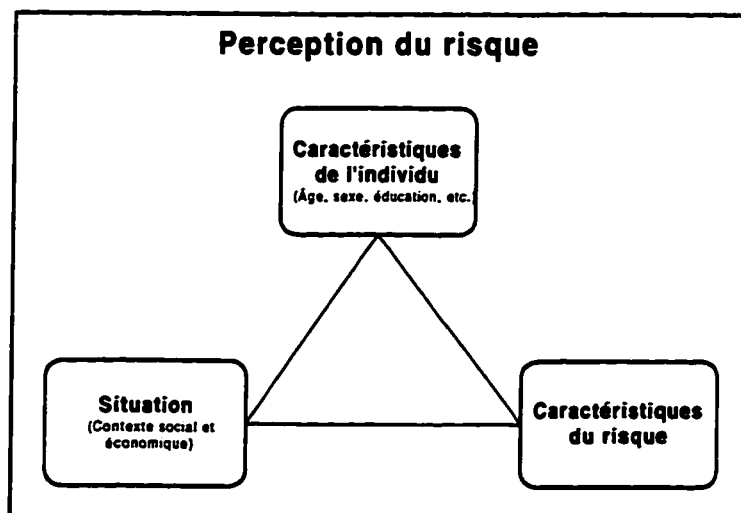


Figure 6.17: Composantes de la perception du risque

Transport des matières dangereuses et risques

D'autres facteurs sont aussi susceptibles d'influencer la perception du risque en le rendant plus intolérable pour la population. On retrouve entre autres (LEVESQUE, 1995, p. 21):

1) Potentiel de catastrophe	7) Attention des médias
2) Incertitude	8) Historique des accidents
3) Effet sur les enfants	9) Équité
4) Effet sur les générations futures	10) Présence de bénéfices
5) Personnalisation des victimes	11) Réversibilité des effets
6) Crédibilité des institutions	12) Degré d'implication personnelle

Certains de ces facteurs peuvent avoir un effet puissant, comme la personnalisation des victimes, l'irréversibilité des effets ou encore l'effet sur les enfants, ces derniers étant considérés comme des victimes innocentes. Cette notion de risque intolérable nous amène à un autre aspect du domaine du risque: le risque acceptable. La définition du niveau de risque acceptable constitue la question fondamentale de tout le processus d'évaluation du risque. Que ce soit par fatalisme, ou en retour d'un bénéfice technologique ou autre, la société accepte volontairement et consciemment le risque comme un prix "raisonnable" à payer pour certains bénéfices. L'usage de l'automobile et les milliers de décès annuels qu'il occasionne en est un exemple parfait. Le jugement sur l'acceptabilité d'un risque ne relève pas uniquement de facteurs quantifiables mais aussi d'une multitude de dimensions et de facteurs qualitatifs.

Espérer régler le problème du risque acceptable en ce qui concerne le transport des matières dangereuses uniquement sur une base quantitative (probabilités et inférence statistique) est illusoire dès l'instant où le public a à intervenir. Ce dernier est souvent moins sensible au côté objectif du quantifiable qu'au caractère équitable du risque. Du point de vue éthique, le risque doit également être équitable quant à sa répartition sur la population. De façon typique, ceux

qui profitent des bénéfices d'une technologie ne sont pas nécessairement ceux qui supportent les risques de cette même technologie.

Plusieurs types d'iniquités peuvent se produire (KASPER et KASPER, 1983, p. 138): l'iniquité entre le public et les travailleurs (ceux dont le travail implique des risques importants); l'iniquité entre les générations qui consiste à léguer les problèmes aux générations futures; l'iniquité entre groupes sociaux et enfin, l'iniquité du point de vue géographique qui occupe une place particulière en transport des matières dangereuses. Cette dernière forme d'iniquité est à mettre en relation avec le phénomène du "pas dans ma cours" qui illustre de façon évidente que ce sont les groupes et les communautés, via le processus politique qui, en dernière instance, se prononcent sur l'acceptation ou le rejet d'un risque. Désormais, l'acceptabilité du risque, sans exclure l'apport des spécialistes, relève également d'une acceptabilité sociale et l'on ne peut omettre cet aspect dans le processus d'évaluation et de mitigation des risques.

6.4.2 Évaluation et mitigation des risques

Dès lors que l'on tient compte de l'acceptabilité sociale dans le processus d'évaluation des risques, comment peut-on en arriver à réduire ces risques en alliant à la fois des critères quantitatifs et qualitatifs? Tout d'abord, il faut admettre que l'on ne peut faire disparaître complètement le risque lié au transport des matières dangereuses, on ne peut que le réduire. De plus, cette réduction des risques s'accompagne nécessairement de coûts (détours des camions, itinéraires allongés, contrôle policier, etc.), coûts qui croissent en fonction du degré de sécurité que l'on désire atteindre et des solutions retenues (figure 6.18).

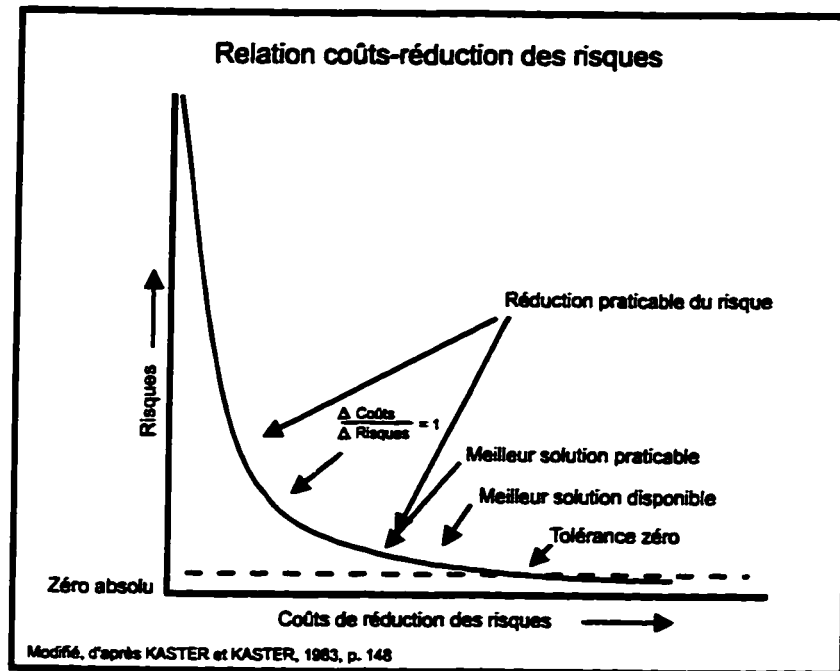


Figure 6.18: Coûts de la réduction des risques

Comme le démontre la figure 6.18, il n'y a aucun doute que la réduction absolue des risques liés au transport des matières dangereuses dépasse la capacité de payer des différents intervenants. En fait, l'étendue de réduction praticable du risque exclut d'emblée la meilleure solution disponible (solution parfaite) et le seuil de tolérance zéro. Ces coûts externes à l'activité de transport qui viennent en contradiction avec le système économique doivent être évalués sérieusement avant de prendre une décision qui pourrait pénaliser indûment les transporteurs. Cette évaluation relève d'un compromis entre le système de valeurs (choix de la société); le système économique (coûts et impacts économiques) et le système physique (le niveau de risque réel) (figure 6.19). Les gouvernements et les spécialistes de l'évaluation du risque jouent un rôle central dans ce mécanisme puisque l'on doit privilégier un équilibre entre les attentes des différents groupes impliqués. Concrètement, cet équilibre se crée et évolue

dans le temps au gré des fluctuations du pouvoir de chacun des groupes. Il ne peut donc y avoir de "recette magique" et c'est par un consensus que les critères d'évaluation sont établis.

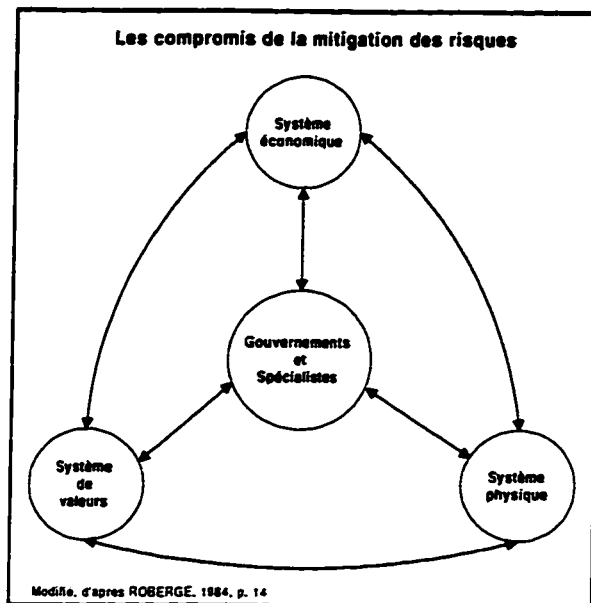


Figure 6.19: Mitigation des risques et compromis

Bien qu'aucune recette magique n'existe, plusieurs auteurs ont présenté des solutions pratiques à ce problème. De façon générale, la plupart des méthodes d'évaluation procèdent par élimination; c'est-à-dire que dans un premier temps, on traite les cas évidents qui font consensus pour l'ensemble des intervenants puis, dans un deuxième temps, on aborde les cas problématiques. On peut synthétiser ces différentes méthodes selon la procédure suivante:

1) En premier lieu, on traite les situations qui font consensus

A) Situations dont les conséquences potentielles d'une erreur seraient particulièrement sévères. Exemple: Accident et déversement de pétrole près des prises d'eau potable.

B) Situations dont les conséquences sont irréversibles. Exemple: Accident avec incendie ou explosion risquant de tuer plusieurs passants (Parcs, cours d'écoles, etc.)

Transport des matières dangereuses et risques

C) Situations à faible probabilité mais à graves conséquences.
Exemple: Accident impliquant des matières fortement radioactives en milieu urbain.

D) Situations qui comportent une large part d'inconnu. Exemple: Les secteurs où l'achalandage de population varie fortement selon les périodes de la journée.

2) On élimine tout risque qui ne rapporte aucun bénéfice ou qui peut être facilement évité. Exemple: Pour deux tronçons de même longueur, choisir celui qui possède la plus faible population ou la population la moins vulnérable.

3) On élimine tout risque important (1/10 000 ou plus annuellement) qui ne rapporte pas de bénéfice majeur. Exemple: Éliminer les tronçons pour lesquels on note un passé lourd en terme d'accidents (points noirs).

4) On ignore, à ce stade, tous les risques mineurs (1/1 000 000 ou moins) qui ne tombent pas dans la catégorie 2.

5) On étudie de manière active les risques qui ne correspondent à aucune des catégories précédentes en évaluant le pour et le contre d'agir ou non en regard de ce risque. A ce stade, le critère d'évaluation consiste à savoir à quels sacrifices la société est prête à consentir pour éviter les conséquences particulières d'un risque.

6) Au besoin, revenir à l'étape 4 et traiter les risques mineurs.

Cette méthode d'élimination des risques basée sur un raisonnement logique et des critères à la fois quantitatifs et qualitatifs peut très bien s'intégrer à un système expert. Dès que les différents groupes se sont entendus sur les risques tolérables et sur ceux devant être éliminés, il n'y a plus qu'à traduire sous forme de règles l'ensemble des critères en y associant, au besoin, un facteur de certitude. C'est donc à partir de ce consensus qu'il devient possible de "nourrir" le système expert et, grâce à son module d'explication des résultats, ce même système sera en mesure de justifier son choix d'itinéraire en fonction des critères établis par consensus.

Transport des matières dangereuses et risques

Évidemment, la détermination d'itinéraires sécuritaires n'est pas le seul moyen de réduire les risques liés au transport des matières dangereuses. La fabrication de conteneurs plus résistants, de réservoirs renforcés à double parois, de camions plus sécuritaires, une meilleure formation du personnel, une législation et un contrôle policier plus efficace sont autant de moyens de réduire les risques de ce type de transport. Toutefois, dans un contexte spatial, l'élaboration d'un modèle de calcul d'itinéraire intégrant les différents points abordés jusqu'à maintenant dans cette thèse, nous apparaît comme une solution potentiellement intéressante. C'est à l'élaboration de ce modèle qu'est consacré le prochain chapitre.

Chapitre VII

Formalisation du modèle de recherche d'itinéraire

“L'évolution ne tire pas ses nouveautés du néant. Elle travaille sur ce qui existe déjà, soit qu'elle transforme un système ancien pour lui donner une fonction nouvelle, soit qu'elle combine plusieurs systèmes pour en échafauder un autre plus complexe.”

- Claude Paquette, *L'effet caméléon*

Les chapitres précédents ont permis de brosser un tableau du contexte de la recherche d'itinéraires sécuritaires pour le transport des matières dangereuses. Nous en sommes maintenant à l'étape de proposer un nouveau modèle de recherche d'itinéraire qui soit en mesure de répondre aux exigences propres au domaine. Mais, tout d'abord, il nous apparaît important de faire un retour sur les principaux éléments des chapitres antérieurs de façon à mieux cerner les exigences auxquelles devra répondre ce modèle. Par la suite, le reste du chapitre sera consacré à la formalisation de ce modèle, sa présentation sous forme de diagramme de flux ainsi qu'à son intégration à un système d'information géographique (SIG) pour en faire un système de support à la décision à référence spatiale (SSDRS).

7.1 Propriétés du modèle de recherche d'itinéraire

Les chapitres 2 et 3 ont démontré que l'approche conventionnelle de recherche d'itinéraires, qu'elle soit à objectif unique ou à objectifs multiples, repose essentiellement sur la recherche d'un itinéraire minimisant la valeur cumulative d'impédance ou de distance. En transport des matières dangereuses, une des méthodes consiste à associer la valeur d'impédance de chaque tronçon à la probabilité d'occurrence d'un accident pour ce tronçon. Toutefois, la rareté des accidents, leur caractère imprévisible, les données manquantes et les très faibles probabilités associées à chaque tronçon limitent grandement cette méthode. De plus, en pratique, cette méthode exclue tous les aspects non quantitatifs du problème.

L'autre approche courante consiste à réduire l'ensemble des caractéristiques de chaque tronçon à une valeur unique d'impédance analogue à la notion de poids. Toutefois, cette variable dérivée ne permet pas toujours de discerner entre 2 ou plusieurs tronçons ayant une impédance identique mais des caractéristiques différentes. De plus, cette approche rend impossible tout raisonnement sur le degré et le type de risque relatif à chaque tronçon puisqu'il n'y a plus accès aux variables originelles lors de la recherche d'itinéraire. L'introduction de données qualitatives pose également problème avec cette méthode.

Les chapitres 4 et 5 ont mis en évidence le fait que c'est via le développement d'un système de support à décision à référence spatiale (SSDRS) que l'on sera en mesure d'apporter une solution plus efficace au problème de la recherche d'itinéraires sécuritaires. En couplant un système d'information géographique, un algorithme de recherche d'itinéraire et un système expert, on sera en mesure d'offrir aux intervenants en transport un outil de support à la décision

Formalisation du modèle de recherche d'itinéraire

beaucoup plus souple. A ce titre, l'apport des SIG se situe principalement au niveau des capacités d'intégration, de gestion et d'analyse des données géographiques (réseaux et attributs), alors que le système expert, grâce à la théorie de la décision, des ensembles flous et de la prise en compte de l'incertitude, peut s'acquitter de la partie "raisonnement" du SSDRS. Ce lien avec un système expert ouvre également la porte au traitement des données quantitatives et qualitatives, contrairement à l'approche conventionnelle par impédance. C'est en se référant à ces grandes lignes directrices que nous avons établi les critères auxquels devrait répondre notre modèle de recherche d'itinéraire. Ce modèle devrait:

Etre capable d'appliquer une démarche déductive (raisonnement) sur l'ensemble des caractéristiques des tronçons.

Contrairement aux modèles conventionnels qui appliquent l'algorithme de recherche d'itinéraire sur la seule valeur d'impédance, le modèle proposé doit accéder à l'ensemble des variables qui caractérisent les tronçons. Ces variables seront utilisées par le système expert pour déterminer quel tronçon retenir dans l'itinéraire.

Etre en mesure de traiter des données quantitatives (probabilités, flux de circulation, etc) et/ou qualitatives (symboliques)

Par exemple si, pour un tronçon donné, on possède des données fiables (probabilité d'accident) en raison d'études particulières ou d'un passé lourd en accidents, le modèle devrait être en mesure d'en tenir compte en tant qu'information significative en plus des autres caractéristiques du tronçon. Ces dernières peuvent être soit quantitatives (dénombrement de population) soit qualitatives (type de route, type de population, présence d'édifices publics, etc.). Si par contre aucune donnée quantitative n'est disponible, on procède avec les caractéristiques du tronçon uniquement.

Rendre possible la prise en compte de plus d'un objectif d'optimisation (équilibre coûts/bénéfices).

Il est très difficile de tracer les limites précises au-delà desquelles les coûts engendrés par un accroissement de sécurité ne peuvent être justifiés. Cependant, un modèle fonctionnel de recherche d'itinéraire doit faire en sorte qu'à chaque fois que 2 ou plusieurs tronçons offrent le même niveau de sécurité, celui résultant en un trajet le plus court (donc le moins dispendieux) soit retenu

Formalisation du modèle de recherche d'itinéraire

S'appuyer sur l'algorithme du plus court chemin.

Le processus de recherche d'itinéraire est nécessairement un processus itératif d'optimisation de parcours (voir chapitre 3). On doit donc se référer à un algorithme issu de la théorie des graphes comme base au modèle.

Tenir compte du type de matière transportée par le véhicule.

Évidemment, chaque type de matière transportée présente des risques particuliers (explosion, incendie, gaz mortel, etc.). C'est d'ailleurs en partie cette caractéristique qui a favorisé le regroupement en classes des matières dangereuses. L'OCDE a produit un tableau résumant les conséquences probables et l'étendue des zones à risques en cas d'échappement ou d'explosion d'une cargaison de matières dangereuses (OCDE, 1988, pp. 52-53). La base de connaissances du système expert devrait tenir compte de ce facteur dans la recherche d'itinéraire.

Etre en mesure d'intégrer l'incertitude liée aux données dans la recherche d'itinéraire.

L'impossibilité d'opérer sur des données totalement complètes et fiables exige que l'on associe au raisonnement menant au choix d'un itinéraire particulier une mesure d'incertitude. Une mesure de certitude doit également accompagner les règles de production du système expert (coefficient de certitude) pour prioriser les connaissances à appliquer au domaine.

Prendre en compte les aspects temporels du transport des matières dangereuses.

Le risque relatif au transport des matières dangereuses est un risque qui fluctue selon les périodes de la journée, principalement en raison des mouvements de population. Un tronçon de route bordé de commerces présente une forte population pendant le jour, alors qu'il sera quasi désert la nuit. De risqué qu'il est en période diurne, ce tronçon devient sécuritaire en période nocturne. Une approche temporelle par plage horaire laisse donc la place à l'établissement de deux catégories d'itinéraires: de jour ou de nuit.

Permettre d'évaluer plusieurs itinéraires prédéterminés.

La méthode classique qui consiste à déterminer un seul itinéraire optimal ne devrait pas être la seule solution possible au problème. Il faut considérer la possibilité que le système expert fasse l'évaluation à posteriori d'un itinéraire déjà établi de façon à conclure sur le degré de risque qu'il représente ainsi que sur les tronçons qui devraient être évités.

Etre integrable à un système d'information géographique (SIG).

La recherche d'itinéraire nécessite la construction d'un réseau topologique (structure arcs-noeuds) ainsi qu'une mise en relation des attributs aux éléments du réseau. La tâche que représente la constitution d'une base de données géographiques (plus de 50% d'un projet) implique l'intégration à un SIG, au niveau même de la structure de données.

Déboucher sur la création d'un système de support à la décision à référence spatiale (SSDRS)

Ultimement, l'outil qui devrait être remis aux preneurs de décisions est un SSDRS capable de leur fournir un certain nombre d'alternatives à un problème donné, alternatives leur permettant de prendre une décision plus éclairée. Un SSDRS joue également le rôle de "boîte à outils" dans la mesure où il offre tous les outils nécessaires, de la saisie des données jusqu'aux recommandations finales en passant par les analyses.

7.2 Formalisation du modèle de recherche d'itinéraire

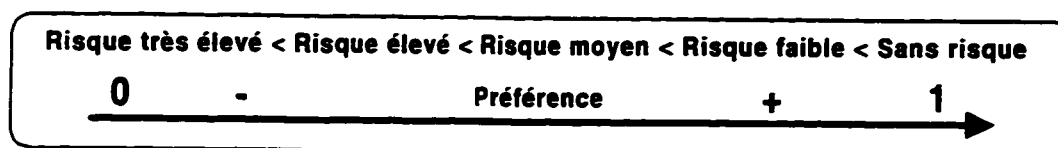
Les critères énumérés dans la section précédente mettent en évidence l'importance du système expert dans le modèle proposé. La réalisation de la plupart des ces exigences passera logiquement par des règles de production qui traduisent sous forme de faits des critères comme par exemple la prise en compte des aspects temporels:

If <école> AND <entre_minuit_et_6_heure> THEN <tronçon_sécuritaire> 0.8

Comme il a été démontré au chapitre 5, les systèmes experts opèrent selon une approche déductive (chaînage avant) analogue à celle d'un expert humain. L'évaluation du risque que représentent 2 ou plusieurs tronçons passera donc par une *procédure d'évaluation des préférences* par chaînage avant; c'est-à-dire des faits (caractéristiques du tronçon, règles de production) vers la conclusion (tronçon préféré). Étant donné que nous sommes dans un environnement opérationnel où il y aura données incomplètes et/ou incertaines, la conclusion

Formalisation du modèle de recherche d'itinéraire

qui sera atteinte par le système expert s'établira sur un "spectre de préférence" allant du tronçon totalement sécuritaire au tronçon à risque très élevé. Ce spectre de préférence correspond en fait à un ensemble flou (chapitre 5) où un degré d'appartenance de "1" indique que le tronçon considéré est assurément un tronçon sécuritaire alors qu'un degré d'appartenance de "0" nous indique que ce tronçon n'appartient assurément pas à cet ensemble. Pour faciliter l'interprétation des résultats, ce spectre de préférence peut être reconverti en 5 *classes de préférences* selon la représentation symbolique suivante:



Dans cette optique, le modèle de recherche d'itinéraire a pour objectif de déterminer l'itinéraire ayant la plus grande préférence cumulative avec comme priorité: 1)d'éviter les tronçons présentant un risque trop élevé (selon l'avis du système expert) 2)en présence de tronçons à risques similaires ou en cas d'absence de risque, sélectionner le tronçon le plus court. Chaque tronçon parcouru dans le réseau doit être évalué par le système expert (procédure d'évaluation des préférences) et dépendamment du type de matière transportée et de la période de la journée, la conclusion atteinte par le système expert pourra différer pour un même tronçon.

En modifiant une ou plusieurs règles de production de la base de connaissances du système expert, l'utilisateur est en mesure de visualiser l'impact de cette modification sur le choix de l'itinéraire et, grâce au module d'explication des résultats du système, ce dernier pourra justifier ses conclusions. Avec une telle approche, il devient beaucoup plus facile d'expérimenter des

scénarios pour faciliter la tâche des preneurs de décisions puisque seul l'ajout ou le retrait des règles de production est requis.

Pour mieux répondre aux exigences de la recherche d'itinéraires sécuritaires, trois stratégies de recherche ont été développées dans le cadre de cette thèse. Il s'agit de la sélection dynamique des tronçons sécuritaires; de l'approche par "K" plus courts chemins et enfin de la sélection à priori des tronçons sécuritaires.

7.2.1 Sélection dynamique des tronçons sécuritaires

La stratégie de sélection dynamique des tronçons sécuritaires consiste à procéder à l'évaluation des préférences à l'intérieur même du processus itératif de recherche d'itinéraire. Dans l'approche classique telle que proposée par l'algorithme de Dijkstra (voir chapitre 3), l'heuristique consiste à visiter chaque noeud du réseau en retenant à chaque itération l'arc offrant la valeur cumulative d'impédance la plus faible (figure 7.1). De cette manière on s'assure, après avoir visité chaque noeud du réseau, d'avoir découvert le plus court chemin entre 2 points donnés.

C'est au niveau de cette heuristique que nous avons apporté des modifications pour bâtir notre modèle. Il s'agit en fait de remplacer l'heuristique classique par la procédure d'évaluation des préférences. De façon pratique, à cette étape de l'itération, un appel est fait au système expert qui, à l'aide des caractéristiques de chaque tronçon et par chaînage avant, détermine lequel des tronçons est le plus sécuritaire (figure 7.2).

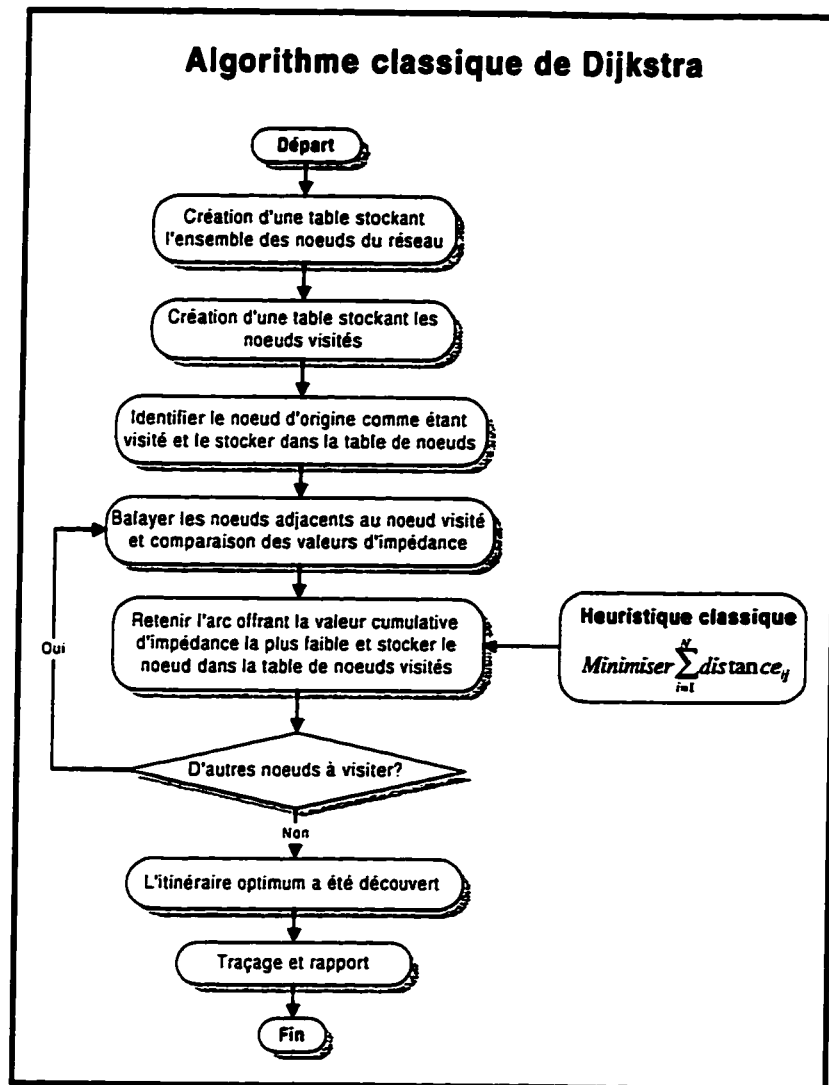


Figure 7.1: Heuristique classique de l'algorithme de Dijkstra

Alors que, dans l'approche classique, l'objectif consiste à minimiser la distance parcourue,

$$\text{Minimiser } \sum_{i=1}^N \text{distance}_{ij}$$

où:

N = Nombre de sommets

Distance_{ij} = Distance entre les sommets i et j

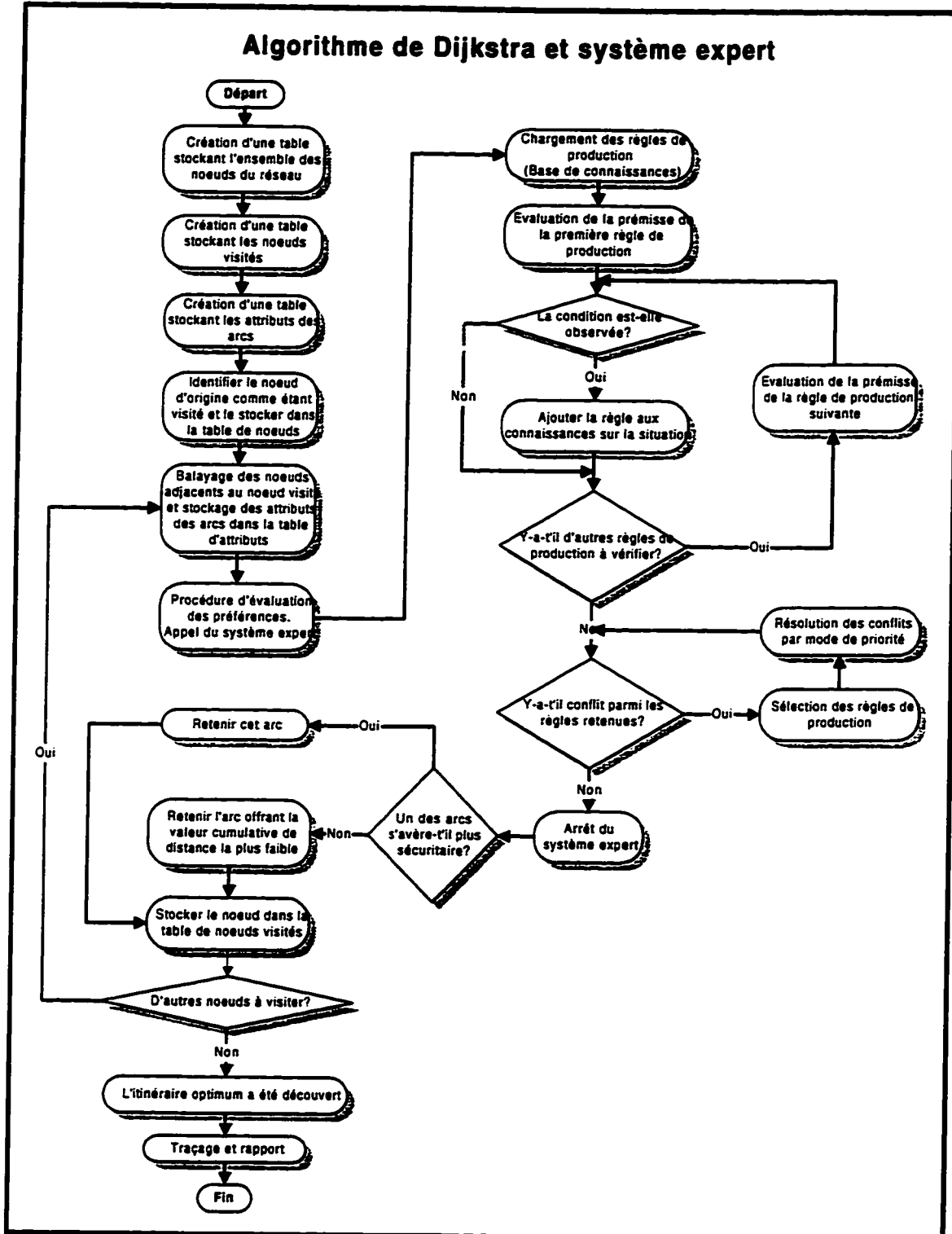


Figure 7.2: Modèle de recherche d'itinéraire à l'aide d'un système expert

Formalisation du modèle de recherche d'itinéraire

dans l'approche par système expert, on vise à maximiser la préférence, cette dernière étant issue du processus d'inférence du système expert.

$$\text{Maximiser } \sum_{i=1}^N \text{Préférence}_{ij}$$

où:

N = Nombre de sommets

Préférence_{ij} = Tronçon le plus sécuritaire selon l'évaluation du système expert

La sélection dynamique est une approche qui favorise nettement l'aspect sécuritaire de la recherche d'itinéraire puisqu'à chaque itération le tronçon retenu est le tronçon offrant la plus grande sécurité. Le tronçon le plus court ne sera retenu qu'en cas d'égalité en terme de sécurité. Une autre particularité de ce modèle réside dans la possibilité de l'appliquer à un système de guidage en temps réel (voir chapitre 2). Le système expert pourrait recevoir, via un système de collecte des paramètres routiers, toute l'information en temps réel sur chacun des tronçons routiers et réagir à des changements soudains de conditions routières tels qu'un embouteillage ou un tronçon bloqué en raison d'un accident.

L'approche par sélection dynamique représente cependant un inconvénient important en ce qui concerne la puissance de traitement requise. Dans sa forme standard, l'algorithme de Dijkstra nécessite N^2 comparaisons (N =nombre de noeuds) donc un nombre équivalent d'appels au système expert (ARLINGHAUS *et al*, 1990, p. 352). Pour un réseau de taille modeste d'environ 500 noeuds, cela représente $500 \times 500 = 250\,000$ appels à la procédure d'évaluation des préférences! Pour opérer ce modèle sur un réseau urbain complexe (environ 5000 noeuds), un traitement sur machine à architecture parallèle (transputer) paraît tout indiqué. Pour remédier

Formalisation du modèle de recherche d'itinéraire

au problème du temps de calcul nécessaire à l'approche par sélection dynamique ainsi que pour présenter un modèle qui ne pénalise pas outre mesure les transporteurs routiers, un modèle de sélection à postériori des itinéraires sécuritaires a été développé.

7.2.2 "K" plus courts chemins: sélection à postériori de l'itinéraire sécuritaire

Alors que la sélection dynamique s'attarde à déterminer un seul itinéraire optimum basé sur la maximisation des préférences, l'approche par "K" plus courts chemins consiste à calculer "K" plus courts chemins en introduisant des déviations à l'itinéraire optimum de départ. Donc, à partir d'une solution optimum, on recherche un ensemble d'itinéraires qui, sans être optimum, s'en approche sensiblement. Le résultat est une série de solutions potentielles ("K" plus courts chemins -> "K" solutions) qui peuvent par la suite être analysées par le système expert qui sélectionnera l'itinéraire le plus sécuritaire.

L'algorithme que nous avons développé se distingue de l'approche que l'on retrouve habituellement dans la littérature traitant du problème des "K" plus courts chemins. Notre "recette maison" consiste en une extension de l'algorithme de Dijkstra, ce qui en facilite la compréhension et la comparaison avec l'approche par sélection dynamique. Une fois le nombre de "K" plus courts chemins fixé par l'utilisateur, la première étape consiste à déterminer le chemin le plus court (K_1) à l'aide de l'algorithme de Dijkstra (figure 7.3). Une fois obtenu, on stocke cet itinéraire optimum qui servira de référence aux étapes subséquentes. A l'étape suivante, on calcule les $K^{\text{ième}}$ plus courts chemins en "déconstruisant" à rebours l'itinéraire optimum K_1 .

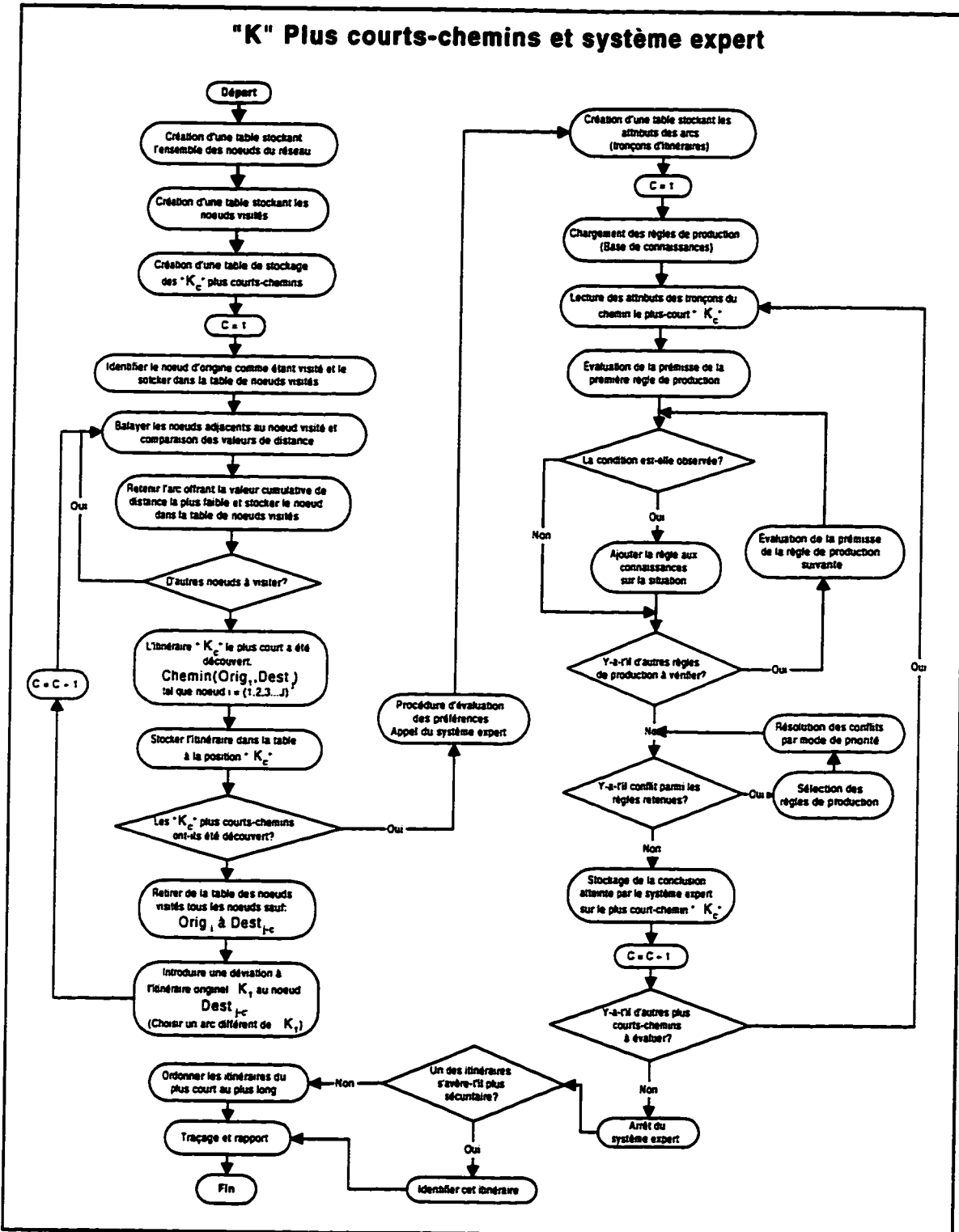


Figure 7.3: "K" plus courts chemins et système expert

Formalisation du modèle de recherche d'itinéraire

Si l'on note $Dest_i$ comme étant le noeud de destination, le calcul du deuxième plus court chemin consiste à introduire une déviation à l'itinéraire optimum K_1 au noeud $Dest_{i-1}$, le troisième plus court-chemin au noeud $Dest_{i-2}$, ... $Dest_{i-3}$, etc. Cette déviation consiste à choisir un tronçon, au noeud $Dest_{i-c}$, différent de celui ayant été retenu à l'itinéraire K_1 . A l'aide de l'algorithme de Dijkstra, on recalcule alors le plus court chemin, à partir de cette déviation, jusqu'au noeud de destination. Ce processus itératif se poursuit jusqu'à ce que les "K" plus courts chemins aient été découverts ou que l'ensemble de l'itinéraire K_1 ai été déconstruit. Avec cette approche, la valeur maximum que peut prendre "K" est le nombre de noeuds-1 composant l'itinéraire K_1 . Il serait possible d'augmenter cette valeur limite de "K" en déconstruisant les itinéraires $K_1, K_2, K_3, \dots K_c$, mais le temps de calcul nécessaire ainsi que la tendance à s'éloigner graduellement de l'itinéraire optimum en limite l'intérêt.

Une fois les "K" plus courts chemins calculés, la procédure d'évaluation des préférences est enclenchée et le système expert entre en fonction. Chaque itinéraire est évalué, tronçon par tronçon, à l'aide des attributs stockés préalablement dans une table. La conclusion atteinte par le système expert est stockée dans un fichier et le résultat final est une évaluation du degré de sécurité que présente chaque itinéraire. Pour les preneurs de décisions, l'un des avantages de cette approche réside dans le fait qu'ils ont en mains non pas une solution unique mais un ensemble de solutions potentielles envisageables.

Il s'agit également d'une méthode qui ne défavorise pas les transporteurs puisqu'à la base on retrouve le calcul d'un certain nombre de plus courts chemins. On peut supposer qu'un itinéraire le plus court possible présente un minimum de dépenses pour les transporteurs. Cette

Formalisation du modèle de recherche d'itinéraire

approche prend tout autant en considération les aspects sécurité puisque le système expert porte un diagnostic en tenant compte des caractéristiques des tronçons de chaque itinéraire. Au besoin, si aucun des "K" plus courts chemins calculés n'est satisfaisant au point de vue sécurité, l'analyse peut être reprise en augmentant la valeur de "K" pour ainsi avoir un plus grand nombre d'itinéraires potentiels. Comparativement à la sélection dynamique, la sélection à postériori des itinéraires est beaucoup moins exigeante en terme de traitement informatique. Le système expert n'est exécuté qu'après l'accomplissement des nombreuses itérations nécessaires pour balayer l'ensemble du réseau. Dans ce cas, le nombre d'appels au système expert équivaut au nombre "K" de plus courts chemins (figure 7.3).

Le principe sous-jacent à la méthode de sélection à postériori est que si l'itinéraire K_1 (le plus court) est inacceptable pour des raisons de sécurité dues à un ou plusieurs tronçons risqués, l'introduction systématique d'une déviation à chaque noeud devrait conduire à un itinéraire acceptable du point de vue sécurité. Une déviation qui aurait pour effet d'éviter de passer devant un hôpital ou un centre d'achats fera diminuer le risque, donc hausser la préférence de cet itinéraire. Cependant, en pratique, la performance de ce modèle varie selon la densité du réseau de transport auquel il est appliqué. Dans les réseaux denses, de type urbain, l'écart entre la longueur totale de l'itinéraire K_1 et les $K^{\text{ième}}$ itinéraires est généralement minime car l'on est en présence de tronçons routiers très courts et à forte densité d'intersections. Une déviation donnée pourra n'engendrer que quelques centaines de mètres ou quelques kilomètres de plus à parcourir. Par contre, en milieu rural, la faible densité du réseau agit de façon inverse. Une déviation donnée pourra introduire une distance supplémentaire à parcourir de plusieurs dizaines de kilomètres ou plus encore. Certaines déviations peuvent également aboutir à des

situations impraticables. Par exemple, en région frontalière, une déviation qui aura pour effet de traverser sans raison la frontière est à exclure par le système expert.¹ D'autres déviations qui impliquent de franchir des obstacles naturels majeurs peuvent aussi déboucher sur des itinéraires désavantageux. Par exemple, une déviation qui ferait en sorte d'emprunter la rive nord du fleuve St-Laurent plutôt que la rive sud pourrait conduire à un itinéraire des centaines de kilomètres plus long que prévu!

Bien que le modèle de sélection à postériori des itinéraires ne se prête pas aussi bien que le modèle par sélection dynamique aux systèmes de guidage en temps réel, il présente plusieurs avantages sur ce dernier. Sa moins grande exigence en terme de traitement informatique, son ensemble d'itinéraires offerts aux preneurs de décisions ainsi que son compromis entre plus court chemin et sécurité des itinéraires le rendent particulièrement intéressant. Contrairement à la sélection à postériori, il est également possible de faire une sélection au préalable des tronçons avant de procéder au calcul d'itinéraire. On parle alors de sélection à priori des tronçons sécuritaires.

7.2.3 Sélection à priori des tronçons sécuritaires

En présence d'un réseau complexe et dense, la sélection dynamique et même la sélection par "K" plus courts chemins peuvent exiger des temps de traitement considérables. Il nous est apparu à propos de développer une méthode "expéditive" permettant de simplifier un réseau de transport préalablement à son analyse. Le moyen le plus simple d'y arriver est de procéder à une opération de filtrage sur chaque arc du réseau en éliminant tout arc ne répondant pas à un

¹On peut contourner ce problème en bloquant tous les tronçons menant à la frontière. De cette manière, les plus courts chemins calculés exclueront automatiquement ces tronçons "culs-de-sac".

Formalisation du modèle de recherche d'itinéraire

seuil de sécurité minimal. La procédure d'évaluation des préférences est appliquée à chaque arc par le système expert et tout arc obtenant la classe de préférence <risque élevé>, <risque très élevé> est rejeté. Un nouveau réseau est reconstruit sans les arcs rejetés, mais composé des arcs <sans risque>, <risque faible>, <risque moyen>. Par la suite, l'algorithme de Dijkstra est appliqué au réseau élagué et le chemin le plus court est déterminé (figure 7.4).

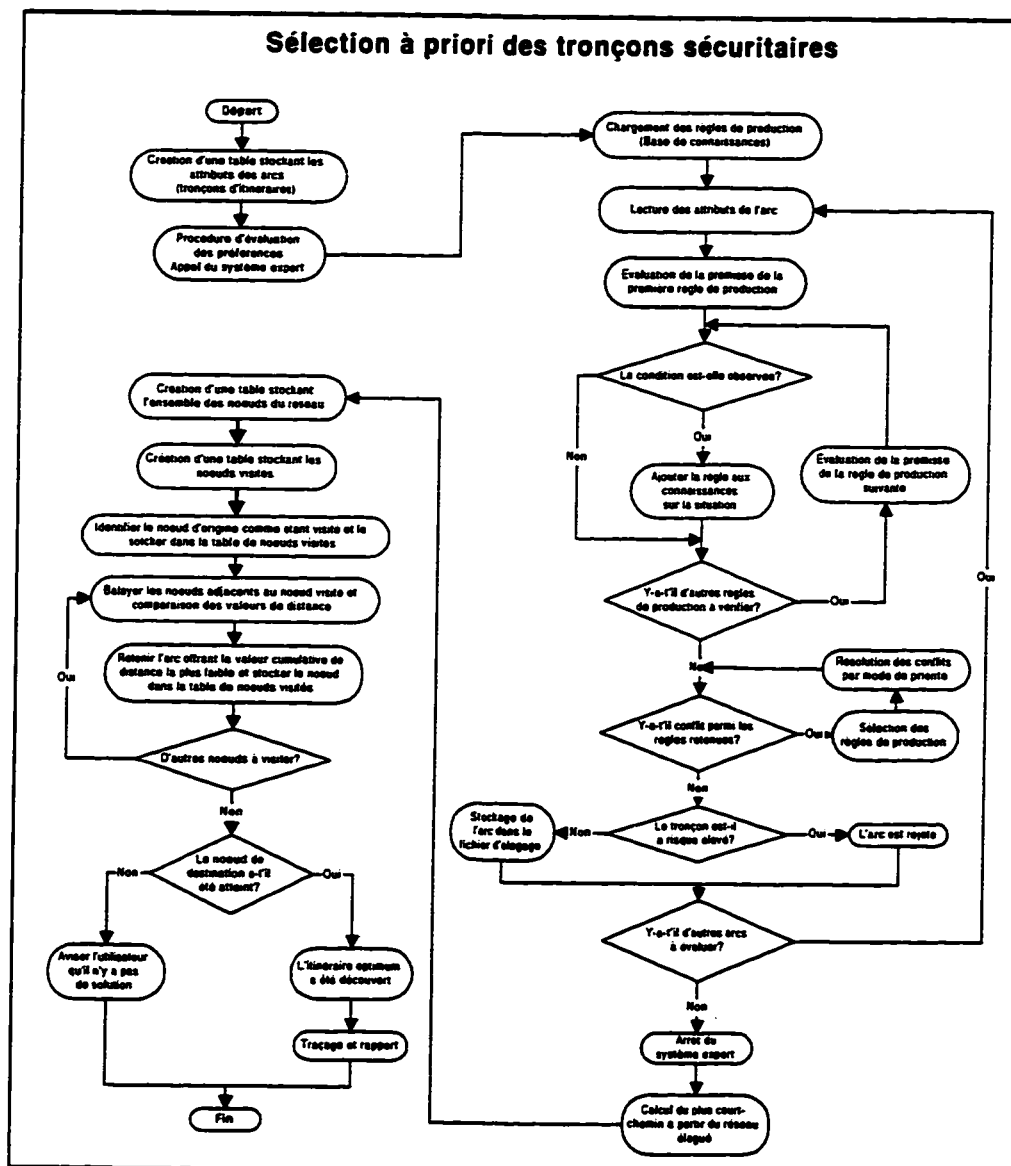


Figure 7.4: Sélection à priori des tronçons sécuritaires

Formalisation du modèle de recherche d'itinéraire

Une fois le réseau élagué produit, il est possible d'y calculer autant d'itinéraires que nécessaire en modifiant origine et destination puisqu'il s'agit d'un nouveau réseau sans les tronçons à risque élevé. Il y a un gain d'efficacité majeur sur les deux modèles précédents si l'on doit calculer de nombreux itinéraires pour une même matière dangereuse et un même niveau de sécurité. Ce gain est essentiellement dû à l'absence d'appel au système expert qui a pour effet de ralentir l'analyse. Cette approche expéditive n'est cependant pas sans problème. Dans certaines situations le processus d'élagage aura pour effet de rendre des portions du réseau inaccessibles (figure 7.5)

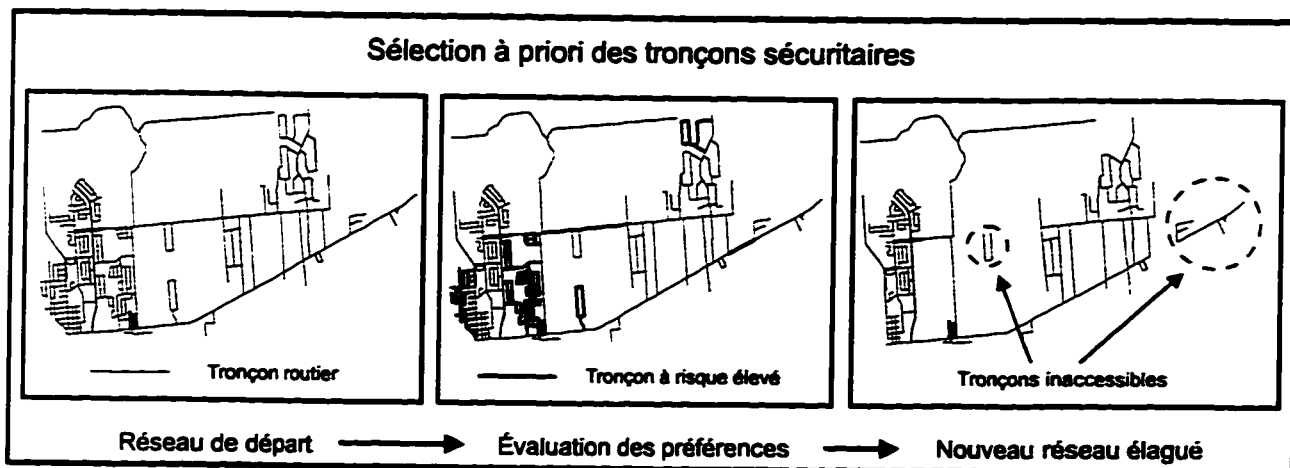


Figure 7.5: Limitations de la sélection à priori des tronçons sécuritaires

Si l'origine ou la destination de l'itinéraire se situent dans l'une de ces sections, aucune solution d'itinéraire n'est possible. On doit alors reprendre le processus d'élagage avec une classe de préférence moins élevée (moins grande sécurité) en espérant que les portions isolées du réseau seront reconnectées. Il s'agit donc d'une méthode à privilégier seulement si les gains en traitement sont considérables (réseaux complexes, nombreux itinéraires) en

gardant toujours en tête que si aucune solution n'est possible, on doit reprendre le tout avec l'approche par sélection dynamique ou par sélection à postériori.

Que l'on préconise l'une ou l'autre de ces approches, la réalisation d'un système de support à la décision à référence spatiale exige l'intégration des modules d'analyse aux autres composantes que sont le système expert et le système d'information géographique. Plus cette intégration sera complète, meilleur sera le système de support à la décision à référence spatiale.

7.3 Architecture du système de support à la décision à référence spatiale

La fonction principale d'un système de support à la décision à référence spatiale est de permettre aux preneurs de décisions de choisir en toute connaissance de cause, en tenant compte du plus grand nombre de facteurs possibles. Ce système doit permettre d'intégrer les différentes sources d'information et de les exploiter adéquatement. Il doit, par ses différentes composantes, offrir de visualiser le réseau et son territoire, d'analyser spatialement et logiquement les données et effectuer des choix. La figure 7.6 illustre la structure du SSDRS proposé. Essentiellement, cette structure reprend les différentes composantes SIG-SE vues aux chapitres 4 et 5. La seule différence réside dans l'ajout d'une interface SIG-SE qui agit comme lien entre le SIG et le système expert.

Ce lien repose sur une structure de données commune qui correspond à la "syntaxe de base" des différentes relations topologiques et spatiales dans les réseaux. Un SSDRS adapté à la

Formalisation du modèle de recherche d'itinéraire

recherche d'itinéraire doit être en mesure de traiter 5 grandes catégories de relations: 1) la connectivité 2) l'orientation 3) les relations arc-arc 4) les relations arc-noeud 5) la succession d'arcs (route).

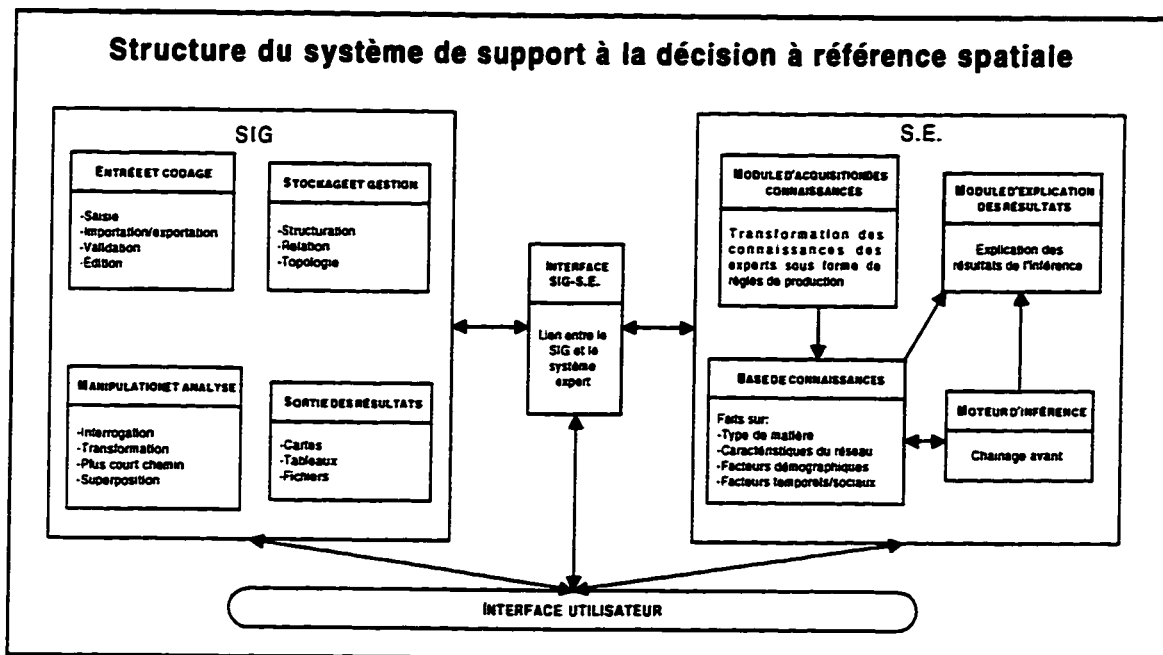


Figure 7.6: Structure du système de support à la décision à référence spatiale

Chacune de ces relations peut se voir associer toute une série d'attributs décrivant les caractéristiques de la relation. A ce niveau, le SIG occupe une place importante puisque c'est par son entremise que l'on parvient à construire ces relations et à y rattacher les attributs correspondants. Par la suite, ce sont également ces attributs qu'utilise le système expert pour sélectionner les tronçons sécuritaires.

Connectivité:

La connectivité indique s'il y a présence d'un lien (route) entre deux noeuds du réseau. Il s'agit d'une relation fondamentale, à la base de la recherche d'itinéraire.

LINK (Arc_ID, Node_ID, Node_ID), [ATTRIBUTS...]

Formalisation du modèle de recherche d'itinéraire

Orientation:

La relation d'orientation indique la direction du flux dans les arcs. Selon le type de réseau en question, les déplacements pourront être à sens unique ou dans les deux directions. De plus, certaines données sont directionnelles, comme par exemple les débits de circulation.

LINK (Arc_ID, From_Node_ID, To_Node_ID), [ATTRIBUTS...]

Relation arc-arc:

La relation arc-arc sert à déterminer les mouvements possibles ou permis aux intersections. On retrouve cette relation dans les SIG sous le terme de "Turn table", c'est-à-dire un tableau indiquant s'il est possible d'effectuer un virage à gauche ou à droite à une intersection donnée. Cette information est essentielle puisque dans certains cas, deux tronçons "topologiquement reliés" peuvent être inaccessibles en raison d'un virage interdit. Le nombre total de mouvements possibles à une intersection donnée est égal à L^2 , où L = nombre d'arcs composant l'intersection.

DIRECTION (Node_ID, From_Arc_ID, To_Arc_ID), [ATTRIBUTS...]

Relation arc-noeud:

La combinaison arc-noeud sert à représenter les obstacles ponctuels ou les barrières que l'on retrouve communément dans les réseaux routiers. Par exemple, un arrêt obligatoire ou un pont relèvent de ce genre de relation. Ce type de noeud se retrouve également sous le terme de pseudo-noeud.

NODE (Node_ID, Arc_ID), [ATTRIBUTS...]

Succession d'arcs (route):

Il peut s'agir d'un itinéraire préétabli ou encore d'un ensemble d'arcs de même catégorie ou ayant un attribut commun. Par exemple, tous les tronçons de route formant une route nationale.

TO_ARC (Arc_ID, To_Arc_ID), [ATTRIBUTS...]

FROM_ARC (Arc_ID, From_Arc_ID), [ATTRIBUTS...]

Ces relations topologiques ne sont pas nécessairement toutes présentes de façon explicite dans la structure de données. Par exemple, connectivité et orientation vont de pair puisque l'orientation d'un arc implique obligatoirement la connectivité. De plus, les relations arc-arc et arc-noeud peuvent être représentées dans la même table de virage car il s'agit de relations

Formalisation du modèle de recherche d'itinéraire

impliquant au moins 2 arcs et un noeud. Qu'elles soient implicites ou explicites, ces relations demeurent à la base même de la modélisation des mouvements dans un réseau. Outre la structure de données et ses relations topologiques, le SSDRS doit également comporter une base de connaissances dont la structure reflète le type de connaissances nécessaires au système expert pour résoudre le problème du transport des matières dangereuses.

7.4 Structure de la base de connaissances du système expert

La base de connaissances d'un système expert constitue un modèle de la représentation que se font les experts "humains" du monde réel. Pour élaborer cette base il faut, dans un premier temps, identifier les caractéristiques du problème puis, dans un deuxième temps, organiser les connaissances en groupes homogènes permettant par la suite de formaliser et d'implanter le modèle. Dans les projets de grande envergure, ce travail d'élaboration est effectué en collaboration avec un ou des experts du domaine d'application ainsi que par un spécialiste en système expert qui verra à traduire sous forme acceptable par le système les faits et les règles du domaine d'application (figure 7.7)

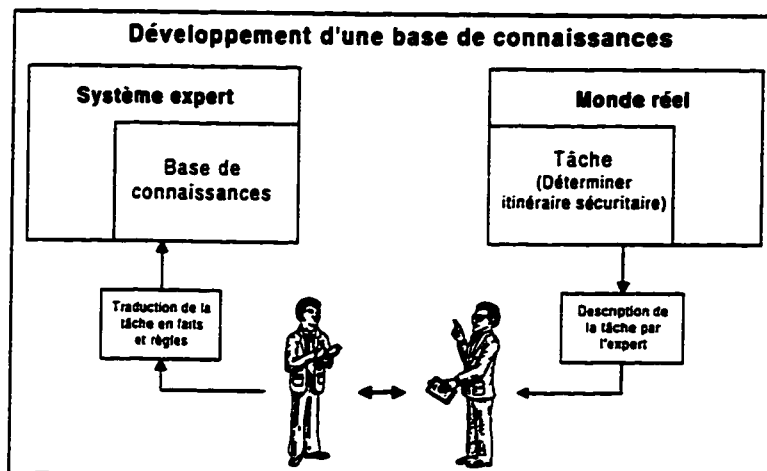


Figure 7.7: Développement d'une base de connaissances

Formalisation du modèle de recherche d'itinéraire

Dans le cadre de ce projet nous avons regroupé en 5 grands blocs homogènes les connaissances nécessaires à la sélection d'un itinéraire de transport sécuritaire. Ces connaissances peuvent être de nature qualitative ou quantitative et permettre de raisonner sur les: 1)caractéristiques du réseau 2)caractéristiques de la matière transportée 3)caractéristiques de la population à risque 4)règlements de transport 5)installations et sites vulnérables.

Caractéristiques du réseau.	Tous les faits et les règles de production relatifs aux attributs des tronçons tels que la classification fonctionnelle, le nombre de voies, le débit de circulation, le taux d'accidents, la présence de ponts ou de tunnels et l'utilisation du sol en bordure des tronçons.
Caractéristiques de la matière dangereuse.	S'agit-il d'une substance fortement acide ou corrosive, d'un explosif ou d'une substance radioactive? Par exemple, une substance explosive présente peu de danger pour l'environnement mais un danger considérable pour la population. A l'inverse, une substance corrosive est relativement inoffensive pour la population mais très néfaste pour l'environnement s'il y a déversement dans un cours d'eau.
Caractéristiques de la population à risque.	Dénombrement de la population aux abords de chaque tronçon. Type de population impliquée (résidents/travailleurs) ainsi que type et densité du bâti (édifices à appartements, tours, unifamiliale, etc.).
Installations et sites vulnérables.	Tous les sites présentant une vulnérabilité particulière aux matières dangereuses, soit en raison de la population qui s'y trouve (écoles, hôpitaux, parcs, etc), soit en raison de critères écologiques ou économiques (zone écologique sensible, source d'eau potable, plan d'eau, etc.).
Règlements de transport	Restrictions et interdictions stipulées par les règlements fédéraux, provinciaux et municipaux. Certaines routes ou tronçons de routes sont interdits aux camions ou aux matières dangereuses.

Évidemment, ces "blocs de connaissances" ne sont pas tous au même niveau dans le processus d'inférence. Une hiérarchisation est nécessaire car certains faits et certaines règles de production auront un impact sur l'ensemble du raisonnement devant conduire au choix d'un

Formalisation du modèle de recherche d'itinéraire

itinéraire. C'est le cas notamment du type de matière transportée et des règlements de transport qui doivent être considérés en premier lieu (figure 7.8).

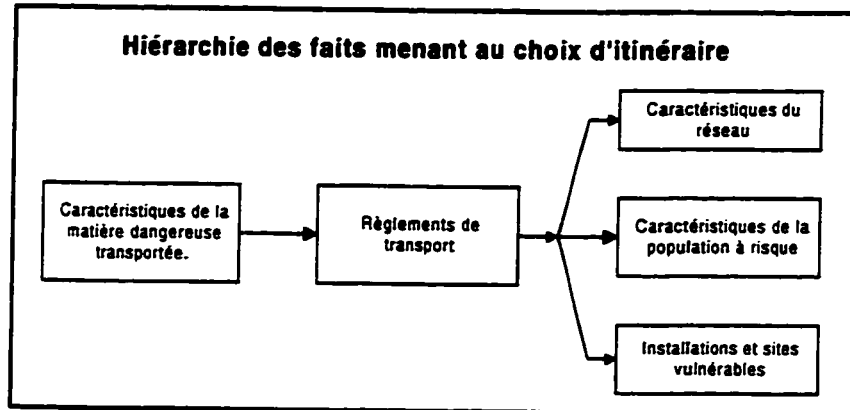


Figure 7.8: Hiérarchisation des faits menant au choix d'un itinéraire

En effet, le type de matière transportée doit être en amont du processus d'inférence puisque c'est de lui que découle un danger potentiel (explosion, incendie, etc.) ou certaines interdictions (exemple: les explosifs sont interdits dans les tunnels). Quant aux règlements de transport, certaines interdictions ou restrictions auront pour effet d'éliminer dès le départ certains tronçons ou certaines routes. En ce qui concerne les trois blocs restants, ils peuvent être considérés au même niveau, ou en séquence selon les désirs des preneurs de décisions ou de la population.

Il faut toutefois bien garder à l'esprit que cette hiérarchisation n'implique pas que les caractéristiques de la matière transportée ont plus de poids dans le modèle que les règlements de transport. Cette hiérarchie, ou séquence, indique simplement que, logiquement, il faut tout d'abord initier le processus d'inférence sur le type de matière transportée avant de balayer les faits concernant les règlements de transport. De façon pratique, cette hiérarchie se traduira par

Formalisation du modèle de recherche d'itinéraire

une séquence précise des règles de production dans la base de connaissances du système expert.

Maintenant que le survol des différentes notions théoriques nécessaires à la compréhension et à l'implantation d'un système de support à la décision a été réalisé, la prochaine étape consiste à procéder à son implantation. Bien que le but de cette recherche ne soit pas de réaliser un logiciel parfaitement fonctionnel (tâche qui relève du domaine technique de la programmation), le prochain chapitre a pour objectif de démontrer la démarche d'implantation des différentes stratégies de recherche d'itinéraire proposées.

Chapitre VIII

Élaboration du système de support à la décision à référence spatiale

“L'Homme n'est rien d'autre que la série de ses actes.”

- G.W.F Hegel

La réalisation d'un projet impliquant un système d'information géographique ou un système de support à la décision nécessite des centaines d'heures de travail. Il est courant, en ce qui concerne les SIG, de consacrer de 50% à 75% des ressources financières et du temps de travail à l'élaboration des bases de données qui formeront l'assise du système. Développer un système parfaitement fonctionnel “clefs-en-mains” n'était donc pas l'objectif de ce projet de recherche, non plus que de régler le problème du transport des matières dangereuses pour une région donnée. La démarche d'implantation visait plutôt à démontrer comment le modèle proposé peut-être implanté en mettant l'emphase sur les aspects méthodologiques. De plus, n'ayant pas les ressources financières nécessaires pour la consultation d'un panel d'experts en transport des matières dangereuses, nous ne pouvons

prétendre avoir réalisé une base de connaissances exhaustive sur le domaine d'application. Malgré ces limitations, cet exercice d'implantation nous a permis de dégager les différents problèmes inhérents aux aspects pratiques de la démarche. Ce chapitre présente donc la démarche d'implantation, les sources et l'élaboration de la base de données nécessaire à un tel projet ainsi que les différents modules développés pour répondre à nos besoins.

8.1 Démarche d'implantation du système de support à la décision à référence spatiale

Notre démarche d'implantation du SSDRS peut être divisée en sept grandes étapes (figure 8.1).

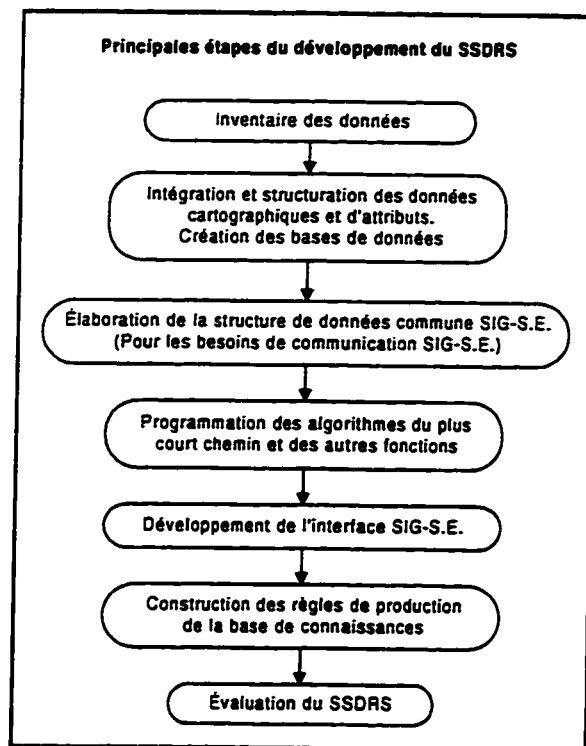


Figure 8.1: Étapes de l'implantation du SSDRS

Ces étapes ne font que présenter de façon globale le déroulement général du processus puisqu'en pratique chacune de ces phases peut être subdivisée en de nombreuses sous-étapes et boucles de rétroaction. Certaines de ces étapes revêtent un caractère plus technique, notamment l'inventaire des données cartographiques et d'attributs ainsi que l'intégration et la structuration des bases de données. D'autres phases ont nécessité de faire des choix ayant un impact sur l'ensemble du système. C'est le cas de la structure de données commune SIG-S.E. et des règles de production de la base de connaissances. Enfin, deux étapes ont exigé de nombreuses heures de programmation, soit le développement de l'interface SIG-S.E. et la programmation des algorithmes de calcul du plus court chemin. Fait à souligner, dans le cadre de ce projet, nous avons réduit au minimum la phase de construction des règles de production de la base de connaissances. A elle seule cette étape forme une discipline en intelligence artificielle, le génie des connaissances (knowledge engineering), avec ses spécialistes et ses méthodes parfois fort complexes.

8.2 Région à l'étude et sources de données

C'est la région de la capitale nationale qui a été retenue comme région d'étude et ce principalement en raison de sa taille, de sa proximité et de l'accès aux données. La zone couverte comprend les municipalités de la Municipalité Régionale d'Ottawa-Carleton (MROC) du côté ontarien ainsi qu'une partie des municipalités d'Aylmer, Hull et Gatineau du côté de l'Outaouais québécois (figure 8.2). Avec une population totale de plus de 800 000 habitants, la région de la capitale offre toutes les caractéristiques d'une grande agglomération. On y retrouve un noyau central fortement urbanisé mais aussi une frange rurale importante en raison de la ceinture verte (figure 8.3).

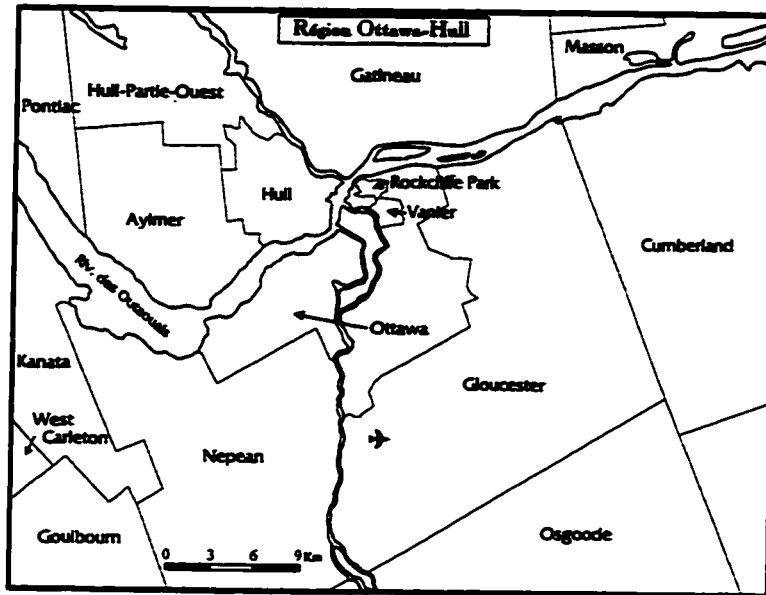


Figure 8.2: Région à l'étude

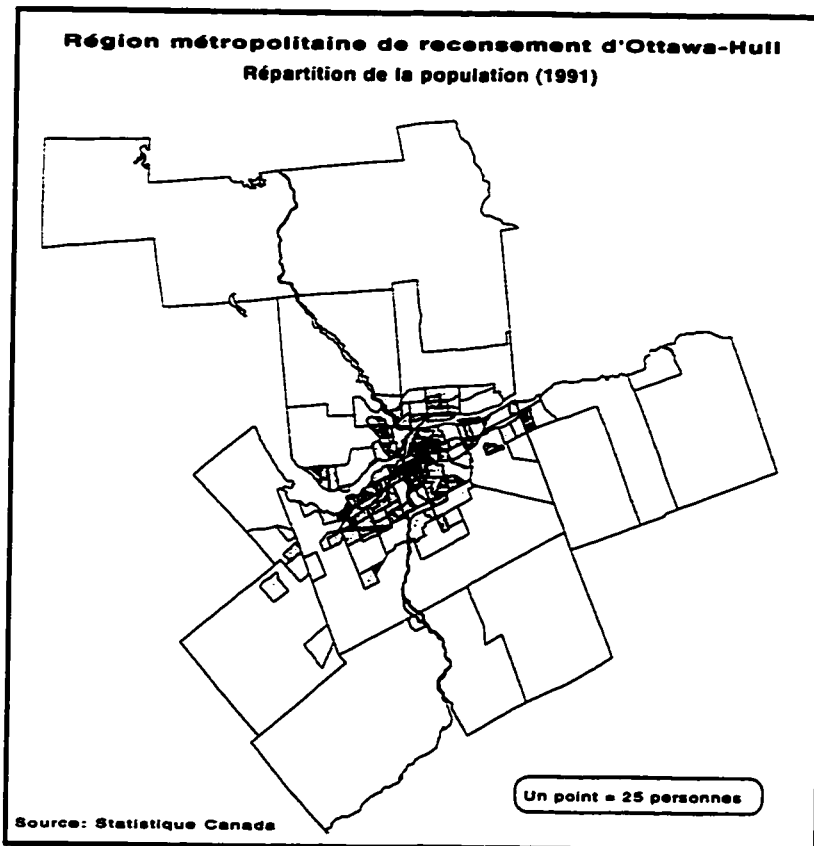


Figure 8.3: Répartition de la population

Parmi les principales utilisations du sol on note le résidentiel de faible à forte densité, des aires commerciales ou industrielles, des espaces de bureau ainsi que des zones agricoles et de loisirs. En raison de son statut de capitale, la présence de sites touristiques et historiques y est également accentuée. Cette grande variabilité dans les types d'activités et d'occupation du sol offre une large gamme de population à risque dans le cadre du transport des matières dangereuses.

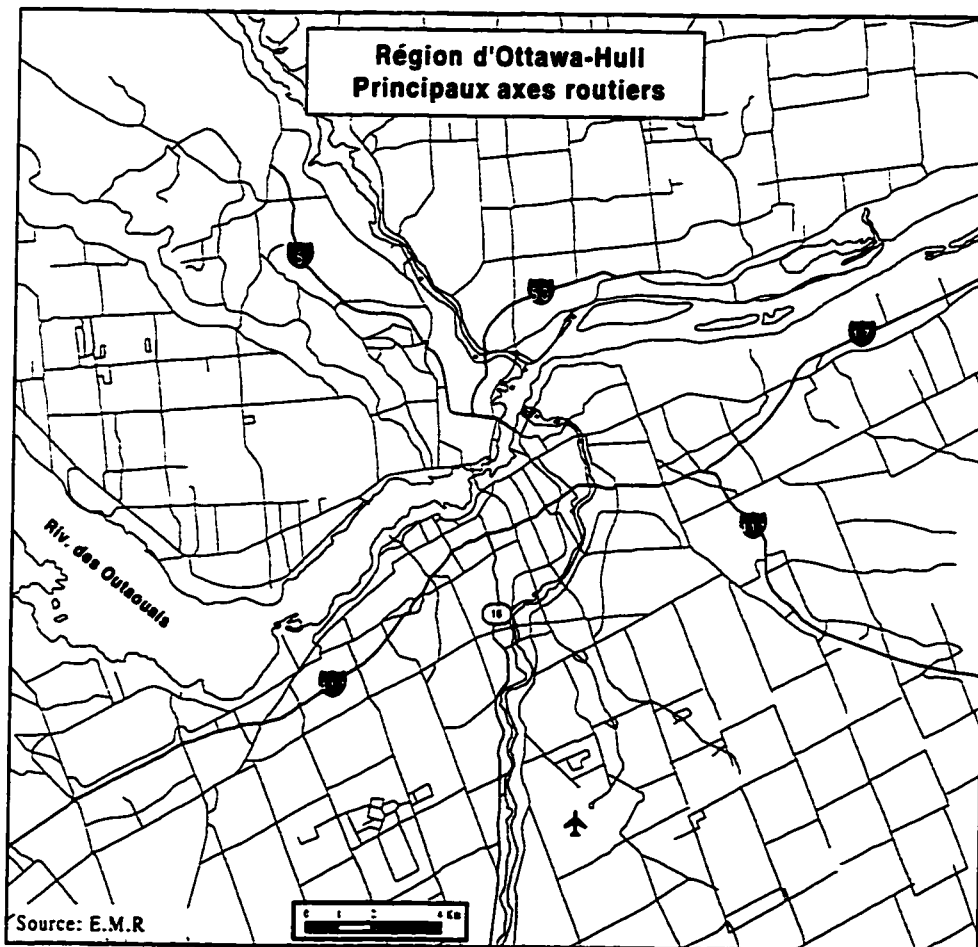


Figure 8.4: Grands axes routiers de la région

Du point de vue des infrastructures de transport, à l'échelle du 1:50 000, le réseau routier de la région s'étend sur plus de 4300 kilomètres de routes de toutes classifications. La présence de

la rivière des Outaouais a pour effet de scinder en deux parties le réseau de transport. L'autoroute 417, l'axe routier majeur, parcourt d'ouest en est la région en reliant Ottawa à Montréal, tandis que dans l'axe nord-sud, l'autoroute 5 est présente du côté québécois. Dans la portion ontarienne, c'est la route 16 qui fait office d'axe nord-sud en reliant Ottawa à l'autoroute 401 vers Toronto ou Cornwall (figure 8.4).

8.2.1 Les sources de données

Il faut distinguer deux grands types de données nécessaires au SSDRS, soit les données cartographiques pour la localisation spatiale des phénomènes et les données d'attributs permettant de caractériser ces mêmes phénomènes. Idéalement, plus un décideur aura accès à un grand nombre de données, plus la décision qu'il prendra se fera de façon éclairée. Le tableau 8.1 présente un inventaire des données qui pourraient faire partie du processus de

Réseau routier	<ul style="list-style-type: none"> -Carte du réseau routier -Classification fonctionnelle du réseau -Numéro de route (nom) -Nombre de voies -Direction (présence de sens unique) -Débit de circulation/tronçon -Vitesse permise -Taux d'accidents/tronçon
Caractéristiques de la population	<ul style="list-style-type: none"> -Carte des secteurs de dénombrement -Décompte de la population/tronçon (densité) -Type de population (résidents/travailleurs) -Caractéristiques du bâti <ul style="list-style-type: none"> -Nombre de logements -Type de logement (unfamiliale, édifices) -Utilisation du sol
Installations et sites vulnérables	<ul style="list-style-type: none"> -Inventaire et localisation des: <ul style="list-style-type: none"> -Écoles, hôpitaux, centres d'achats, bâtiments récréatifs, musés, parcs, passages à niveau, etc. -Zones écologiquement sensibles -Points d'approvisionnement en eau

Tableau 8.1: Inventaire des données pour la prise de décision.

recherche d'un itinéraire sécuritaire. Ces données sont regroupées selon trois des grands blocs de connaissances définis au chapitre précédent (tableau 8.1).

Évidemment, l'accès à toutes ces données n'est pas aisé puisque parfois certaines d'entre elles sont inexistantes ou incomplètes (débits de circulation, taux d'accidents) ou encore fort onéreuses à acquérir comme cela est le cas des données de Statistiques Canada¹. Dans le cadre de ce projet, nous nous sommes limités à un sous-ensemble de ces données en ne retenant que les éléments essentiels à la démarche d'implantation.

C'est la carte topographique sous format numérique à l'échelle du 1:50 000 (feuille 31 G/5) qui a servi de carte de base pour le réseau routier (figure 8.5). Un travail très laborieux de structuration de données a été nécessaire puisque ce type de fichier, distribué par Ressources Naturelles Canada, est fourni à l'état brut (nous discuterons ultérieurement de la structuration). Les éléments relatifs au réseau routier ont été extraits à partir des indicatifs d'éléments pour former une couche unique. La deuxième carte ayant servi au projet est celle de la région métropolitaine de recensement d'Ottawa-Hull (RMR) provenant de la base de données CARTLIB de Statistique Canada. C'est à partir de cette carte qu'ont été compilées les données relatives à la population et au type d'utilisation du sol. À ces documents cartographiques de base, plusieurs attributs ont pu être rattachés:

¹Le seul fichier du réseau de rue (Street Network File) de la région d'Ottawa-Hull implique des coûts d'acquisition de plusieurs milliers de dollars, coûts qui limitent fortement l'accès aux données.

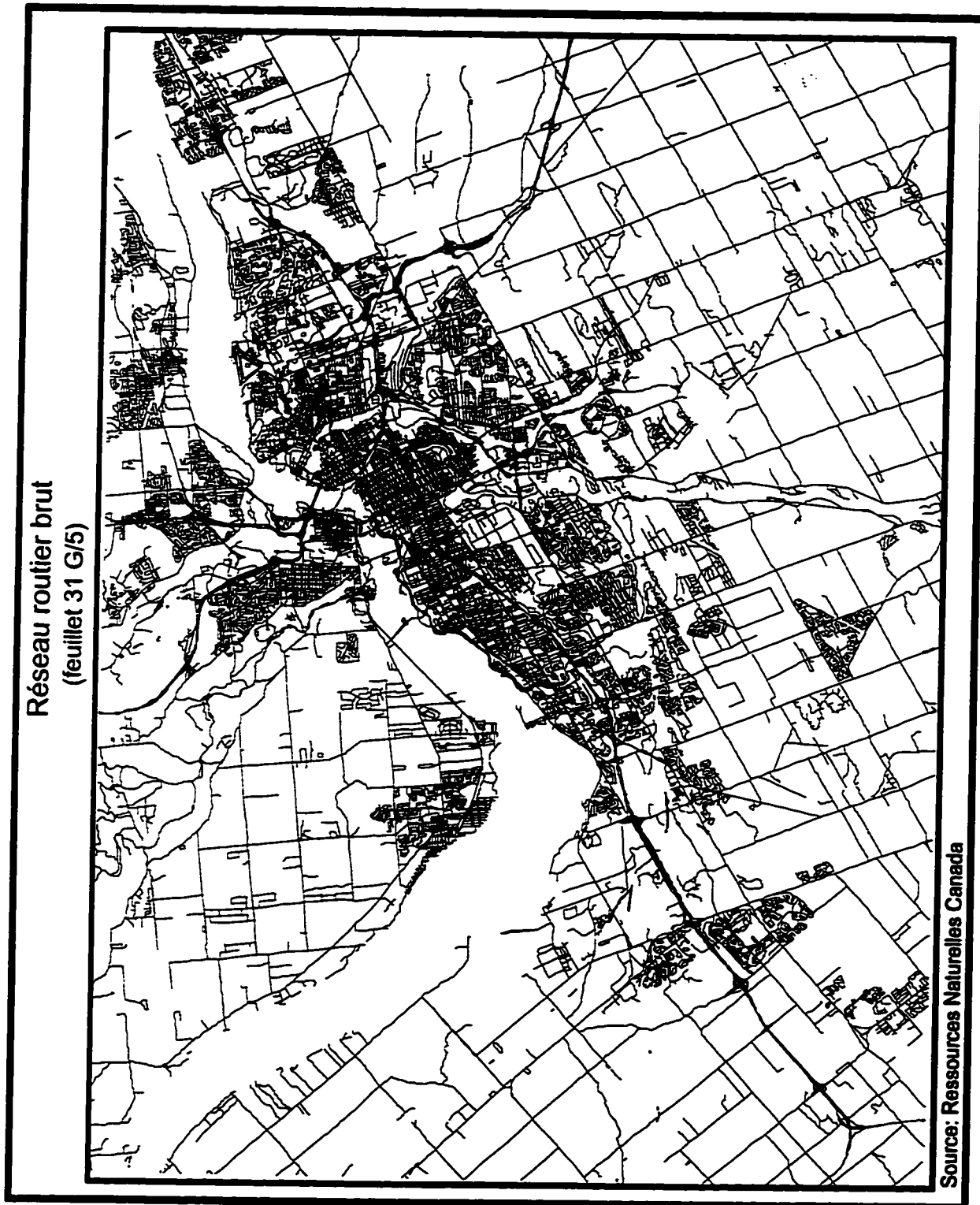


Figure 8.5: Réseau routier brut, région Ottawa-Hull

- Classification des routes (selon la carte topographique)**
 - Autoroutes à deux voies séparées
 - Routes pavées à plus de deux voies
 - Routes pavées à deux voies
 - Routes pavées à une voie

- Numéro de la route (selon la carte routière)**

- Nombre de voies (selon la carte topographique)**

- Direction (selon la carte routière)**

- Débit de circulation, lorsque disponible (Ministère des Transports de l'Ontario)**
 - Moyenne annuelle du trafic quotidien (information sporadique)

- Décompte de population par secteur de recensement (Statistique Canada)**
 - Selon le recensement de 1991 de Statistique Canada

- Nombre de logements par secteur de recensement (Statistique Canada)**
 - Selon le recensement de 1991

- Type de logement par secteur de recensement (Statistique Canada)**
 - Maisons unifamiliales
 - Édifices de plus de 5 étages

- Inventaire des sites vulnérables (diverses cartes de la région)**
 - Hôpitaux, écoles, parcs, plans d'eau, centres commerciaux

Il est à noter que ces données ne sont pas uniformes pour l'ensemble de la région à l'étude. En ce qui concerne les débits de circulation, ils sont très facilement disponibles (sous forme de listing papier) du côté ontarien, mais inaccessibles du côté de l'Outaouais québécois. Ce manque d'homogénéité dans les données n'est cependant pas unique à la région d'Ottawa-Hull et représente plutôt une caractéristique commune à ce genre de projet. En plus des données, la réalisation d'un SSDRS exige également le recours à plusieurs outils informatiques, notamment un système d'information géographique et un système expert.

8.2.2 Les outils essentiels au projet

Arc/Info-PC v3.4.2 est le logiciel de SIG qui a été retenu pour ce projet. Il est développé, avec une approche boîte à outils, par la firme ESRI (Environmental Systems Research Institute) depuis 1982. Arc/Info est le logiciel de SIG le plus répandu dans le monde (plus de 60 000 utilisateurs) et offre une gamme de fonctionnalités très riches et très complètes qui ne cessent de s'accroître en raison de son aspect modulaire. La version complète d'Arc/Info est fonctionnelle dans un environnement Unix (Sun ou Risk) tandis que la version IBM-PC est une version allégée avec plusieurs fonctions en moins et certaines limites au niveau de la gestion de la mémoire.

Arc/Info est également pourvu de son propre langage de programmation pour le développement ou la personnalisation d'outils d'analyse. En version Unix, ce langage est l'Arc Macro Language (AML) alors qu'en version PC, le Simple Macro Language (SML) est disponible en version allégée. Ce langage permet, entre autres, l'appel de fonctions externes dans la mesure où la mémoire disponible est suffisante. Du point de vue architecture, Arc/Info est composé de deux sous systèmes. Le premier, ARC, prend en charge les entités cartographiques tandis qu'INFO gère en mode relationnel les attributs de ces entités. Le lien entre ces deux sous systèmes est maintenu grâce à un identificateur commun (figure 8.6).

Le module ARCEDIT est dédié à la création et à l'édition des entités cartographiques; le module ARCPLLOT permet la rédaction et la production cartographique; le module OVERLAY la superposition des couches d'information (couvertures) et le module NETWORK les opérations sur les réseaux. Ce dernier module comporte une fonction de calcul d'itinéraire optimum basée

sur la notion d'impédance et l'algorithme de Dijkstra. Pour représenter de façon exhaustive un réseau routier dans Arc/Info, quatre types de fichiers sont nécessaires.

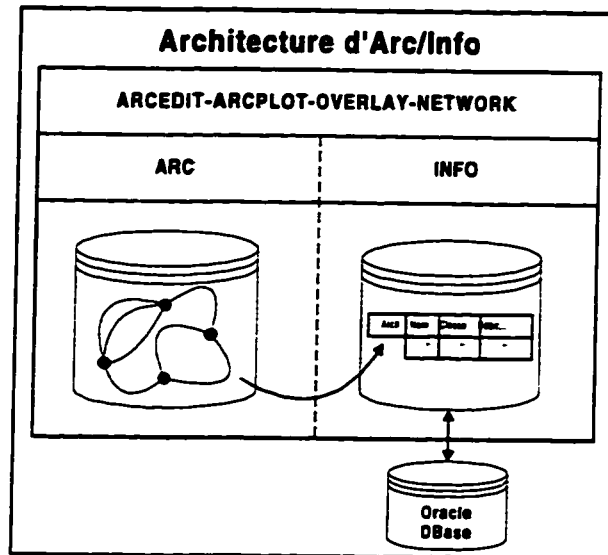


Figure 8.6: Architecture d'Arc/Info

Il faut un fichier **.ARC** qui contient la description géométrique des arcs; un fichier **.AAT** (Arc Attribute Table) renfermant les attributs de chaque arc; un fichier **.NAT** (Node Attribute Table) stockant les attributs de chaque noeud et enfin un fichier "Turntable" décrivant les mouvements possibles aux intersections. Accessoirement, un fichier **.PAT** (Polygon Attribute Table) est nécessaire pour traiter les attributs des entités polygonales (secteurs de recensement).

Outre le SIG Arc/Info, nous avons également eu recours à un système expert pour élaborer le SSDRS. Le système expert qui a été retenu est CLIPS v5.1 (C Language Integrated Production System) en raison de son coût d'acquisition minime et de l'accès au code source. CLIPS est le fruit de recherches en intelligence artificielle effectuées par le *Johnson Space Center* de la

NASA dans le but de réaliser un système expert portable et pouvant s'intégrer à des applications externes. CLIPS est développé en langage C et est en mesure de représenter les connaissances sous forme de règles de production ou par l'approche orientée-objet. Il traite les règles de production par chaînage avant, ce qui correspond aux besoins de la recherche d'itinéraire sécuritaire. En raison de l'accès au code source, CLIPS peut être compilé et intégré à un système en développement ou encore être portable sur différentes machines (Sun, PC, Macintosh, Risc, etc.).

Du point de vue architecture, CLIPS est en fait une librairie de fonctions avec laquelle l'utilisateur peut interagir de deux façons. Le mode commande exige de l'utilisateur qu'il tape les commandes de façon séquentielle, commandes qui seront interprétées par le système. Ces commandes peuvent être des appels aux fonctions, des règles, des variables ou des constantes. En mode traitement par lot (batch mode), un fichier texte contenant l'ensemble des instructions requises au fonctionnement du système est passé en paramètre lors de l'appel de CLIPS. Ce dernier retourne le résultat du traitement en paramètres après exécution.

En plus du SIG et du système expert, d'autres outils ont été utilisés pour ce projet. Mentionnons simplement le compilateur TURBO C++ qui a permis le développement des différentes routines essentielles au SSDRS et plus particulièrement de l'application "NodArc" dont nous discuterons ultérieurement.

8.3 Production des bases de données

Le but de l'étape de production des bases de données est de réaliser l'intégration des données de diverses sources et de s'assurer de l'intégrité topologique et sémantique des données qui serviront au SSDRS. Il s'agit d'une étape longue et fastidieuse qui recèle bon nombre d'embûches au niveau technique. Dans ce projet, l'utilisation de la carte de base au 1:50 000 plutôt que le fichier de réseau de rues de Statistique Canada a eu pour effet d'alourdir considérablement le travail d'édition. On peut diviser en trois phases les tâches reliées à la production des bases de données. Dans un premier temps, il est nécessaire de traiter les données cartographiques, dans un deuxième temps on doit rattacher aux entités cartographiques les attributs appropriés et, dans un troisième temps, élaborer le contenu de la base de connaissances.

8.3.1 Structuration des données cartographiques

Bien que la carte topographique à l'échelle du 1:50 000 représente une source d'information considérable à l'échelle canadienne, elle présente plusieurs problèmes au niveau topologique. La correction des ces problèmes ainsi que la structuration du fichier cartographique de Statistique Canada a nécessité de nombreuses étapes illustrées à la figure 8.7

Une fois les fichiers numériques en mains, la première étape a été la conversion des fichiers du format SIF (Standard Interchange Format) en un format lisible par Arc/Info-PC. Arc/Info-PC ne possédant pas d'utilitaire de lecture du format SIF, nous avons dû développer une routine en langage "C" permettant d'extraire par couche les éléments du fichier originel puis sauvegarder le résultat en format UNGEN d'Arc/Info-PC. De cette façon, toutes les entités de la couche #6

(réseau routier) du fichier SIF ont été transférées vers Arc/Info avec leur code d'identification respectif.

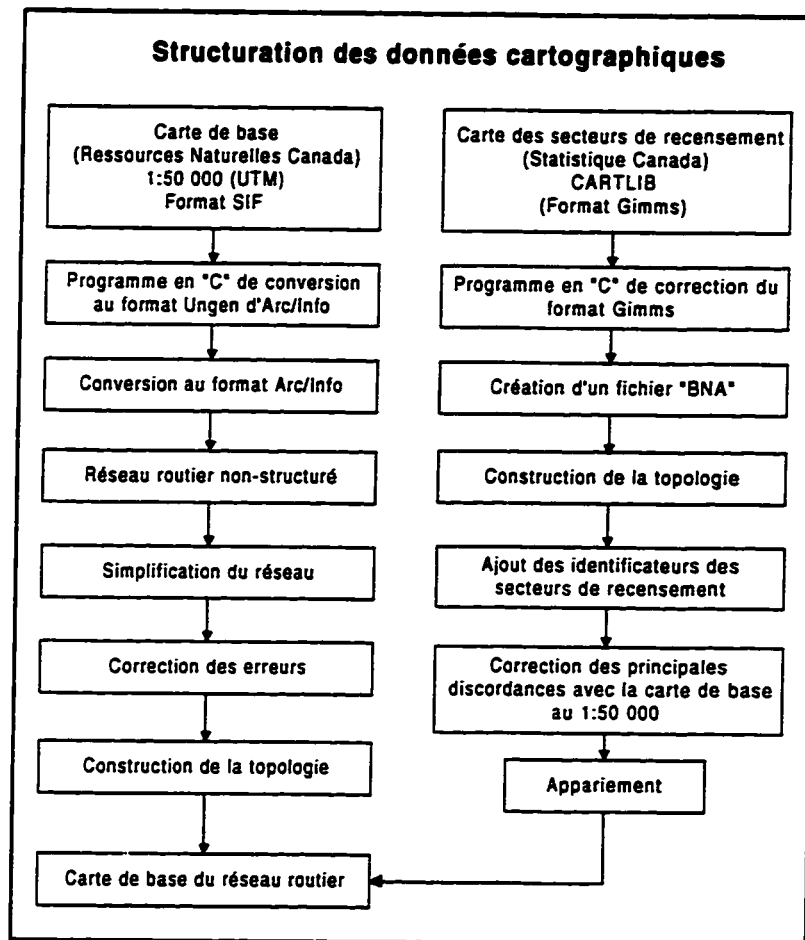


Figure 8.7: Structuration des données cartographiques

Cette première phase complétée, nous avons en mains un réseau routier non structuré composé de 55 967 segments, 10418 arcs pour un total de 4853 kilomètres d'éléments linéaires sans aucune structure topologique (figure 8.5). Un bref examen visuel du réseau nous a permis de relever un très grand nombre d'erreurs de saisie dues au balayage optique (scanner). Il nous est donc apparu évident que la somme de travail nécessaire pour corriger l'ensemble du fichier (centaines d'heures) dépassait largement le cadre de ce projet. Nous

avons donc procédé à une simplification du réseau routier avant de passer à l'étape de la correction des erreurs. Cette simplification a consisté à ne retenir que les routes de niveau supérieur (autoroutes, routes pavées à plus de 2 voies, routes pavées à 2 voies ou moins) en excluant les rues locales. Cette opération de filtrage s'est effectuée sur le code d'identification des tronçons (code à 4 chiffres, de 2010 à 2080) et le résultat est présenté à la figure 8.8. Ce fichier renferme 5401 arcs, 3867 noeuds pour un total de 2090 kilomètres de routes.

C'est à partir du réseau simplifié que le processus de correction des erreurs a débuté. On peut regrouper en deux grandes classes les erreurs ayant nécessité des corrections. Tout d'abord, on retrouve les erreurs liées au mode de saisie des données ayant entraîné un captage incomplet ou erroné du réseau. Parmi les cas les plus flagrants, on note les échangeurs autoroutiers, les arcs trop longs, les arcs trop courts ou les routes incomplètes (figure 8.9).

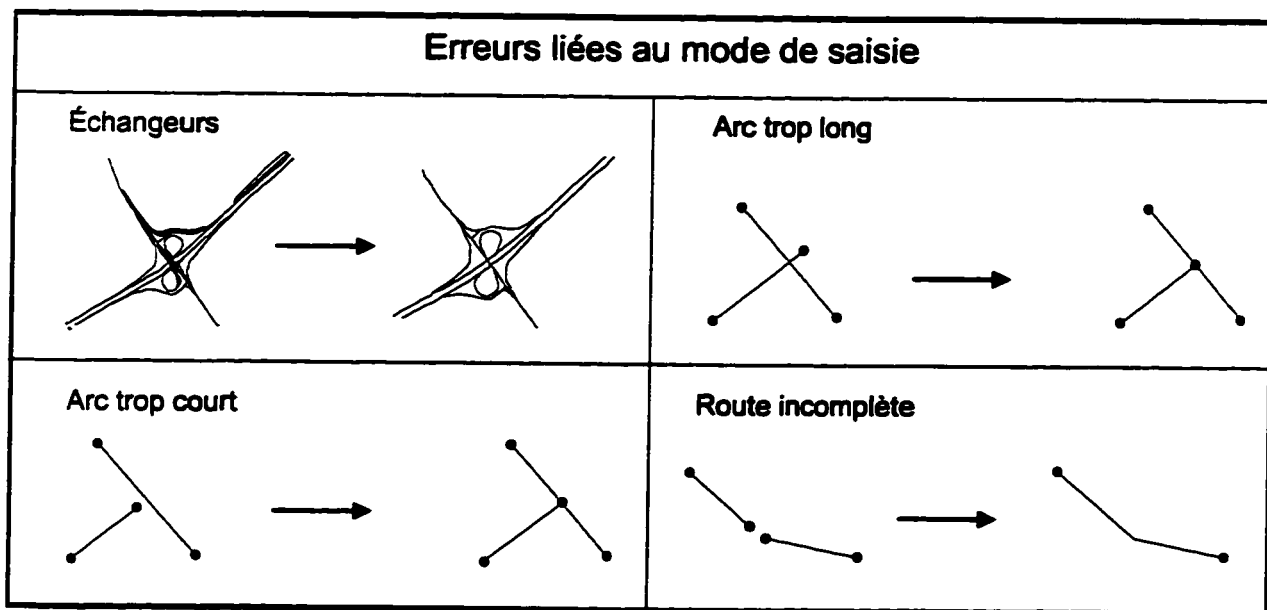


Figure 8.9: Erreurs liées au mode de saisie

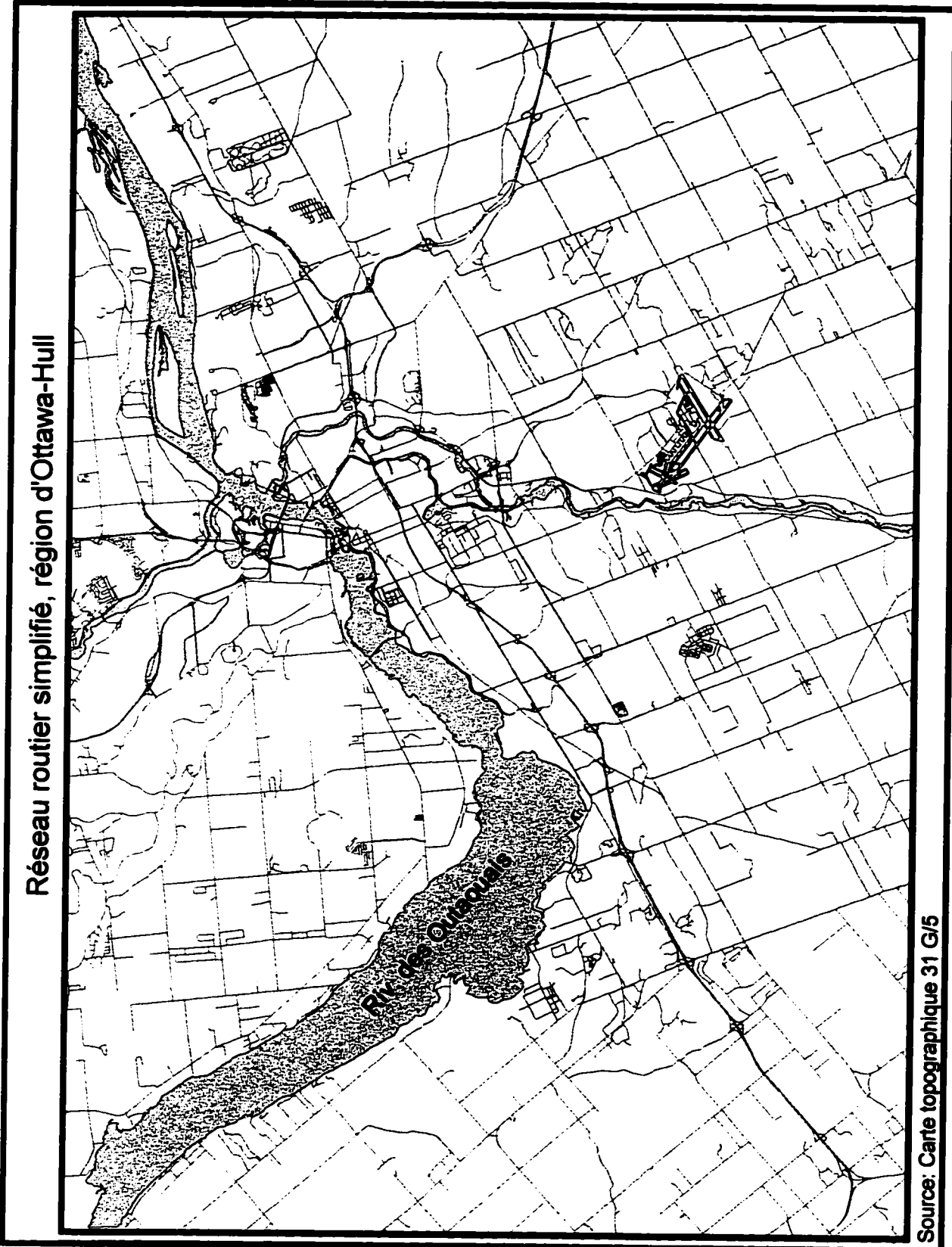


Figure 8.8: Réseau routier simplifié, région d'Ottawa-Hull

Les échangeurs routiers ont été nettoyés manuellement à l'aide des outils d'édition du module ARCEDIT d'Arc/Info, la carte papier au 1:50 000 faisant office de document de référence. Cette étape est difficilement automatisable puisque le jugement humain et la référence au document original est nécessaire pour produire un réseau conforme à la réalité. Par contre, la correction des arcs trop longs, trop courts ou des routes incomplètes a pu bénéficier d'une approche semi-automatisée à l'aide de la fonction CLEAN d'Arc/Info. Cette fonction permet, en tenant compte d'un seuil de tolérance défini par l'utilisateur, de tronçonner les arcs trop longs ou d'allonger les arcs trop courts. Le résultat de cette opération doit cependant être contre-vérifié manuellement par l'opérateur car, dans certaines situations, des raccordements erronés sont créés lors du traitement.

Le deuxième groupe d'erreurs fait référence à l'intégrité topologique du réseau routier. Il peut s'agir de pseudo-nœuds, d'intersections manquantes ou erronées ou encore d'une mauvaise orientation de la numérisation (figure 8.10).

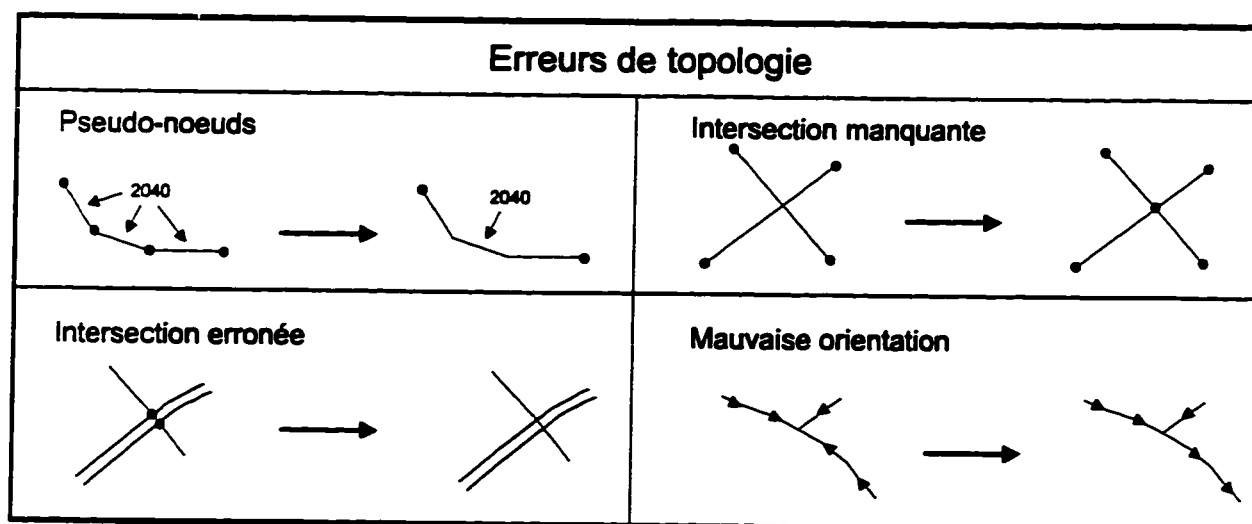


Figure 8.10: Erreurs de topologie

Dans le cas des pseudo-noeuds, la fonction DISSOLVE d'Arc/Info a permis de traiter de façon expéditive le problème. En spécifiant le code d'identification numérique comme critère de dissolution, tous les arcs ayant un code identique et formés de pseudo-noeuds ont été fusionnés. Par exemple, tous les arcs consécutifs sans intersection et ayant le même code "2040" sont fusionnés en un seul arc au code "2040" (figure 8.10). Par contre, les intersections manquantes ou erronées n'ont pu être traitées de façon automatisée puisque chaque cas représente une situation particulière exigeant le jugement humain. Les cas les plus communs d'intersections erronées se sont produits à la jonction du réseau d'autoroute où la présence de viaducs nous place en situation de graphe non-planaire. Ces fausses intersections ont été éliminées à l'aide de la fonction UNSPLIT d'Arc/Info.

Enfin, une erreur topologique peu apparente mais ayant un impact considérable sur la recherche d'itinéraire concerne la mauvaise orientation des arcs à sens unique. Par convention, l'orientation de la numérisation, c'est-à-dire la séquence des paires de coordonnées X,Y, détermine l'orientation du flux de circulation en présence de sens unique. La saisie par balayage optique a eu pour effet de créer un grand nombre de situations conflictuelles dans le réseau, conflits résultant généralement en deux tronçons parallèles dans la même direction (autoroutes) ou encore par des échangeurs orientés en direction inverse du trafic. La fonction FLIP d'Arc/Info a permis de modifier l'orientation des arcs affectés.

Outre le traitement du fichier du réseau routier, le fichier des secteurs de recensement a aussi été traité (figure 8.7). C'est à partir du fichier de la librairie CARTLIB de Statistique Canada que nous avons procédé. Une routine en langage "C" de lecture et de conversion du format GIMMS

au format "BNA" a été réalisée pour faciliter le transfert. La construction de la topologie (création de polygones) et l'ajout des identificateurs des secteurs de recensement s'est faite à l'aide d'Arc/Info. Enfin, le module ARCEDIT d'Arc/Info a été mis à profit pour la correction des principales discordances entre les limites des secteurs de recensement et le réseau routier au 1:50 000 (figure 8.11).

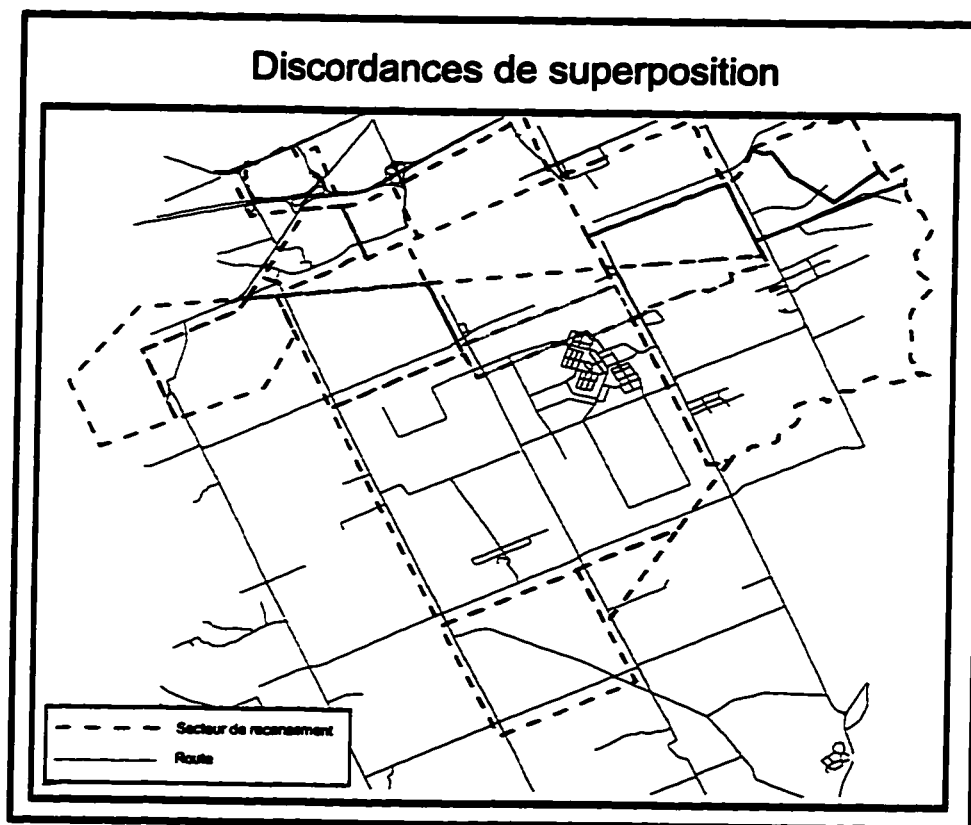


Figure 8.11: Discordances de superposition

Ces discordances proviennent d'une part de l'échelle de travail mais surtout de l'erreur de positionnement inhérente aux fichiers CARTLIB (plus de 300 mètres d'erreur) qui empêche une superposition parfaite des deux cartes². Il est à noter que les limites de la RMR d'Ottawa-Hull

²La superposition des deux cartes est nécessaire pour permettre d'associer les données sur la population et le logement aux tronçons routiers de chaque secteur de recensement.

dépassent largement celles de la carte topographique au 1:50 000 (figure 8.12). En pratique, seulement la partie centrale de la RMR a été utilisée lors de la superposition des cartes.

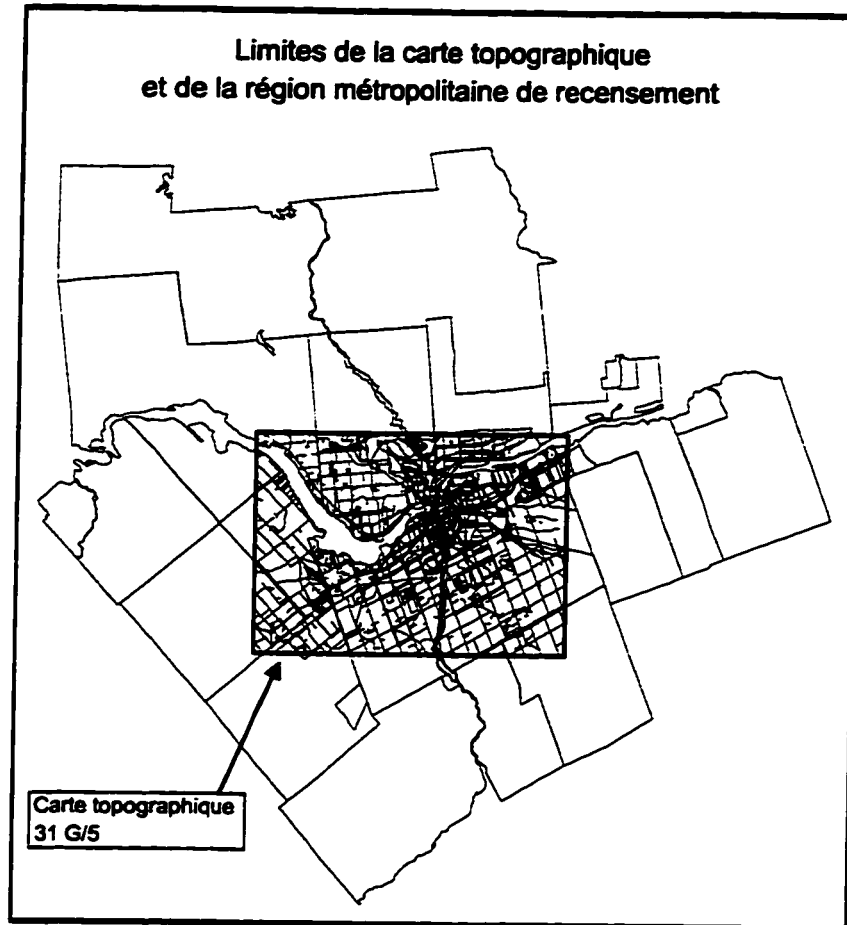


Figure 8.12: Limites de la RMR et de la carte topographique

8.3.2 Intégration des données d'attribut

En ce qui a trait aux données d'attribut, l'intégration a consisté à rattacher aux tronçons routiers les attributs de trois grands groupes principaux (figure 8.13). Les principaux sites vulnérables, à savoir les hôpitaux, les écoles, les parcs, les musées, les cinémas, les centres commerciaux, les édifices publics ont été localisés à l'aide de diverses cartes de la région d'Ottawa-Hull ainsi

que du logiciel *PC-Pathfinder*³. Après localisation visuelle, les sites vulnérables ont été rattachés manuellement à leur tronçon respectif à l'aide de la souris (Arc/Info). Un même site vulnérable pouvait être rattaché à plus d'un tronçon routier selon sa position dans la configuration du réseau (figure 8.14).

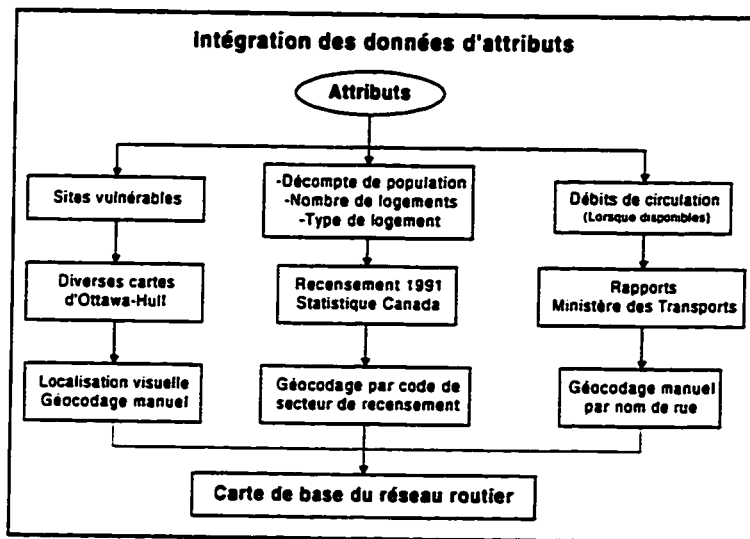


Figure 8.13: Intégration des données d'attributs

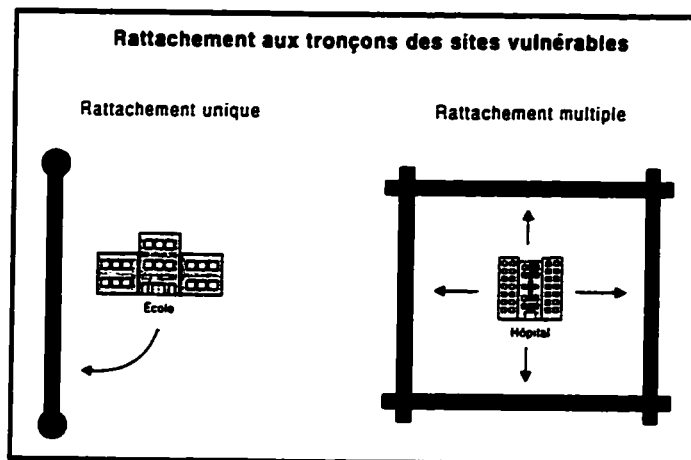


Figure 8.14: Rattachement des attributs

³Contrairement à ce que laisse présager son nom, le logiciel *PC-Pathfinder* n'offre aucune fonction de recherche d'itinéraire. Cependant, il renferme pour la région d'Ottawa-Hull la localisation d'un grand nombre de sites d'intérêt général ou touristique.

Les données relatives à la population et au logement ont été recueillies à partir des codes de secteurs de recensement de Statistique Canada. Du recensement de 1991, nous avons retenu le décompte de population pour en dériver la densité de population au kilomètre carré (figure 8.15), le nombre total de logement pour évaluer la population résidente (figure 8.16) et le type de logement (unifamiliale, tour d'habitation de plus de 5 étages) pour juger du type de bâti (figure 8.17).

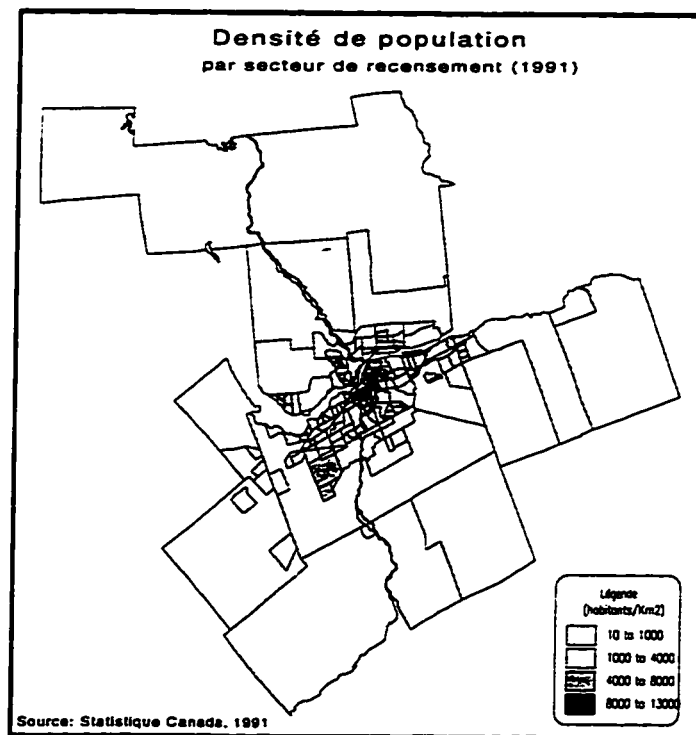


Figure 8.15: Densité de population, RMR d'Ottawa-Hull

Ces données ont été rattachées aux tronçons routiers à l'aide du module OVERLAY d'Arc/Info. Chaque tronçon situé dans un secteur de recensement donné se voyait attribuer les caractéristiques de ce secteur. Quant aux données sur les débits de circulation, leur disponibilité est limitée à certains points névralgiques du réseau, selon les relevés du Ministère des Transports. Comme il s'agit d'un débit de véhicule calculé d'intersection à

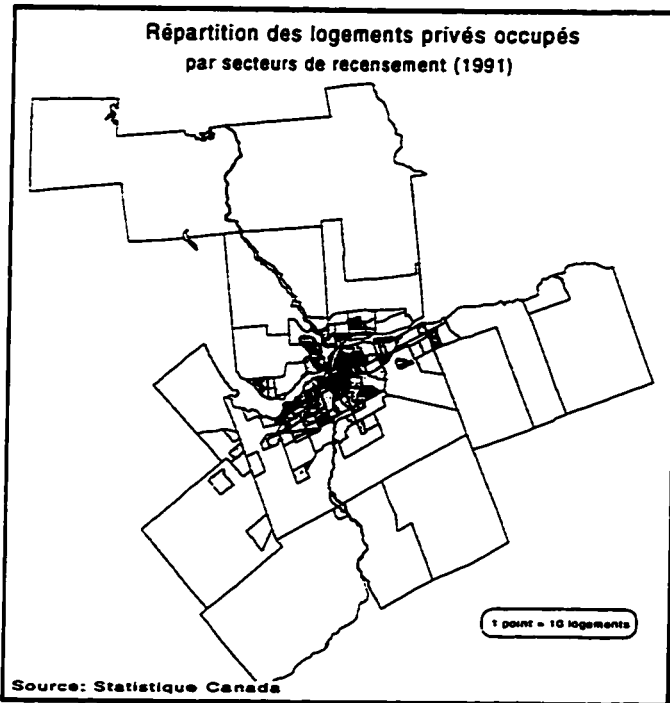


Figure 8.16: Répartition des logements privés, RMR d'Ottawa-Hull

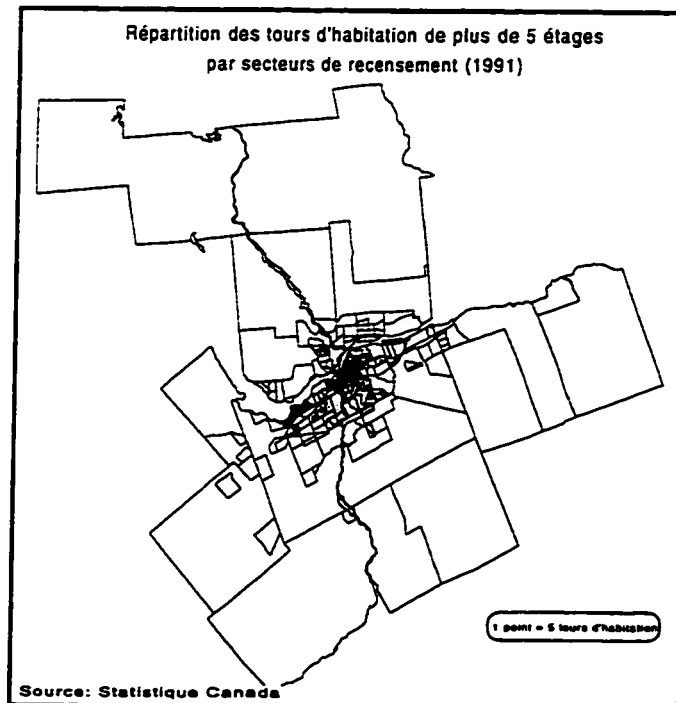


Figure 8.17: Répartition des tours d'habitation de + de 5 étages

intersection, il est relativement aisé (bien que fastidieux!) de géocoder cette information. Seule une bonne carte routière est nécessaire.

8.3.3 Élaboration de la base de connaissances

Comme il a été mentionné au chapitre précédent (section 7.4), la structuration de la base de connaissances revêt un caractère important dans un SSDRS. Bien que chaque fait et chaque règle de production soit un “morceau de connaissances” indépendant, il est préférable de déterminer au préalable, sous forme d’organigramme, la logique sous-jacente à la base de connaissances. La figure 8.18 synthétise le schéma déductif général des règles de production. L’objectif ultime est évidemment d’en arriver à déterminer, pour chaque tronçon, quel risque il comporte (classe de préférence, section 7.2) pour le transport d’une matière donnée, à un moment donné.

L’aspect temporel est pris en compte en considérant simplement si le transport aura lieu de jour ou de nuit (figure 8.18). Pour simplifier la base de connaissances, nous avons regroupé, selon la méthode de l’OCDE, les matières dangereuses selon leurs effets les plus probables en cas d’accident (explosion, gaz toxiques, radioactivité, autres). Les aspects légaux ne sont considérés qu’en cas d’interdiction formelle. L’interdiction peut faire référence spécifiquement à la matière dangereuse (les explosifs sont interdits dans les tunnels) ou encore à une interdiction de circuler pour les camions (qu’ils transportent ou non des matières dangereuses). A ce stade de la démarche déductive, un tronçon peut être rejeté pour des raisons strictement légales.

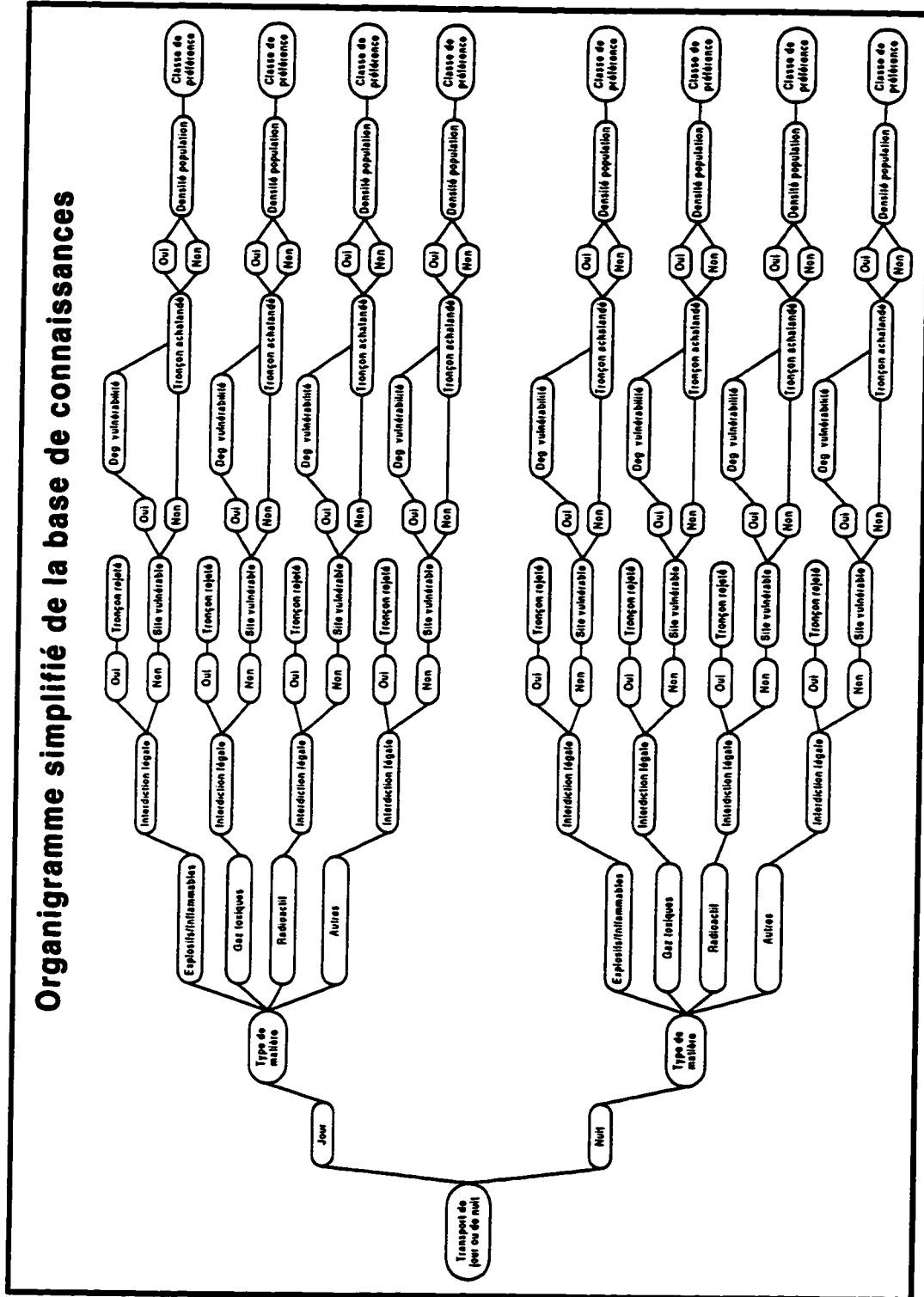


Figure 8.18: Organigramme simplifié de la base de connaissances

Si le tronçon n'est pas frappé d'interdiction de circuler, la démarche se poursuit avec l'évaluation des sites vulnérables le long du tronçon (figure 8.18). Si le tronçon possède un site vulnérable, son degré de vulnérabilité sera évalué en fonction de la matière transportée et de la période de la journée. Un site à risque très élevé de jour peut devenir à risque faible la nuit (exemple: musée). De même, un site tel qu'un plan d'eau sera à risque faible pour les gaz toxiques, mais à risque élevé pour les inflammables (essence). Si les données sur le débit de circulation pour le tronçon sont disponibles, l'achalandage est évalué. Tout tronçon ayant un débit de circulation au-dessus de la moyenne de débit est considéré comme achalandé. Enfin, les densités de population et de logements complètent la démarche de déduction.⁴

Pour traduire ce processus déductif de façon à ce que le logiciel CLIPS puisse l'appliquer au réseau routier, il faut le coder selon une syntaxe particulière. Pour y parvenir, quatre aspects sont importants: le codage des faits, le codage des règles de production, la salience et le codage de l'incertitude. CLIPS fait une distinction entre les faits et les règles de production. Les faits sont associés aux données ou aux informations sur lesquelles le système pourra raisonner. Chaque fait est composé d'un identificateur symbolique suivi d'un ou plusieurs items (slots) décrivant la connaissance:

(musee (site "musee")
 (risque "moyen")
 (période "jour")
 (matiere "inflammables")
 (certitude 0.8))

⁴Les données sur la population et le logement se sont avérées moins significatives que prévu dans ce contexte puisque tous les tronçons à l'intérieur d'un même secteur de recensement avaient les mêmes valeurs de densité, diminuant ainsi de beaucoup le pouvoir discriminant de cette variable.

Dans cet exemple, "musee" est l'identificateur symbolique du fait et les items "site", "risque", "periode", "matiere" et "certitude" renferment des connaissances sur le risque associé au transport d'une matière inflammable, de jour, à proximité d'un musée. A l'aide de cette syntaxe, on peut définir toutes les connaissances nécessaires au raisonnement en les divisant en morceaux indépendants les uns des autres.

```
(musee      (site "musee")
            (risque "faible")
            (periode "nuit")
            (matiere "inflammable")
            (certitude 0.8))
(musee      (site "musee")
            (risque "sans risque")
            (periode "nuit")
            (matiere "autres")
            (certitude 0.65))
(plan_eau   (site "plan d'eau")
            (risque "sans risque")
            (periode "jour")
            (matiere "gaz toxique")
            (certitude 0.9))
etc...
```

L'ensemble des faits forme la *liste des faits*, liste qui peut être stockée dans un fichier et chargée en mémoire par le système expert au début du processus d'inférence. CLIPS possède également des outils d'édition de la liste de faits permettant d'ajouter, de retrancher ou de modifier des faits au besoin. En modifiant ce fichier, l'opérateur est en mesure de simuler différents scénarios et différentes situations sans avoir à modifier tout le SSDRS.

Les faits ne sont que des connaissances passives. Pour raisonner, le système expert doit aussi renfermer des règles de production qui lui indiqueront comment réagir dans une situation donnée. Tout comme les faits, les règles de production sont définies à l'aide d'une syntaxe

propre à CLIPS. La partie gauche de la règle (prémisse) décrit une situation et la partie droite, l'action à entreprendre dans cette circonstance. Dans CLIPS, le symbole "=>" sépare la prémisse de son action. Chaque règle est également repérable à l'aide d'un identificateur symbolique unique précédant la prémisse.

```
(defrule site_vulnérable_hopital " "
  (declare (salience 80))
  (site hopital)
=>
  (assert (risque tres eleve)))

(defrule interdiction_legale " "
  (declare (salience 100))
  (matiere explosif_inflammables)
  (interdiction tunnel)
=>
  (assert (troncon rejete)))

(defrule achalandage " "
  (declare (salience 50))
  (donnee vrai)
=>
  if (debit > moy_debit)
  then .
    (assert (troncon achalande))
    (assert (risque eleve))
  else
    (assert (troncon peu achalande))
    (assert(risque faible)))
```

Par exemple, la règle de production "site_vulnérable_hopital" indique au système expert qu'un tronçon routier avec un hôpital est à risque très élevé en tout temps. Un seul critère est nécessaire dans la prémisse (site hopital) pour déclarer le tronçon à risque très élevé (assert (risque tres eleve)). Le deuxième exemple de règle (interdiction_legale) possède une prémisse à deux critères, soit la présence d'explosifs/inflammables et d'un tunnel. Quant à la "salience" (declare (salience 100)), elle est un paramètre optionnel des règles de production. Sa valeur

peut varier de -10 000 à 10 000 et elle permet de fixer l'ordre de priorité dans lequel les règles seront exécutées. Dans l'exemple ci-dessus, la règle "interdiction_legale" sera exécutée la première puisqu'elle possède la salience la plus élevée. L'avantage principal d'utiliser la salience est qu'elle permet d'orienter le raisonnement du système tout en évitant d'avoir à se préoccuper de l'ordre des règles de production dans la base de connaissances.

De plus, la définition des règles de production dans CLIPS peut s'adapter à des situations complexes. Par exemple, la règle "achalandage" détermine le risque d'un tronçon si des données sont disponibles (donnee vrai) et selon les critères IF-THEN-ELSE. Plusieurs branchements conditionnels peuvent être imbriqués selon les besoins. CLIPS possède d'ailleurs un langage complet permettant de traiter les données symboliques (qualitatives) et quantitatives. Quant à l'incertitude, nous l'avons introduit de façon explicite dans les faits de la base de connaissances. Elle est considérée au même titre que les autres informations de la base.

Tout comme en programmation conventionnelle, le développement de la base de connaissances d'un système expert doit être suivi de tests pour s'assurer du bon fonctionnement du système. Contrairement à la programmation conventionnelle cependant, le seul moyen de contrôler la justesse des conclusions générées par le système est de le confronter à des experts humains. Faire tourner le système le plus grand nombre de fois

possible dans des situations différentes est le seul moyen de s'assurer de la valeur d'une base de connaissances.⁵

8.4 Le système de support à la décision

A l'origine, le système de support à la décision devait être développé à l'intérieur même du SIG Arc/Info à l'aide du langage SML. Cependant, plusieurs tests à partir du logiciel nous ont convaincus des limites de cette approche. Dans sa version IBM-PC, Arc/Info possède des capacités limitées de gestion de la mémoire, incompatible avec le fonctionnement intégré du système expert CLIPS. Contrairement à la version Unix d'Arc/Info, quasi illimitée en gestion mémoire, la version PC ne permet que le développement de petites applications externes⁶. Pour vérifier nos propos concernant la recherche d'itinéraires et les systèmes experts, malgré ce problème purement technique, nous avons développé plusieurs routines en langage "C" qui ont formé la base du logiciel NodArc, notre "mini SSDRS".

8.4.1 NodArc

Bien que le développement de NodArc ait nécessité plusieurs milliers de lignes de programmation, il ne possède aucune fonction d'édition de données. Il permet toutefois la conversion et l'intégration des fichiers Arc/Info (Ungen), la création des fichiers topologiques et d'attributs nécessaires à la recherche d'itinéraires, l'affichage du réseau routier et de l'itinéraire

⁵Dans les systèmes à plusieurs centaines ou milliers de règles, la confiance absolue dans le système demeure un idéal inatteignable. Cette remarque est toutefois valable pour les logiciels conventionnels qui renferment tous des "bugs" cachés.

⁶Cette limite ne fait pas référence à la taille des fichiers mais plutôt à la taille des programmes externes pouvant être exécutés à l'intérieur de l'interpréteur de commande (Arc/Info prompt) d'Arc/Info.

calculé et assure également le lien avec le système expert CLIPS. NodArc possède aussi plusieurs autres fonctions telle que le démontrent les saisies d'écran des figures 8.19 à 8.25.

L'écran de travail de NodArc se divise en trois fenêtres (figure 8.19). Une fenêtre principale permet l'affichage du réseau routier et de l'itinéraire calculé; une fenêtre secondaire (File Informations) renferme les informations sur les fichiers en traitement et la fenêtre "System Informations" informe l'utilisateur de la durée de l'opération réalisée. Les menus "NodArc" et "System" comprennent les diverses fonctions de configuration du logiciel et d'affichage des fichiers de résultats (figure 8.20 et 8.21)

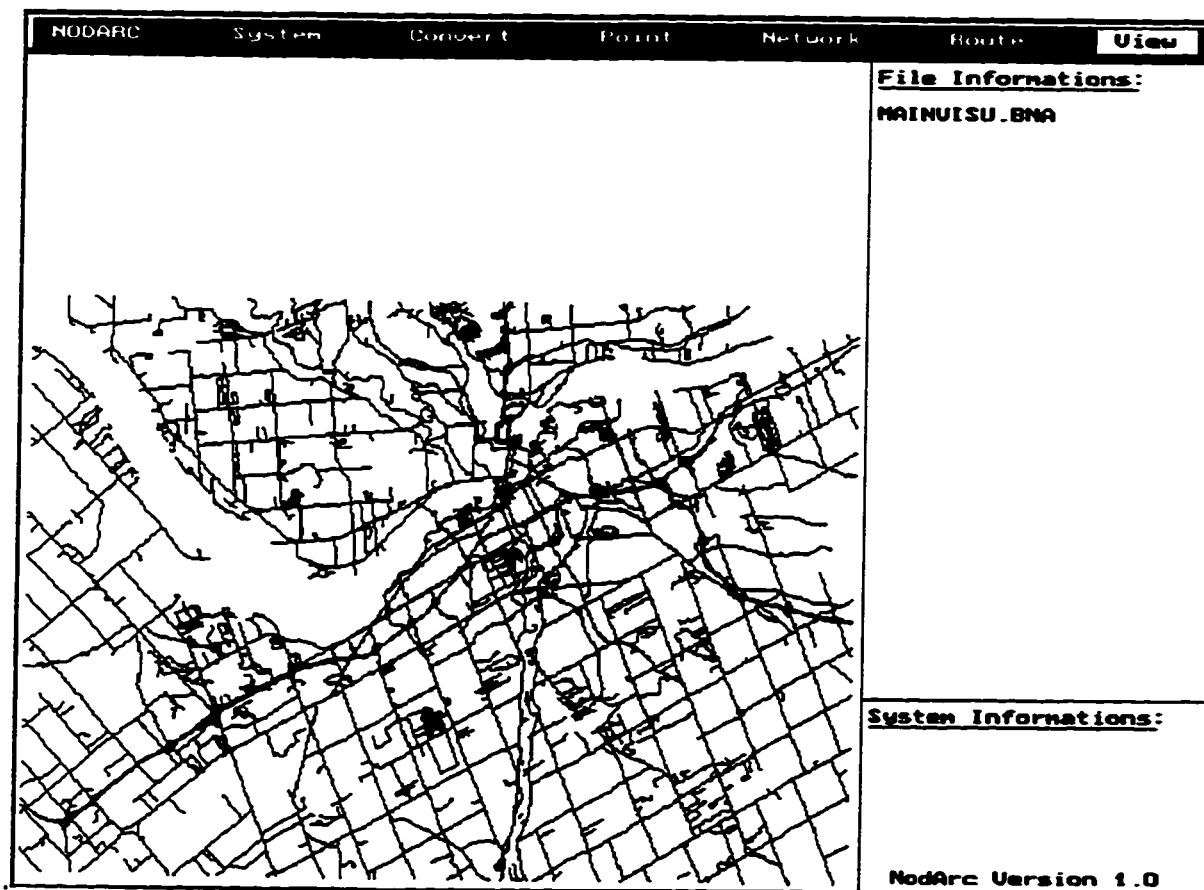


Figure 8.19: NodArc, fenêtres d'affichage

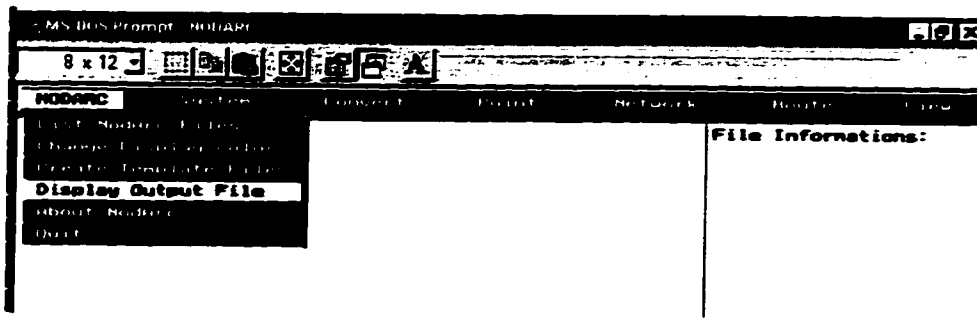


Figure 8.20: NodArc, menu NODARC

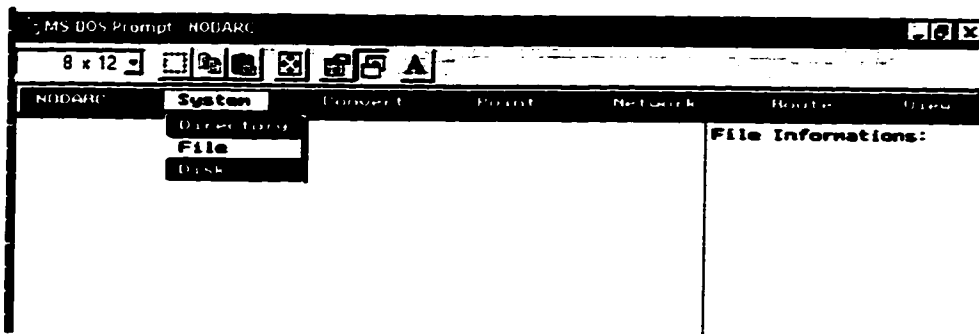


Figure 8.21: NodArc, menu SYSTEM

Les fonctions d'intégration de fichiers et de topologie se retrouvent dans les menus "Convert" et "Network" (figure 8.22 et 8.23), tandis que le menu "View" renferme les options d'affichage des fichiers cartographiques et du réseau routier (figure 8.24).

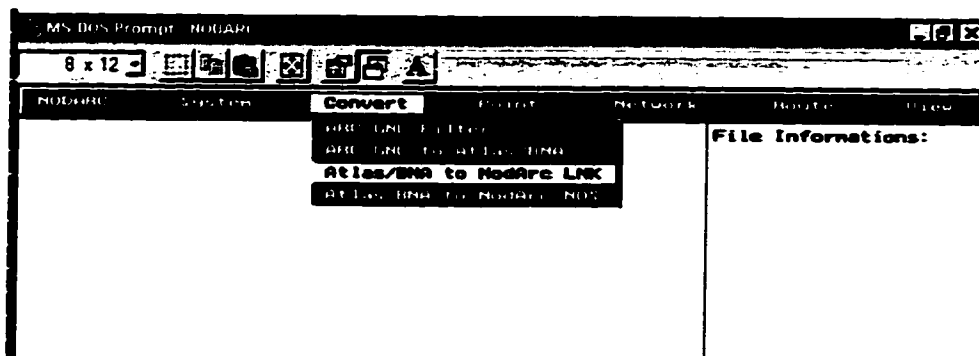


Figure 8.22: NodArc, menu CONVERT

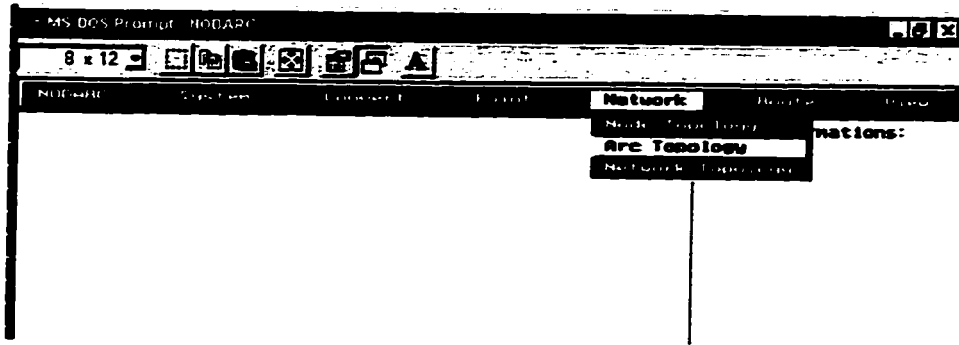


Figure 8.23: NodArc, menu NETWORK

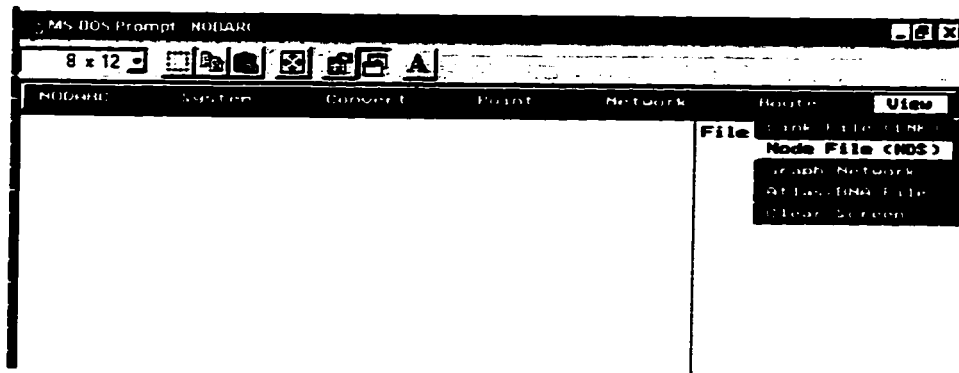


Figure 8.24: NodArc, menu VIEW

Quant au menu "Route", il permet le calcul d'itinéraire à partir d'un algorithme du plus court chemin, le calcul de K-plus courts chemins et le lien avec le système expert avec l'option "Expert Path" (figure 8.25).

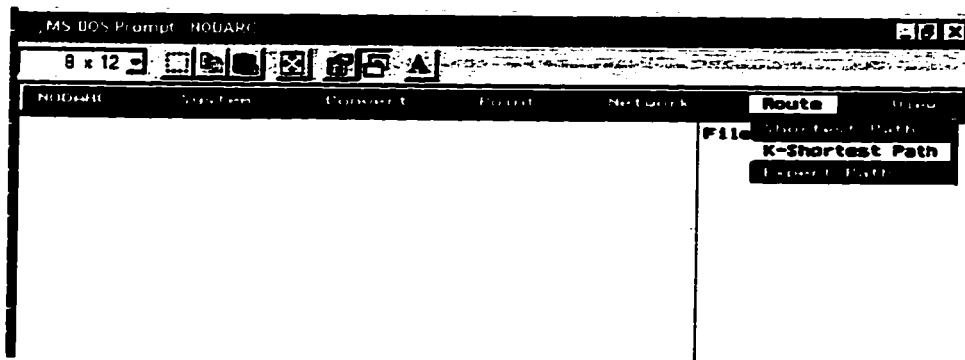


Figure 8.25: NodArc, menu ROUTE

8.4.2 Structure de données de NodArc

Les bases de données ayant été structurées préalablement à l'aide d'Arc/Info, nous avons choisi de développer dans NodArc des fonctions d'intégration des fichiers provenant d'Arc/Info. L'annexe 1 contient le code source des principales fonctions de conversion et de création des fichiers NodArc tandis que la figure 8.26 illustre comment ce dernier intègre le réseau routier construit dans Arc/Info pour le convertir en son propre format pour le calcul d'itinéraires (figure 8.26).

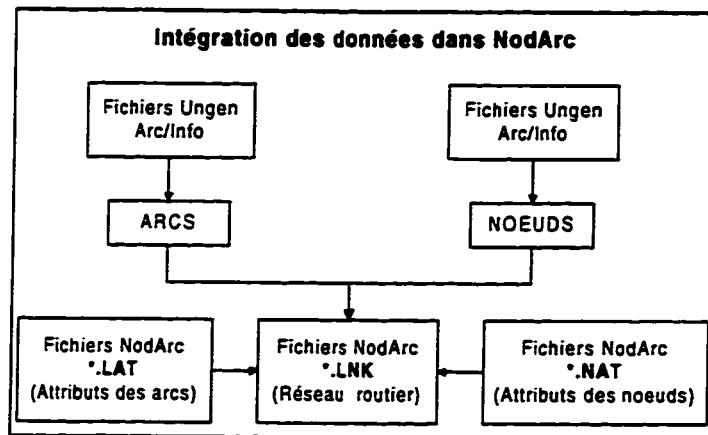


Figure 8.26: Intégration des données dans NodArc

C'est à partir d'un fichier d'arcs et d'un fichier de noeuds que NodArc crée un fichier *.LNK (Link file) renfermant la topologie du réseau routier. À ce fichier *.LNK sont associés deux fichiers, un fichier *.LAT (Link Attribute Table) contenant les attributs des arcs et, un fichier *.NAT (Node Attribute Table) renfermant les attributs des noeuds du réseau. Le lien entre ces fichiers est assuré à partir d'un identificateur unique, généré par NodArc au moment de la conversion, pour chaque élément du réseau routier (arcs et noeuds). Toutes les informations nécessaires au calcul d'itinéraire se retrouvent donc dans ces trois fichiers.

8.4.3 Algorithme du plus court chemin

C'est à partir de la formalisation du modèle présentée au chapitre 7 (section 7.2) que nous avons programmé l'algorithme de recherche d'itinéraire de NodArc. Bien que l'algorithme de Dijkstra ait été retenu dès le départ, nous avons procédé à quelques expérimentations sur des variantes de l'algorithme de Dijkstra ainsi que sur d'autres algorithmes, comme celui de la cascade et de "K" plus courts chemins. Présenté sous forme de pseudo-code, l'algorithme de Dijkstra prend l'allure suivante:

Considérons un réseau de "n" noeuds numérotés de 0 à "n-1"

A partir:

- D'une liste d'adjacence "c" où $c[i,j]$ nous donne la distance réelle à parcourir du noeud "i" au noeud "j" (ou encore ∞ si aucun arc)
- D'une liste de préférence $lp[i,j]$ nous donnant la classe de préférence provenant du système expert pour l'arc reliant les noeuds "i" et "j".
- D'un vecteur $d[i]$ stockant la distance connue la plus courte du noeud 0 (départ) au noeud "i".
- D'un vecteur de préférence $p[i]$ stockant la somme des préférences du noeud 0 au noeud "i"
- D'un vecteur $v[i]$ stockant la liste de noeuds ayant été visités
- D'un vecteur $r[i]$ stockant le noeud précédant "i" dans le chemin le plus court.

Pour trouver le chemin le plus sécuritaire du noeud 0 au noeud "m":

algorithme dijkstra (c)

$v[0] = \text{visité}$

Incrémenter i de 1 à n

$v[i] = \text{non visité}$

$r[0] = 0$

Pour chaque i où $c[0,i] \neq \infty$

$r[i] = 0$

```

Incrémenter i de 0 à n
  d[i] = c[0,i]
  p[i] = Appel du système expert (lp[0,i])
  jusqu'à ce que le noeud "m" soit visité
Trouver les noeuds non visités i tel que p[i] est maximum
v[i] = visité
Pour chaque noeud non visité "j" connecté à "i"
  si p[i] + lp[i,j] > p[j]
    p[j] = p[i] + lp[i,j]
    d[j] = d[i] + c[i,j]
    r[j] = i

```

Évidemment, il s'agit ici d'une version fortement épurée de l'algorithme qui ne présente pas toutes les conditions devant être vérifiées pour la recherche d'itinéraire. L'annexe 2 contient le code source en langage "C" de l'algorithme de base développé pour ce projet.

Étant donné la lenteur considérable de traitement engendrée par l'appel au système expert, nous avons également procédé à une modification de l'algorithme. Dans la version modifiée (voir annexe 2), nous procédons au calcul d'itinéraire simultanément dans les deux directions à partir du noeud d'origine et du noeud de destination. Cette approche nous permet de réduire considérablement le nombre d'itérations puisque, dans les meilleures situations, les deux portions d'itinéraires se rejoignent au sommet central de l'itinéraire, réduisant ainsi de moitié le nombre d'itérations nécessaires pour parvenir à une solution. Il faut toutefois préciser que, même avec cet algorithme, l'ajout de la composante système expert alourdit notablement la recherche d'itinéraire⁷.

⁷Certaines versions d'algorithmes que nous avons testés exigeaient plusieurs heures de temps de calcul, même des journées entières dans le cas de l'algorithme de la cascade.

8.4.4 Recherche d'itinéraire

C'est à partir du menu ROUTE de NodArc que l'on peut effectuer une recherche d'itinéraire. Avant de débiter une recherche, NodArc doit connaître 5 paramètres qui sont demandés séquentiellement à l'utilisateur à partir de fenêtres de saisie (figure 8.27).

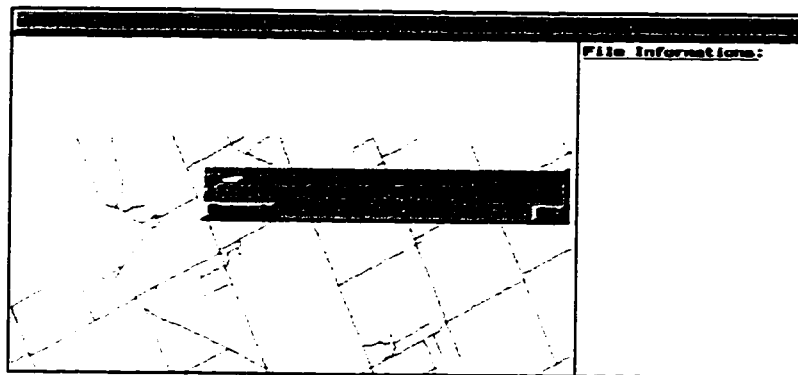


Figure 8.27: Fenêtre de saisie des paramètres

Ces paramètres sont le noeud de départ, le noeud d'arrivée, la classe de marchandises transportées, la période (jour/nuit) et le nom du fichier de sortie où sera stockée la table de routage. Dans sa version 1.0, NodArc ne calcule les itinéraires que de noeud à noeud dans le réseau. Un itinéraire doit donc débuter et se terminer à une intersection ou à une extrémité de tronçon. De plus, NodArc ne reconnaît que l'identificateur numérique pour localiser les noeuds. L'opérateur doit donc connaître l'identificateur numérique du noeud de départ et du noeud d'arrivée préalablement au calcul d'itinéraire.

Une fois les paramètres en mémoire, NodArc procède au calcul d'itinéraire. Les matrices d'adjacence sont créées puis balayées et les tronçons sont évalués. Le résultat final prend deux formes. Tout d'abord, une carte est affichée à l'écran. Celle-ci contient le réseau routier et l'itinéraire calculé qui apparaît dans une couleur distincte. De plus, un fichier de sortie, créé

lors du traitement, renferme des informations telles que la longueur du trajet retenu et la table de routage. Cette dernière est en fait la liste des noeuds (avec identificateur numérique) par lesquels passe l'itinéraire retenu. Dans une prochaine version de NodArc, cette table contiendra également l'explication du choix de l'itinéraire générée par le système expert.

Pour démontrer le comportement de NodArc en ce qui a trait à la recherche d'itinéraire, nous avons procédé à une simulation sur la partie sud-est de la région d'Ottawa-Hull (Gloucester). Le point de départ (noeud 10) se situe au nord-est sur *Base Line Road* (près de *Leitrim Road*) et le point d'arrivée sur *Limebank Road* au sud-ouest (noeud 200). Dans un premier temps, nous avons retiré de la base de données l'*Hippodrome Rideau* situé près de l'intersection des routes *Albion Road* et *Rideau Road*. La figure 8.28 illustre le trajet retenu sans tenir compte de la présence de l'hippodrome.

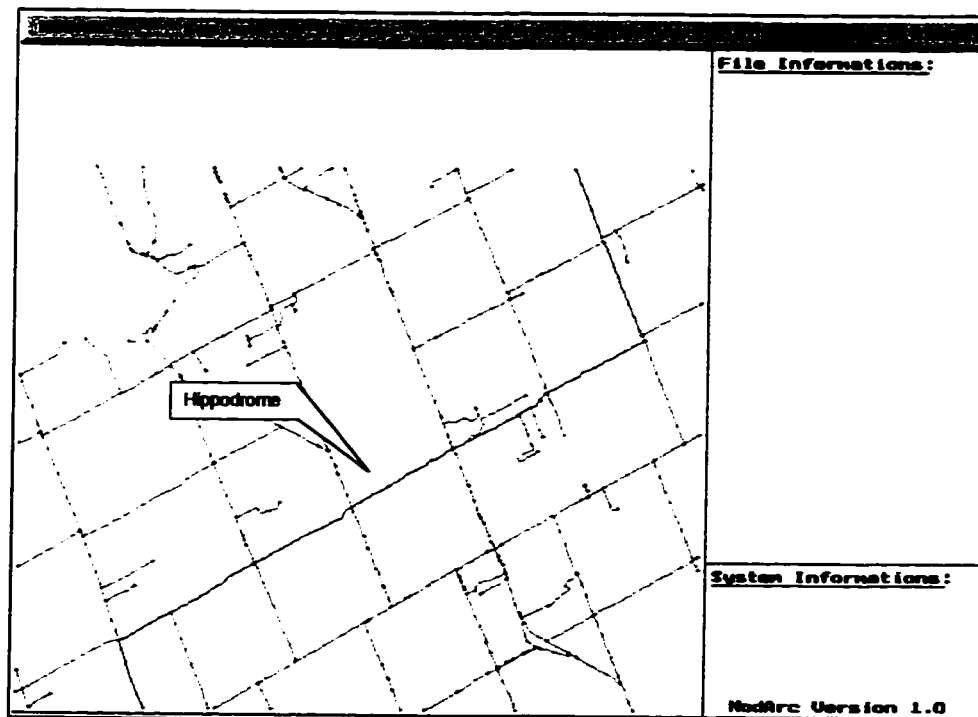


Figure 8.28: Itinéraire retenu sans tenir compte de l'hippodrome

Le trajet retenu emprunte le tronçon de route passant devant l'hippodrome (Rideau Road) puisqu'il permet de minimiser la distance à parcourir. Le fichier de sortie (figure 8.29) nous indique que la longueur totale du trajet est de 16.08 kilomètres (16086 mètres).

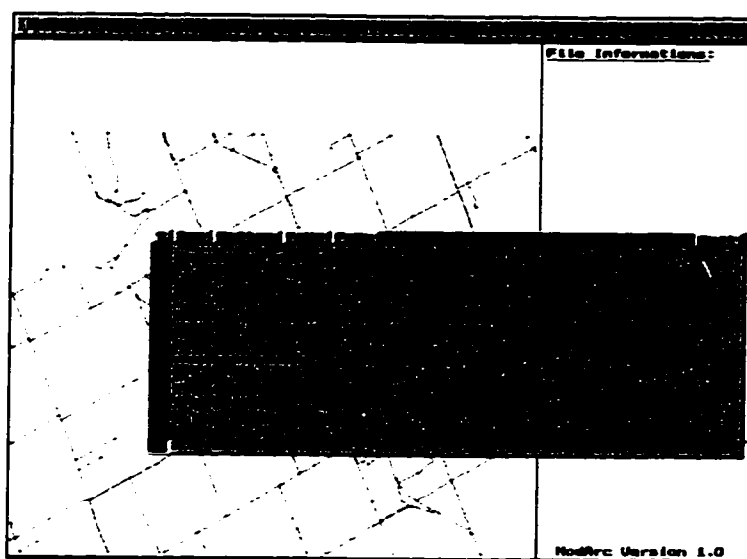


Figure 8.29: Fichier de sortie et table de routage

Par contre, si l'on reprend la même recherche d'itinéraire mais en indiquant dans la base de données la présence de l'hippodrome, on obtient un trajet différent (figure 8.30).

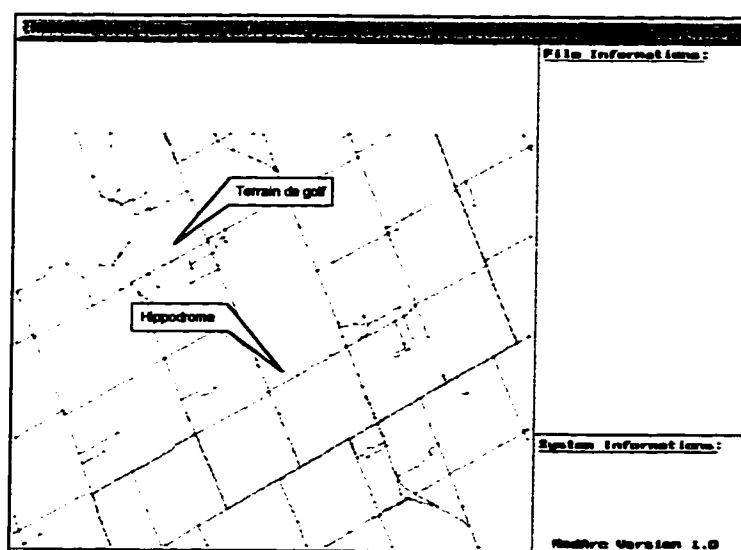


Figure 8.30: Trajet retenu en tenant compte de l'hippodrome

L'itinéraire contourne par le sud l'hippodrome, puis revient sur *Rideau Road* pour finalement se rendre à destination. Il est à noter que la raison pour laquelle l'itinéraire ne contourne pas par le nord l'hippodrome est la présence d'un terrain de golf (Club de golf d'Ottawa) au nord-ouest. Comme on pouvait s'y attendre, le deuxième itinéraire implique une distance à parcourir plus longue avec 20 kilomètres (figure 8.31).

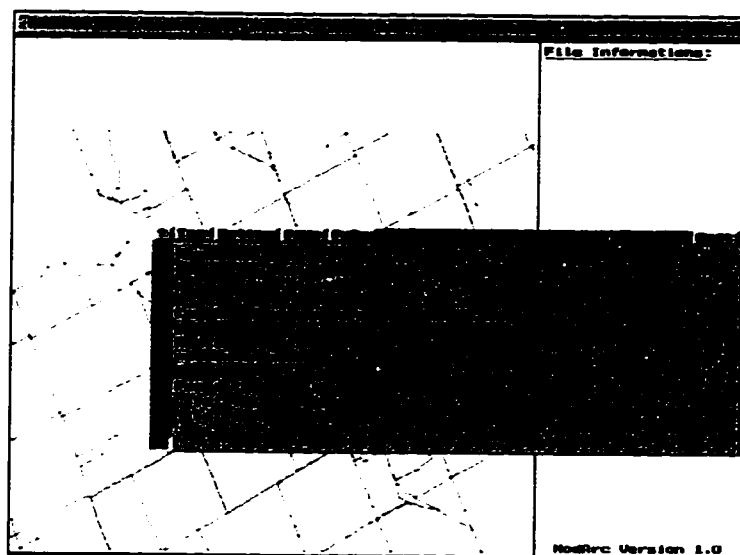


Figure 8.31: Fichier de sortie et table de routage

Certains secteurs, en milieu plus densément peuplé et aux utilisations du sol variées peuvent également influencer fortement les résultats. La figure 8.32 illustre une portion du réseau routier située au centre-ville d'Aylmer. On retrouve dans ce secteur le centre commercial *Les Galeries d'Aylmer*, le *Club de golf de Gatineau* et de nombreuses écoles (8 au total). La position de ces infrastructures a pour effet de rendre inaccessible le jour, par sélection à priori des tronçons sécuritaires, tout le territoire situé au sud des *Galeries d'Aylmer*. Selon le système expert, certains tronçons du *Chemin Vanier*, du *Boulevard Lucerne*, de l'*Avenue Frank-Robinson*, du *Chemin d'Aylmer* et du *Boulevard Wilfrid-Lavigne* sont jugés à risque. Par contre, la nuit, tout le secteur redevient accessible.

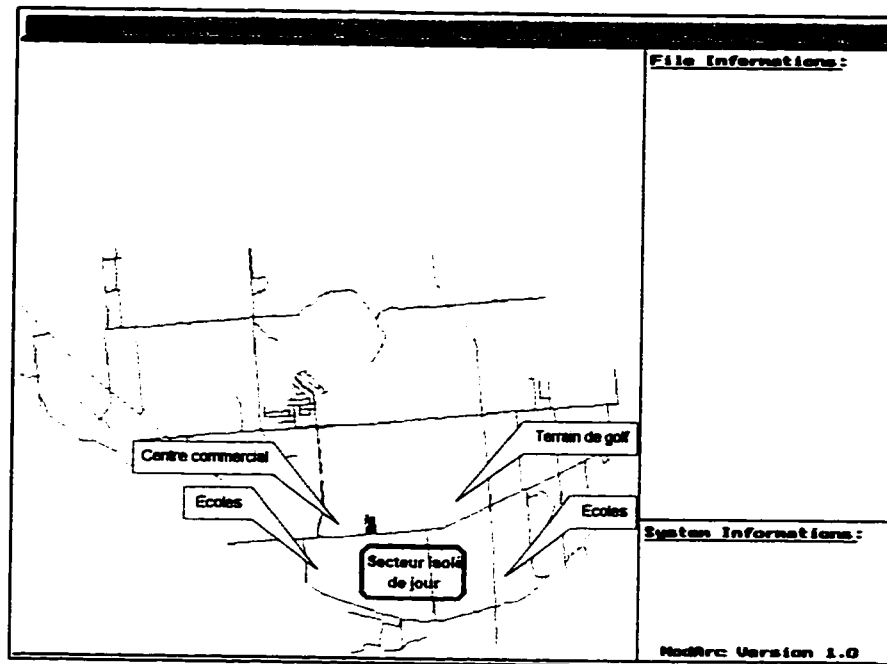


Figure 8.32: Secteur inaccessible le jour

8.5 Évaluation du système: les sources d'erreurs

En situation d'application réelle, la valeur d'un SSDRS dépend de son degré de fiabilité. Dans l'élaboration d'un tel système, des erreurs peuvent se glisser à toutes les étapes du projet. Dans l'élaboration des bases de données, des erreurs peuvent se produire lors de la saisie et de la structuration des données cartographiques et d'attributs. Les nombreuses opérations manuelles et répétitives sont sujettes aux erreurs de codage. A ce niveau, seul un examen minutieux des bases de données par plusieurs opérateurs peut offrir une garantie suffisante de fiabilité.

Plus subtiles et beaucoup plus difficiles à repérer sont les erreurs associées à la base de connaissances du système expert. Au-delà de la simple erreur de syntaxe qui sera détectée

automatiquement par l'interpréteur de commandes du système expert, des erreurs de consistance, de complétude et de logique peuvent réduire considérablement la fiabilité d'une approche par système expert. En situation réelle, la validation de la base de connaissances et des liens logiques entre les différents éléments de cette base se fera par un groupe d'experts humains dans le domaine d'application. La tâche de ces experts sera alors de confronter le système dans un éventail le plus vaste possible de situations usuelles et de "cas problèmes". Le degré de fiabilité d'un SSDRS sera d'autant plus élevé qu'il aura été scruté de façon exhaustive par un groupe d'experts aux compétences reconnues.

Évidemment, dans le cadre de cette thèse, nous n'avons pas eu accès à un groupe d'experts tant pour bâtir la base de connaissances que pour la tester. Nous ne pouvons donc pas prétendre que NodArc détermine de façon absolue les itinéraires les plus sécuritaires, mais plutôt qu'il opère selon les données et les connaissances liées à notre propre "compétence en la matière". Pour s'assurer de la vraisemblance des itinéraires calculés par NodArc, nous les avons confrontés à une carte du réseau routier de la région. A partir du tracé et de la base de données d'Arc/Info, il nous a été relativement aisé de produire une liste des attributs de chaque tronçon constituant l'itinéraire et d'ainsi vérifier la logique du système. La table de routage nous a également permis d'obtenir l'itinéraire calculé de noeud en noeud, de l'origine à la destination. Malgré la lenteur du déroulement d'une séance, nous avons pu procéder à quelques tests pour nous assurer de la cohérence des résultats.

Enfin, il est possible, dans le cadre spécifique de notre projet, que des erreurs se soient aussi glissées lors de la programmation des diverses fonctions de NodArc. Ce n'est que pour des

Élaboration du SSDRS

raisons techniques (limitations de la gestion mémoire d'Arc/Info-PC) que NodArc a dû être développé. Nous croyons qu'une intégration complète du SSDRS incluant le système expert, à l'intérieur même de la version UNIX d'Arc/Info, limiterait de beaucoup les sources d'erreurs potentielles en plus de faciliter énormément le développement et l'utilisation du système.

Chapitre IX

Conclusion

"[...]L'écrit, c'est comme la guillotine. Tranchant!"

- Daniel Latouche

Le but principal de cette thèse était de démontrer qu'il est possible d'introduire une capacité de raisonnement dans les modèles de recherche d'itinéraires en les couplant à un système expert et à un système d'information géographique pour en faire un système de support à la décision à référence spatiale. En abordant cette démonstration dans le cadre du transport des matières dangereuses, nous avons également voulu illustrer comment l'établissement d'itinéraires sécuritaires pourrait bénéficier de cette capacité de raisonnement.

La quasi-totalité des modèles de recherche d'itinéraires ont recours à l'approche par impédance pour déterminer un trajet optimum. Cette approche est toutefois très limitative puisque cette valeur de pondération unique pour chaque tronçon ne permet pas le raisonnement sur l'ensemble des caractéristiques du tronçon. En outre, en transport des

matières dangereuses, il est très courant de recourir à des modèles probabilistes pour déterminer le degré de risque associé à chaque tronçon. Un tronçon se verra attribuer une probabilité d'occurrence d'accident selon une série de critères préétablis. Cette méthode pose toutefois problème puisque, d'une part, elle exige une connaissance parfaite du réseau de transport. D'autre part, la nature même des accidents dont la probabilité d'occurrence est très faible et difficile à quantifier, diminue considérablement la crédibilité des probabilités qui sont calculées. Nous croyons, grâce à cette recherche, avoir apporté un éclairage nouveau sur ce problème. Nous croyons de plus que la solution proposée permet de palier en grande partie aux lacunes que comportent les modèles traditionnels de recherche d'itinéraires.

Le premier objectif de cette thèse était de faire le point sur les modèles de recherche d'itinéraires basés sur la théorie des graphes et d'en dégager les fondements théoriques et les faiblesses. La théorie des graphes s'avère être un outil puissant de modélisation des réseaux de transport. Outre ses indices morphologiques et topologiques permettant de synthétiser les caractéristiques d'un réseau, la théorie des graphes ouvre la porte, grâce à la représentation matricielle, aux modèles d'optimisation tel que celui du plus court chemin. Parmi les principaux algorithmes du plus court chemin, l'algorithme de Dijkstra s'est avéré le plus efficace pour répondre à nos besoins, tant du point de vue du traitement que de la possibilité d'y intégrer un système expert.

Notre deuxième objectif était de voir comment la théorie de la décision et les systèmes experts sont en mesure d'améliorer l'approche conventionnelle de recherche d'itinéraire. Tout d'abord, la théorie de la décision offre, grâce aux arbres de décision, un outil de formalisation des

problèmes de décision permettant de représenter les événements et leurs conséquences ainsi que les relations qui les unissent. Mais surtout, cette théorie propose une solution alternative à l'approche probabiliste conventionnelle de la prise de décision: le raisonnement qualitatif. Ce type de raisonnement, qui s'apparente à celui d'un expert humain, permet d'inclure des éléments quantitatifs et qualitatifs dans le processus décisionnel, selon la disponibilité des informations.

L'arithmétique d'intervalle et les ensembles flous sont deux outils d'application du raisonnement qualitatif (semi-quantitatif). Alors que l'arithmétique d'intervalle traduit les situations d'informations manquantes ou incertaines sous forme d'intervalles plus ou moins grands mais aux limites précisément définies, les ensembles flous offrent plus de souplesse à cet égard. Ces derniers utilisent le degré d'appartenance à un ensemble pour qualifier les éléments du système. Cette approche est intéressante à plusieurs niveaux. D'une part, elle donne plus de souplesse en éliminant les limites de classes arbitrairement définies. D'autre part, les ensembles flous sont parfaitement adaptés au traitement de l'incertitude et à la représentation des connaissances tel qu'on les retrouve dans les systèmes experts.

De par leur capacité de simuler l'expert humain, les systèmes experts sont en mesure d'introduire un raisonnement quantitatif et qualitatif dans le processus de recherche d'un itinéraire optimum. La représentation des connaissances à l'aide de règles de production où l'on retrouve une prémisse et une action semble tout à fait appropriée au domaine du transport des matières dangereuses. En outre, le mode d'inférence par chaînage avant, qui s'appuie sur les faits pour aboutir à une conclusion, s'avère la meilleure approche pour sélectionner des

tronçons sécuritaires. Enfin, grâce au facteur de certitude, il est possible d'assurer aux connaissances du système expert un degré de certitude (incertitude) permettant d'évaluer le niveau de croyance ou de rejet d'un raisonnement.

Sélectionner un itinéraire sécuritaire pour le transport des matières dangereuses est une décision de nature spatiale qui implique le traitement de données géoréférencées. Un système d'information géographique constitue l'outil idéal à partir duquel bâtir les bases de données nécessaires au modèle de recherche d'itinéraire. Un SIG vectoriel tel qu'Arc/Info offre toutes les fonctions d'intégration, d'édition et de structuration des données en plus d'être un bon outil de gestion globale des réseaux de transport. Ultimement, en couplant un système décisionnel (système expert) à un SIG, on obtient un système de support à la décision à référence spatiale (SSDRS) adapté au transport des matières dangereuses et permettant aux preneurs de décisions d'élaborer des scénarios facilitant la prise de décisions.

Le troisième objectif de cette thèse était d'examiner la notion de risque, son évaluation et sa mitigation ainsi que d'explorer le domaine du transport des matières dangereuses en mettant en évidence les principales caractéristiques des accidents en territoire canadien. La notion de risque est un concept complexe qui fait référence au risque réel et au risque perçu. Alors que le risque réel peut se définir comme étant le produit de la probabilité de survenue par la sévérité des préjudices, le risque perçu, quant à lui, fait appel à un ensemble de critères. La perception du risque varie selon les caractéristiques des individus, le contexte social et économique et les caractéristiques du risque. Déterminer un risque acceptable en transport des matières dangereuses ne relève pas uniquement du risque réel mais aussi du risque perçu par la

société. Dès lors, des critères quantitatifs et qualitatifs doivent faire partie du modèle de recherche d'itinéraire. Le but ultime d'un itinéraire sécuritaire est de minimiser les risques auxquels sont confrontées les populations. Faire disparaître totalement un risque est cependant illusoire. La mitigation des risques est un compromis entre la tolérance zéro et un risque très élevé aux yeux de la population. Ce compromis implique un équilibre entre le système de valeurs d'une société, son système économique et le système physique dans lequel évolue l'activité de transport. En définitive, ce compromis mènera à une hiérarchisation des risques, des risques les plus sévères aux plus mineures, et c'est à partir de cette hiérarchisation que des choix d'actions seront faits.

Malgré le fait que le transport des matières dangereuses occupe une place importante au Canada, la chance a voulu que peu de catastrophes majeures se soient produites dans le passé. Si l'on considère que le plus grand nombre de mouvements de marchandises dangereuses a lieu dans les grands centres urbains du pays, on peut toutefois croire qu'il y aura un jour une ou des catastrophes majeures. Au Canada, ce sont le rail et la route qui s'accaparent de la très grande majorité des accidents. Tous modes de transport confondus, les accidents sont concentrés dans trois régions géographiques soient le corridor Windsor-Québec, l'Alberta et la région de Vancouver. Les matières dangereuses les plus souvent mises en cause sont les liquides inflammables, les gaz comprimés et les matières corrosives.

Dans le cas du transport routier, les matières les plus souvent impliquées dans un accident sont les mêmes, mais avec une part plus importante des liquides inflammables. La répartition géographique des accidents laisse voir une très forte concentration en Alberta en raison de

l'activité pétrolière et une concentration de moindre importance dans l'axe Windsor-Québec. Alors que la majorité des accidents routiers se sont produits en zone rurale, ceux-ci se sont généralement soldés par un déversement ou une fuite.

La contamination de l'environnement est l'une des conséquences les plus fréquentes des accidents bien que le degré de sévérité ait été généralement mineur. Nous croyons cependant que ce portrait plutôt positif du bilan accidentel est un peu trompeur car plus de la moitié des accidents routiers présentaient un degré de sévérité potentielle de modérée à majeure.

Le dernier objectif de cette thèse était de formaliser un modèle de recherche d'itinéraire alliant algorithme traditionnel et système expert et d'en démontrer la démarche d'implantation à la région d'Ottawa-Hull. Trois stratégies ont été développées dans le cadre de cette recherche: la sélection dynamique des tronçons sécuritaires, l'approche par "K" plus courts chemins et la sélection à priori des tronçons sécuritaires. Toutes ces approches sont basées sur un couplage algorithme de Dijkstra - système expert.

L'approche par sélection à priori des tronçons s'est avérée la plus simple et la plus rapide mais présente de sérieuses limitations en ce qui a trait à la fiabilité des résultats. Dans certaines situations, des portions du réseau routier peuvent être rendues inaccessibles. Si l'origine ou la destination de l'itinéraire se situent dans l'une de ces portions, aucune solution d'itinéraire n'est alors possible.

Conclusion

L'approche par sélection dynamique des tronçons sécuritaires s'est révélée plus fiable que la sélection à priori. Cependant, celle-ci favorise nettement l'aspect sécurité de la recherche d'itinéraires au détriment des contraintes économiques puisque le tronçon le plus court n'est retenu qu'en cas d'égalité en terme de sécurité. Bien que ce modèle puisse être appliqué à un système de guidage en temps réel, le rendant ainsi apte à réagir à des changements soudains de conditions routières, la capacité de traitement qu'il exige (lenteur) en limite l'implantation.

C'est la méthode des "K" plus courts chemins qui s'est montrée le meilleur compromis entre la sécurité et les coûts de transport. Des "K" plus courts chemins calculés, le plus sécuritaire est retenu suite à l'évaluation des tronçons par le système expert. Chacun des itinéraires calculés constitue également une solution alternative envisageable pour les preneurs de décisions, ce qui peut favoriser le processus décisionnel. Il s'agit cependant d'une approche dont les performances varient selon la densité et la configuration du réseau. Dans certaines situations, la déconstruction de l'itinéraire optimal peut conduire à des situations désavantageuses.

L'élaboration et l'implantation du système de support à la décision s'est avérée une étape longue et laborieuse. Cette expérience nous a permis de conclure que, dans sa version actuelle, Arc/Info-PC comporte des limitations importantes en ce qui a trait à l'ajout d'applications externes. En outre, l'utilisation de la carte de base au 1:50 000 a exigé un travail considérable de correction des erreurs et de structuration des données pour s'assurer de l'intégrité topologique du réseau. Une bonne partie de ce travail aurait pu être évitée avec l'accès à un fichier déjà structuré pour des applications en transport.

NodArc, bien qu'incomplet, représente un exemple de certaines fonctions que devrait posséder un SSDRS en transport. Le développement d'un tel système aurait toutefois avantage à se faire à l'intérieur même d'un système d'information géographique, préférablement sur la plate-forme Unix d'Arc/Info en raison de sa plus grande flexibilité. Le temps que nous avons consacré à la conception et au développement de cette application nous porte à croire que le développement d'un SSDRS complet ne peut être envisagé que dans un projet à long terme. En effet, une longue période de rodage est nécessaire pour s'assurer de la validation de la base de connaissances et de la fiabilité des conclusions générées par le système expert.

Bien que le but de cette recherche ait été d'apporter un éclairage nouveau sur le problème de la sélection d'itinéraires optimums, plusieurs questions demeurent à explorer. Nous avons délibérément omis de traiter des aspects légaux et moraux du transport des matières dangereuses. Déterminer des itinéraires sécuritaires est une chose, faire respecter ces itinéraires par les camionneurs en est une autre. Des recherches sont actuellement en cours aux États-Unis sur le suivi, par poste de contrôle automatisé, des flottes de camions. En établissant des itinéraires désignés pour le transport des matières dangereuses, on accroît inévitablement le niveau de risque le long de ces routes. On peut sans doute s'interroger sur l'équitabilité d'une telle situation. La réponse à ces questions de nature légale et morale ne peut provenir que d'un choix social (ou de l'absence de choix...). L'acceptabilité d'une situation à risque ne peut se définir uniquement à l'aide de la technique.

L'axe de recherche le plus prometteur nous apparaît être le développement sur plate-forme Unix d'un SSDRS parfaitement intégré. Nos expérimentations nous portent à croire que

l'algorithme des "K" plus courts chemins couplé à un système expert à base de règles de production serait l'approche à privilégier dans de futures recherches. Nous avons d'ailleurs entrepris cette recherche au laboratoire LIAGE de l'Université du Québec à Trois-Rivières.

La représentation des connaissances à caractère spatial dans un système expert nous apparaît également un axe de recherche important à explorer. Si la représentation des connaissances "générales" ne pose pas de problème, ce n'est pas toujours le cas des connaissances géographiques. Par exemple, représenter les mouvements possibles aux intersections (virages à gauche, à droite, interdits) nécessite un lien très étroit entre le système expert et le système d'information géographique. Beaucoup demeure à faire dans ce domaine. Il faut toutefois préciser que, contrairement au début des années 90, l'application des systèmes experts en transport n'est plus une "vue de l'esprit". A titre d'exemple, le Ministère des Transports du Québec (MTQ) possède depuis quelque temps un système expert pour la détection et l'analyse des sites potentiellement dangereux.

Le traitement intégré de l'incertitude à toutes les étapes du processus décisionnel nécessitera également plus d'attention. Dans sa version 5.1, CLIPS ne possédait aucune fonction spécifique de prise en compte de l'incertitude. Cette dernière était une variable au même titre que toutes les autres. Pour produire un SSDRS vraiment utile aux preneurs de décisions, ces derniers devront pouvoir être informés de la valeur des conclusions générées par le système.

Enfin, nous croyons que l'ajout d'une composante qualitative aux modèles de recherche d'itinéraires pourrait déboucher sur d'autres champs de recherche en géographie des

Conclusion

transports, notamment en ce qui concerne la perception de l'espace. Chez l'humain, l'orientation dans l'espace et dans les réseaux s'appuie sur des connaissances procédurales et une carte mentale particulière à chaque individu. Ce type de modèle pourrait donc aider à mieux comprendre le comportement des usagers du réseau routier.

Bibliographie

ABKOWITZ, Mark, DER-MING CHENG, Paul, LEPOFSKY, Mark (1990). *Use of Geographic Information Systems in Managing Hazardous Materials Shipments.* Transportation Research Record No 1261, Washington, USA, pp 35-43.

ABKOWITZ, M., WALSH, S., HAUSER, E., MINOR, L. (1990). Adaptation of Geographic Information Systems to Highway Management. *Journal of Transportation Engineering*, vol 116, no 3, pp. 310-327.

ABKOWITZ, Mark et al (1991). *State and Local Issues in Transportation of Hazardous Materials: Toward a National Strategy.* Transportation Research Record No 1313, Washington, USA, pp 49-54.

ADAMSON, P., TICK, E. (1991). Greedy Partitioned Algorithms for the Shortest-Path Problem. *International Journal of Parallel Programming*, vol 20, no 4, pp. 271-298.

(AFCA) Association française des concepteurs d'abris (1989). *Les risques majeurs et la protection des populations.* Éditions du Moniteur, Paris, France, 272 pages.

ALLBURN, Edward (1991). *Graph Decomposition. Imposing Order on Chaos.* Dr Dobb's Journal, Vol. 16, No. 1, USA, pp 88-96.

ANDERSON, Sven-Eric (1994). Safe Transport of Dangerous Goods: Road, Rail or Sea? A Screening of technical and Administrative Factors. in *European Journal of Operational Research*, vol 75, no 3, pp. 499-507.

ANG, A., BRISCOE, J. et al. (1989). *Development of a System Risk Methodology for Single and Multimodal Transportation Systems.* U.S. DOT, Washington.

ANSELIN, Luc (1989). *What Is Special About Spatial Data ? Alternative Perspectives on Spatial Data Analysis.* Proceedings Symposium on Spatial Statistics, Past, Present and Future, Syracuse, USA, pp 63-77.

ARLINGHAUS, Sandra, ARLINGHAUS, W.C., NYSTUEN, J.D. (1990). The Hedetniemi Matrix Sum: An Algorithm for Shortest Path and Shortest Distance. *Geographical Analysis*, Vol. 22, No. 4, pp. 351-360.

- ARONOFF, Stan (1989).** *Geographic Information Systems: A Management Perspective*. WDL Publications, Ottawa, Canada, 294 pages.
- ASSOCIATION CANADIENNE DE NORMALISATION-ACNOR (1992).** *Lignes directrices sur la gestion des déchets biomédicaux au Canada*. Gouvernement du Canada, Ottawa, 48 pages.
- BAMFORD, C.G., ROBINSON, H. (1978).** *Geography of Transport*. MacDonald and Evans, Londres, Angleterre, 430 pages.
- BEETLE, George R. (1990).** Computer-Based Routing and Scheduling in Metropolitan Regions. *Journal of Transportation Engineering*, vol 116, no 6, pp. 809-814.
- BEROGGI, Giampiero E.G. (1994).** A Real-Time Routing Model for Hazardous Materials. in *European Journal of Operational Research*, vol 75, no 3, pp. 508-520.
- BERRY, Brian J., MARBLE, Duane F. (Editeurs) (1968).** *Spatial Analysis. A Reader in Statistical Geography*. Prentice-Hall, USA, 512 pages.
- BIRD, James (1989).** *The Changing Worlds of Geography*. Clarendon Press, Oxford, Angleterre, 299 pages.
- BOUCHARD, M., MOULIN, B. (1988).** *Une approche méthodologique pour la conception de systèmes experts*. Département d'Informatique, Université Laval, Québec, Canada, 116 pages.
- BOURSIN, Jean-Louis (1986).** *Les structures du hasard. Les probabilités et leurs usages*. Editions du Seuil, France, 186 pages.
- BOWEN, William M., WEEKS, Kenneth D., HILL, Timothy (1989).** A Deep Knowledge Planning Decision Support System for Aiding Nuclear Waste Transportation Decisions. in *Computers, Environment and Urban Systems*, vol 13, no 1, pp. 15-27.
- BRACKEN, Ian, WEBSTER, Christopher (1990).** *Information Technology in Geography and Planning*. Routledge, Londres, Angleterre, 444 pages.
- BRADFORD, M.G., KENT, W.A. (1977).** *Human Geography. Theories and Their Applications*. Oxford University Press, Oxford, Angleterre, 178 pages.
- BREHMER, Berndt, ERIKSSON, Erik, WULFF, P. (1994).** Emerging Trends in Risk Management. in *European Journal of Operational Research*, vol 75, no 3, pp 475-476.
- BREHMER, Berndt, (1994)** Psychological aspects of Traffic Safety. in *European Journal of Operational Research*, vol 75, no 3, pp. 540-552.
- BRIDGEHOUSE, Barbara (1989).** *The Potential Uses of the AMF for Emergency Services*. Proceedings National Conference on GIS, Challenge for the '90s, Ottawa, pp 769-773.

- BRIGGS, K. (1974).** *Introducing Transportation Networks*. University of London Press, Londres, Angleterre, 48 pages.
- BUCKLEY, Fred, HARARY, Frank (1990).** *Distance in Graphs*. Addison-Wesley Publishing Company, Californie, USA, 335 pages.
- BUNGE, William (1966).** *Theoretical Geography*. Royal University of Lund, Lund, Suède, 289 pages.
- BUREAU D'AUDIENCES PUBLIQUES SUR L'ENVIRONNEMENT (1990).** *Les déchets dangereux au Québec. Une gestion environnementale*. Gouvernement du Québec, Québec, Canada, 491 pages.
- BUREAU D'AUDIENCES PUBLIQUES SUR L'ENVIRONNEMENT (1990).** *Les déchets dangereux au Québec. Orientations*. Gouvernement du Québec, Québec, Canada, 208 pages.
- BUREAU D'AUDIENCES PUBLIQUES SUR L'ENVIRONNEMENT (1990).** *Les déchets dangereux au Québec. Situation et enjeux*. Gouvernement du Québec, Québec, Canada, 87 pages.
- BURROUGH, P.A. (1986).** *Principles of Geographical Information Systems for Land Resources Assessment*. Clarendon Press, Oxford, Angleterre, 193 pages.
- BURROUGH, P.A. (1992).** Are GIS Data Structures Too Simple Minded ? *Computers & Geosciences*, Vol. 18, No. 4, pp 395-400.
- BURROUGH, P.A. (1992).** Development of Intelligent Geographical Information Systems. *International Journal of Geographical Information Systems*, vol 6, No 1, pp 1-11.
- BURTON, Ian (1963).** The Quantitative Revolution and Theoretical Geography. *The Canadian Geographer*, Vol. 7, No. 4, Canada, pp 151-162.
- CARNINO, Annick; NICOLET, Jean-Louis; WANNER, Jean-Claude (1990).** *Man and Risks. Technological and Human Risk Prevention*. Marcel Dekker Inc., New York, USA, 241 pages.
- CARTER, James R. (1989).** On Defining the Geographic Information System. in *Fundamentals of Geographic Information Systems: A Compendium*. USA, pp 3-7.
- CARTWRIGHT, John (1990).** CAMEO - Present and Future. *GIS World*, Vol. 3, No. 6, USA, pp 78-80.
- CAYLER, G. (1879).** On the Colouring of Maps. *Royal Geographical Society, Proceedings*, vol 1, pp. 259-261.

- CHALMERS, Lex, MACLENNAN, Mark (1990).** *Expert Systems in Geography and Environmental Studies. An Annotated Review of Recent Work in the Field.* Department of Geography, University of Waterloo, Canada, 69 pages.
- CHEN, Gen-Huey, HUNG, Yung-Chen (1994).** Algorithms for the Constrained Quickest Path Problem and the Enumeration of Quickest Paths. *Computers & Operations Research*, vol 21, no 2, pp. 113-118.
- CHEVALIER, J.-J. (1993).** Systèmes d'aide à la décision à référence spatiale (SADRS): méthode de conception et de développement. *The Canadian Conference on GIS, Proceedings*, Ottawa, pp. 561-572.
- CHORLEY, Richard J., HAGGETT, Peter (Editeurs) (1967).** *Models in Geography.* Methuen & Co Ltd, Londres, Angleterre, 816 pages.
- CHORLEY, Richard J., HAGGETT, Peter (Editeurs) (1969).** *Integrated Models in Geography. Parts I and IV of Models in Geography.* Methuen, London, 677 pages.
- CHORLEY, Richard J., HAGGETT, Peter (Editeurs) (1969).** *Physical and Information Models in Geography. Parts I, II and V of Models in Geography.* Methuen, London, 801 pages.
- CHORLEY, Richard J. (1968).** Geography and Analog Theory. in BERRY, Brian *et al* (1968). *Spatial Analysis. A Reader in Statistical Geography.* Prentice-Hall, USA, 512 pages.
- CHORLEY, Roger (1987).** *Handling Geographic Information.* Department of Environment, Londres, Angleterre, 208 pages.
- CICERI, M.F., MARCHAND, B. et RIMBERT, S. (1977).** *Introduction à l'analyse de l'espace.* Masson, Paris, France, 173 pages.
- CLARKE, Keith C. (1990).** *Analytical and Computer Cartography.* Prentice Hall, Englewood Cliffs, USA, 290 pages.
- COFFEY, William J. (1988).** *Geographical Systems and Systems of Geography.* University of Western Ontario, London, Canada, 211 pages.
- COHN, Louis F.; HARRIS, Roswell A. (1992).** *Knowledge-Based Expert Systems in Transportation. Synthesis of Highway Practice #183.* National Research Board, Washington, USA, 52 pages.
- COPPOCK, J.T., RHIND, D.W. (1991).** The History of GIS. in *Geographical Information Systems. Principles and Applications. Volume 1.* (Maguire *et al*, eds), 447 pages.
- COUCLELIS, Helen (1986).** Artificial Intelligence in Geography: Conjectures on the Shape of Things to Come. *The Professional Geographer*, Vol. 38, No. 1, USA, pp 1-11.

- CRACKNELL, Arthur P. (1986).** Editorial - Geographical Information Systems. *International Journal of Remote Sensing*, vol 7, no 6, pp 721-722.
- CROSS, A. (1989).** *Childhood Leukaemia in Northern England*. University of Newcastle upon Tyne.
- CULLEN, Ian (1986).** Expert Systems in Planning Analysis. *Town Planning Review*, Vol. 57, No. 2, USA, pp 239-251.
- CURRENT, John, MARSH, Michael (1993).** Multiobjective Transportation Network Design and Routing Problems: Taxonomy and Annotation. *European Journal of Operational Research*, vol 65, no 1, pp. 4-19.
- CURRENT, John R., REVELLE, C., COHON, J.L. (1990).** An Interactive Approach to Identify the Best Compromise Solution for Two Objective Shortest Path Problems. *Computers and Operations Research*, vol 17, no 2, pp 187-198.
- DALTON, R., GARLICK, J., MINSHULL, R., ROBINSON, A. (1973).** *Networks in Geography*. George Philip and Sons, Londres, Angleterre, 59 pages.
- DANTZIG, G.B. (1960).** On the Shortest Route Through a Network. *Management Science*, no 6, pp. ??-??
- DAUPHINÉ, André (1979).** *Espace, région et système*. Éditions Économica, Collection Géographia, Paris, France, 167 pages.
- DAUPHINÉ, André (1987).** *Les modèles de simulation en géographie*. Éditions Économica, Paris, France, 187 pages.
- DAUPHINÉ, André; VOIRON-CANICIO, Christine (1988).** *Variogrammes et structures spatiales*. GIP Reclus, Montpellier, France, 56 pages.
- DEMERS, Marthe (1993).** *Le projet de recherche au Ph.D.*. Éditions Hélio, Québec, 175 pages.
- DENIS, Hélène (1990).** *La gestion de catastrophe. Le cas d'un incendie dans un entrepôt de BPC à Saint-Basile-le-Grand*. Gouvernement du Québec, Québec, Canada, 115 pages.
- DESROCHERS, M., VERHOOG, T.W. (1989).** *A Matching Based Savings Algorithm for the Vehicle Routing Problem*. École des Hautes Études Commerciales, Montréal, Canada, 16 pages.
- DESROCHERS, M., VERHOOG, T.W. (1990).** *A New Heuristic for the Fleet Size and Mix Vehicle Routing Problem*. École des Hautes Études Commerciales, Montréal, Canada, 19 pages.
- DEWHURST, F.W., GWINNETT, E.A. (1990).** Artificial Intelligence and Decision Analysis. *Journal of Operational Research Society*, vol 41, no 8, pp. 693-701.

- DIGGLE, P.J., GATRELL, A.C., LOVETT, A.A. (1990).** Modelling the Prevalence of Cancer of the Larynx in Part of Lancashire: A New Methodology for Spatial Epidemiology. in *Spatial Epidemiology*, R.W. Thomas (ed.), pp 35-47.
- DIJKSTRA, E. (1959).** Two Problems in Connexion with Graphs. *Numerische Mathematik*, no 1, pp. 269-271
- DING, Yuemin, DENSHAM, Paul J., ARMSTRONG, M. (1992).** Parallel Processing for Network Analysis: Decomposing Shortest Path Algorithms for MIMD Computers. *Proceedings, 5th International Symposium on Spatial Data Handling*, vol 2, USA, pp. 682-691.
- DOOLEY, J., BURTON, I. (1983).** "Risk Assessment for the Transport of Hazardous Materials" in *Risk. A Symposium on the Assessment and Perception of Risk to Human Health in Canada*. The Royal Society of Canada, 225 pages.
- DOUGLAS, David H. (1994).** A Solution to the Least Cost Path Problem in GIS. *The Canadian Conference on GIS*, Proceedings vol 2, Ottawa, Canada, pp. 1083-1089.
- DUCHESSI, Peter, BELARDO, Salvatore, SEAGLE, John P. (1988).** Artificial Intelligence and the Management Science Practitioner: Knowledge Enhancements to a Decision Support System for Vehicle Routing. *Interfaces*, Vol. 18, No. 2, USA, pp 85-93.
- DUCLOS, Denis (1989).** *La peur et le savoir. La société face à la science, la technique et leurs dangers*. Éditions La Découverte, Paris, France, 308 pages.
- DUMAS, Yvan, DESROSIERS, Jacques (1986).** *Problème de plus court chemin pour la confection de tournées de véhicules avec cueillette, livraison et contraintes d'horaires*. École des Hautes Études Commerciales, Montréal, Canada, 26 pages.
- DUMOLARD, Pierre (1981).** *L'espace différencié. Introduction à une géotaxinomie*. Économica, Paris, France, 201 pages.
- EDMUNDS, Robert A. (1988).** *The Prentice Hall Guide to Expert Systems*. Prentice Hall, USA, 440 pages.
- EGENHOFER, Max J., FRANK, Andrew U. (1989).** Object-Oriented Modeling in GIS: Inheritance and Propagation. *Proceedings Auto-Carto 9*, Baltimore, USA, pp 588-598.
- EGENHOFER, Max J., FRANK, Andrew U. (1990).** LOBSTER: Combining AI and Database Techniques for GIS. *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 6, pp 919-926.
- FABBRI, A. G. (1991).** An Introduction to Expert Systems in Spatial Data Analysis. in Belward, A. S. & Valenzuela, C. (eds) (1991). *Remote Sensing and Geographical Information Systems for Resource Management in Developing Countries*. Kluwer Academic Publishers, 506 pages, pp 389-423.

- FAGHRI, Ardeshir, DEMETSKY, Michael J. (1988).** *Knowledge Representation and Software Selection for Expert Systems Design.* Transportation Research Record No 1187, Washington, USA, pp 1-8.
- FARBEY, B.A., LAND, A.H., MURCHLAND, J.D. (1967).** The Cascade Algorithm for Finding All Shortest Distances in a Directed Graph. *Management Science*, no 14, pp. 19-28.
- FEDRA, Kurt (1989).** Computer-Based Information and Decision Support Systems for Management of Hazardous Substances and Industrial Risk. in *Hazardous Waste Management*, MALTEZOU, S, BISWAS, A., SUTTER, H. (éditeurs). Tycooly Press, Londres, pp. 242-257.
- FEDRA, Kurt, REITSMA, Rene F. (1990).** Decision Support and Geographical Information Systems. in *Geographical Information Systems for Urban and Regional Planning*, Scholten & Stillwell (eds), pp 177-188.
- FEUCHTWANGER, Martin (1989).** Geographic Logical Database Model Requirements. *Proceedings Auto-Carto 9*, Baltimore, pp 599-609.
- FIKSEL, Joseph (1987).** The Impact of Artificial Intelligence on the Risk Analysis Profession. *Risk Analysis*, Vol. 7, No. 3, pp 277-280.
- FINLAY, Paul N. (1990).** Decision Support Systems and Expert Systems: A Comparison of Their Components and Design Methodologies. *Computers and Operations Research*, vol 17, no 6, pp. 535-543.
- FINLAY, Paul N., WILSON, J.M. (1990).** Orders of Validation in Mathematical Modelling. *Journal of Operational Research Society*, vol 41, no 2, pp. 103-109.
- FISCHHOFF, B.; LICHTENSTEIN, P.; SLOVIC, P. et al (1981).** *Acceptable Risk.* Cambridge University Press, 185 pages
- FISHER, Manfred M. (1994).** From Conventional to Knowledge-Based Geographic Information Systems. *Computers, Environment and Urban Systems*, vol 18, no 4, pp. 233-242.
- FISHER, Peter F. (1990).** A Primer of Geographic Search Using Artificial Intelligence. *Computers & Geosciences*, Vol. 16, No. 6, pp 753-776
- FOSTER, Harold D. (1980).** *Disaster Planning. The Preservation of Life and Property.* Springer-Verlag, USA, 275 pages.
- FOSTER, Michael J.; SHAND, Peter J (Éditeurs) (1990).** *The Association for Geographic Information Yearbook 1990.* Taylor & Francis, Londres, Angleterre, 386 pages.
- FOTHERINGHAM, Stewart, ROGERSON, Peter (1994).** *Spatial Analysis and GIS.* Taylor & Francis, Great Britain, 281 pages.

- FRANK, Andrew U. (1992).** Spatial Concepts, Geometric Data Models, and Geometric Data Structures. *Computers & Geosciences*, Vol. 18, No. 4, pp 409-417.
- FRENCH, Simon (1988).** *Decision Theory. An Introduction to the Mathematics of Rationality.* John Wiley & Sons, New York, USA, 448 pages.
- FRENZEL, Louis E. (1987).** *Understanding Expert Systems.* Howard W. Sams & Co., Indiana, USA, 225 pages.
- FREUDENBURG, William R. (1988).** Perceived Risk, Real Risk: Social Science and the Art of Probabilistic Risk Assessment. *Science*, vol 242, no 4875, pp. 44-49.
- GAILE, Gary L.; WILMOTT, Cort J. (Éditeurs) (1984).** *Spatial Statistics and Models.* Reidel Publishing, Dordrecht, Pays-Bas, 482 pages.
- GARRISON, William L. (1960).** Connectivity of the Interstate Highway System. *Papers and Proceedings of the Regional Science Association*, No 6, pp. 121-137.
- GATRELL, Anthony (1983).** *Distance and Space. A Geographical Perspective.* Clarendon Press, Oxford, Grande-Bretagne, 195 pages.
- GATRELL, Anthony, VINCENT, Peter (1991).** Managing Natural and Technological Hazards. in *Handling Geographical Information: Methodology and Potential Applications*, Ian Masser et Michael Blakemore (editors), 317 pages, pp 148-180.
- GETIS, Arthur, BOOTS, Barry (1978).** *Models of Spatial Processes. An Approach to the Study of Point, Line and Area Patterns.* Cambridge University Press, Cambridge, Angleterre, 198 pages.
- GIARRATANO, Joseph, RILEY, Gary (1994).** *Expert Systems Principles and Programming.* PWS Publishing Company, Boston, USA, 644 pages.
- GILLIES, Alan C. (1991).** *The Integration of Expert Systems into Mainstream Software.* Chapman & Hall Computing, London, 246 pages.
- GLICKMAN, Theodore S. (1983).** Rerouting Railroad Shipments of Hazardous Materials to Avoid Populated Areas. *Accident Anal. Prevent.*, no 15, pp. 329-335.
- GLICKMAN, Theodore S. (1991).** An Expedient Risk Assessment of the Highway Transportation of Flammable Liquids in Bulk. *Transportation Science*, Vol. 25, No. 2, pp 115-123.
- GLICKMAN, Theodore S., GOLDING, Dominic, SILVERMAN, Emily D. (1992).** *Acts of God and Acts of Man. Recent Trends in Natural Disasters and Major Industrial Accidents.* Resources for the Future, USA, 65 pages.

- GOODALL, Brian (1987).** *Dictionary of Human Geography.* Penguin Books, Grande-Bretagne, 509 pages.
- GOODCHILD, Michael F. (1992).** Geographical Data Modeling. *Computers & Geosciences*, Vol. 18, No. 4, pp 401-408.
- GOODCHILD, Michael F. (1992).** Geographical Information Science. *International Journal of Geographical Information Systems*, Vol. 6, No. 1, pp 31-45.
- GOODCHILD, Michael F.; GOPAL, Sucharita (Éditeurs) (1989).** *Accuracy of Spatial Databases.* Taylor & Francis, Bristol, USA, 290 pages.
- GOPALAN, Ram, BATTÀ, R., KARWAN, M. H. (1990).** The Equity Constrained Shortest Path Problem. *Computers and Operations Research*, vol 17, no 3, pp. 297-307.
- GORYS, Julius (1990).** *Transportation of Dangerous Goods in the Province of Ontario.* Transportation Research Record No 1264, Washington, USA, pp 57-68.
- GORYS, Julius (1991).** 1988 Ontario Commercial Vehicle Survey. *Transportation Research Record*, No 1313, Transportation Research Board, USA, pp. 20-26.
- GREEN, A.E. (Éditeur) (1982).** *High Risk Safety Technology.* John Wiley & Sons, New York, USA, 654 pages.
- GROUPE DE STAGE DU DEA (1986).** *Modélisation systémique.* Université de Grenoble, Grenoble, France, 143 pages.
- GUARISO, G., WERTHNER, H. (1989).** *Environmental Decision Support Systems.* Ellis Horwood Limited, Angleterre, 240 pages.
- GUERRIER, Philippe, GODON, Daniel (1993).** IGLOU 2.0: A Canadian Decision Support System Based on GIS Technology and Designed for Technological Accidents Prevention, Preparedness and Response. *Proceedings, The Canadian Conference on GIS*, Ottawa, pp. 962-967.
- GUIGO, Maryse (1995).** *Gestion de l'environnement et systèmes experts.* Masson Éditeur, Paris, France, 181 pages.
- GUPTILL, Stephen C. (Éditeurs) (1988).** *A Process for Evaluating Geographic Information Systems.* U.S. Geological Survey, Washington, USA, 133 pages.
- HAASTRUP, Palle (1994).** Overview of Problems of Risk Management of Accidents with Dangerous Chemical in Europe. in *European Journal of Operational Research*, vol 75, no 3, pp. 488-498.

- HAGGETT, Peter, CHORLEY, Richard (1972).** *Network Analysis in Geography.* Edward Arnold Publishers, Londres, Angleterre, 348 pages.
- HAGGETT, Peter (1973).** *L'analyse spatiale en géographie humaine.* Armand Colin, Paris, France, 390 pages.
- HAGGETT, Peter, CLIFF, Andrew D., FREY, Allan (1977).** *Locational Analysis in Human Geography.* Edward Arnold Publishers, Londres, Angleterre, 605 pages.
- HARTSFIELD, Nora, RINGEL, Gerhard (1990).** *Pearls in Graph Theory. A Comprehensive Introduction.* Academic Press Inc., Californie, USA, 246 pages.
- HARWOOD, Douglas W., VINER, John G., RUSSELL, Eugene R. (1990).** *Truck Accident Rate Model for Hazardous Materials Routing.* Transportation Research Record No 1264, Washington, USA, pp 12-23.
- HARWOOD, Douglas W., VINER, John G., RUSSELL, Eugene R. (1993).** Procedure for Developing Truck Accident and Release Rates for HAZMAT Routing. *Journal of Transportation Engineering*, vol 119, no 2, pp. 189-199.
- HERRING, John R. (1992).** TIGRIS: A Data Model for an Object-Oriented Geographic Information System. *Computers & Geosciences*, Vol. 18, No. 4, pp 443-452.
- HITCHCOCK, F. L. (1941).** The Distribution of a Product from Several Sources to Numerous Localities. *Journal of Mathematical Physics*, No 20, pp. 224-230.
- HOBEIKA, Antoine, KIM, S., SETHURAMAN, R. (1993).** Characteristics of Hazardous-Material Accidents in Pennsylvania. *Journal of Transportation Engineering*, vol 119, no 2, pp. 226-238.
- HOLT-JENSEN, Arild (1988).** *Geography History and Concepts.* Barnes & Noble Books, New Jersey, USA, 186 pages.
- HORTON, R.E. (1945).** Erosional Development of Streams and Their Drainage Basins: Hydrophysical Approach to Quantitative Morphology. *Bulletin of the Geological Society of America*, No 56, pp. 275-370.
- HUBERT, Philippe, PAGES, Pierre (1989).** Risk Management for Hazardous Materials Transportation: A Local Study in Lyons. *Risk Analysis*, Vol. 9, No. 4, pp 445-451.
- HUGGETT, Richard (1980).** *Systems Analysis in Geography.* Clarendon Press, Oxford, Angleterre, 208 pages.
- HUGGETT, R.J., THOMAS, R.W. (1980).** *Modelling in Geography. A Mathematical Approach.* Harper & Row, Londres, Angleterre, 338 pages.

- HUXHOLD, William E. (1991).** *An Introduction to Urban Geographic Information Systems.* Oxford University Press, New York, USA, 337 pages.
- HWANG, C.L., PAIDY, S.R., YOON, K. (1980).** Mathematical Programming with Multiple Objectives: A Tutorial. *Computers & Operations Research*, vol 7, no 1-2, pp. 5-31.
- IGNIZIO, James P. (1990).** A Brief Introduction to Expert Systems. *Computers and Operations Research*, vol 17, no 6, pp. 523-533.
- IVANCIE, F. (1984).** *Hazardous Materials Routing Study.* Portland Office of Energy Management, Portland, Oregon.
- JACKSON, Peter (1986).** *Introduction to Expert Systems.* Addison-Wesley Publishing, Wokingham, Angleterre, 246 pages.
- JACOB, Hélène, MONGRAIN, Jean-Pierre, SIMOND, Monique (1988).** *Vocabulaire du transport des marchandises dangereuses.* Gouvernement du Canada, Ottawa, 481 pages.
- JANKOWSKI, P., RICHARD, L. (1994).** Integration of GIS-Based Suitability Analysis and Multicriteria Evaluation in a Spatial Decision Support System for Route Selection. in *Environment and Planning B: Planning and Design*, vol 21, no 3, pp. 323-340.
- JENSEN, J.R., CHRISTENSEN, E.J. (1989).** Solid and Hazardous Waste Disposal Sites Selection Using Digital Geographic Information Systems Techniques. in *Fundamentals of Geographic Information Systems: A Compendium*, William J. Ripple (editor), ACSM, ASPRS, 248 pages, pp 195-206.
- JOHNSON, Glenn O. (1992).** GIS Applications in Emergency Management. *URISA Journal*, vol 4, no 1, pp. 66-72.
- KARIMI, Hasan A. (1994).** An Analysis of Neural Networks Models for Route Optimization. *Proceedings, The Canadian Conference on GIS*, vol 1, Ottawa, pp. 502-508.
- KASPERSON, R.E.; KASPERSON, J.X. (1983).** "Determining the Acceptability of Risk: Ethical and Policy Issues". in *Risk A Symposium on the Assessment and Perception of Risks to Human Health in Canada.* The Royal Society of Canada. 225 pages.
- KAST, Robert (1993).** *La théorie de la décision.* Éditions La Découverte, Paris, France, 125 pages.
- KATZ, Robert, CHEVALIER, J.J. (1994).** Using GIS Data for Mathematical Modelling in Transportation Planning. *Proceedings, The Canadian Conference on GIS*, vol 2, Ottawa, pp. 857-866.
- KHOONG, C.M. (1993).** Shortest-Path Reconstruction Algorithms. *The Computer Journal*, vol 36, no 6, pp. 588-592.

- KIM, T.J., WIGGINS, L.L., WRIGHT, J.R., (Editeurs) (1990).** *Expert Systems: Applications to Urban Planning*. Springer-Verlag, New York, USA, 268 pages.
- KIRKBY, M.J., NADEN, P.S., BURT, T.P., BUTCHER, D.P. (1987).** *Computer Simulation in Physical Geography*. John Wiley & Sons, Chichester, Angleterre, 227 pages.
- KNOKE, David, KUKLINSKI, James H. (1982).** *Network Analysis*. Sage Publications, California, USA, 95 pages.
- KORNHAUSER, Alain L. (1984).** Examples of Corporate Distribution and Transportation Planning Using the Princeton Transportation Network Model and Graphic Information System. *Computer, Environment and Urban Systems*, Vol. 9, No. 1, pp 53-71.
- KUIPERS, Benjamin (1994).** *Qualitative Reasoning. Modeling and Simulation with Incomplete Knowledge*. The MIT Press, Cambridge, USA, 418 pages.
- LANTER, David (1992).** *Intelligent Assistants for Filling Critical Gaps in GIS*. Technical Report 92-4, University of California, Santa Barbara, USA, 77 pages
- LASSARE, Sylvain, FEDRA, K., WEIGKRICHT, E. (1993).** Computer-Assisted Routing of Dangerous Goods for Haute-Normandie. *Journal of Transportation Engineering*, vol 119, no 2, pp. 200-210.
- LAURINI, Robert, MILLERET-RAFFORT, Françoise (1992).** *Pre-Applicative Spatial Reasoning: How to Get Well-Done Spatial Data Bases Able to Support Consistent Application Reasonings ?* SORSA '92 Symposium, Ottawa, 26 pages.
- LAURINI, Robert, MILLERET-RAFFORT, Françoise (1993).** *Les bases de données en géomatique*. Hermes, Série Géomatique, Paris, France, 340 pages.
- LAURINI, Robert, THOMPSON, Derek (1992).** *Fundamentals of Spatial Information Systems*. Academic Press, The APIC Series, Grande Bretagne, 680 pages.
- LEE, Y.C., MASRY, S.E. (1989).** On the Limitations of Planar Graphs in Modelling Spatial Features. *Proceedings National Conference on GIS, Challenge for the '90s*, Ottawa, pp 1035-1043.
- LEPOFSKY, Mark, ABKOWITZ, M., CHENG, P. (1993).** Transportation Hazard Analysis in Integrated GIS Environment. *Journal of Transportation Engineering*, vol 119, no 2, pp. 239-254.
- LEROUX, Denis G. (1987).** *Les systèmes d'information à référence spatiale (SIRS). Etude bibliographique*. Université de Sherbrooke, Sherbrooke, Canada, 41 pages.
- LEVESQUE, Guy (1995).** *Nucléaire et perception du risque: le cas de la centrale Gentilly 2*. Université du Québec à Trois-Rivières, Québec, 91 pages.

- LINDER-DUTTON, Laurel, BATA, Rajan, KARWAN, Mark H. (1991).** Equitable Sequencing of a Given Set of Hazardous Materials Shipments. *Transportation Science*, Vol. 25, No. 2, pp 124-137.
- LIST, George, MIRCHANDANI, Pitu (1991).** An Integrated Network/Planar Multiobjective Model for Routing and Siting for Hazardous Materials and Wastes. *Transportation Science*, Vol. 25, No. 2, pp 146-156.
- LIST, George F., MIRCHANDANI, Pitu B., TURNQUIST, Mark A. ZOGRAFOS, Konstantinos (1991).** Modeling and Analysis for Hazardous Materials Transportation: Risk Analysis, Routing/Scheduling and Facility Location. *Transportation Science*, Vol. 25, No. 2, pp 100-114.
- LOSCH, A. (1954).** *The Economics of Location*. Yale University Press, New Haven, USA, 187 pages.
- LOTAN, Tsippy, KOUTSOPOULOS, Haris (1993).** Route Choice in the Presence of Information Using Concepts from Fuzzy Control and Approximate Reasoning. *Transportation Planning and Technology*, vol 17, no 2, pp. 113-126.
- LOUKES, David K. (1994).** Development of Canadian Data Standards for GIS in Transportation (GIS-T). *Proceedings, The Canadian Conference on GIS*, vol 1, Ottawa, pp. 93-104.
- LOWE, John C., MORYADAS, S. (1975).** *The Geography of Movement*. Houghton Mifflin Company, Boston, USA, 333 pages.
- LUPIEN, Anthony E., MORELAND, William H. DANGERMOND, Jack (1987).** Network Analysis in Geographic Information Systems. *Photogrammetric Engineering and Remote Sensing*, Vol. 53, No. 10, pp 1417-1421.
- MACMILLAN, Bill (Editeur) (1989).** *Remodelling Geography*. Basil Blackwell, Oxford, Angleterre, 348 pages.
- MAGGIO, Robert C. (1987).** The Role of the Geographic Information Systems in the Expert System. *GIS'87 Proceedings*, ACSM, ASPRS, USA, pp 685-692.
- MAGUIRE, David J. (1989).** *Computers in Geography*. Longman Scientific & Technical, New York, USA, 248 pages.
- MAGUIRE, David J., GOODCHILD, Michael F. & RHIND, David W. (Éditeurs) (1991).** *Geographical Information Systems. Principles and Applications. Volume 1: Principles*. Longman Scientific & Technical, New York, USA, 649 pages.
- MAGUIRE, David J., GOODCHILD, Michael F. & RHIND, David W. (Éditeurs) (1991).** *Geographical Information Systems. Principles and Applications. Volume 2: Applications*. Longman Scientific & Technical, New York, USA, 447 pages.

- MAGUIRE, David, HICKIN, Bill, LONGLEY, Ian et MASEV, Trajco (1991).** Waste Disposal Site Selection Using Raster and Vector GIS. in *Mapping Awareness and Integrated Spatial Information Systems*, vol. 5, No. 1, pp 24-27.
- MAHMASSANI, Hani S. (1984).** Uncertainty in Transportation Systems Evaluations: Issues and Approaches. *Transportation Planning & Technology*, vol 9, pp. 1-12.
- MALTEZOU, Sonia, BISWAS, Asit K., SUTTER, Hans (eds) (1989).** *Hazardous Waste Management*. Tycooly Publishing, Angleterre, 344 pages.
- MARK, David M. (1993).** Human Spatial Cognition. in *Human Factors in Geographical Information Systems*, Medyckyj, Scott & Hearnshan, H. (éditeurs), Belhaven Press, Grande-Bretagne, 266 pages.
- MARTIN, Larry R.G., LAFOND, Gilbert (Editeurs) (1988).** *Risk Assessment and Management: Emergency Planning Perspectives*. University of Waterloo Press, Waterloo, Canada, 352 pages.
- MASSER, Ian, BLAKEMORE, Michael (eds) (1991).** *Handling Geographical Information: Methodology and Potential Applications*. Longman Scientific & Technical, Grande Bretagne, 317 pages.
- McBRIDE, Richard (1982).** Controlling Left and U-Turns in the Routing of Refuse Collection Vehicles. *Computers & Operations Research*, vol 9, no 2, pp. 145-152.
- McBRIDE, Richard D., O'LEARY, Daniel E. (1993).** The Use of Mathematical Programming with Artificial Intelligence and Expert Systems. *European Journal of Operational Research*, vol 70, no 1, pp. 1-15.
- MEHLDAU, Gerhard, SCHOWENGERDT, Robert A. (1990).** A C-Extension for Rule-Based Image Classification Systems. *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 6, pp 887-892.
- MENON, S., SMITH, Terence R. (1989).** A Declarative Spatial Query Processor for Geographic Information Systems. *Photogrammetric Engineering and Remote Sensing*, Vol. 55, No. 11, pp 1593-1600.
- MERLIN, Pierre (1991).** *Géographie, économie et planification des transports*. Presses Universitaires de France, Paris, 472 pages.
- MILLER, Harver J. (1993).** GIS and the Geometry of Facility Location Problems. *GIS/LIS '93, Proceedings Annual Conference*, vol 2, pp. 530-539.
- MILLER, Marc, THÉRIAULT, Marius (1990).** Les systèmes d'information géographique ou la dernière tentation du géographe. *Cahiers de Géographie du Québec*, Vol. 34, No. 91, Québec, Canada, pp 81-85.

- MINISTÈRE DE L'ENVIRONNEMENT (1990).** *Canada's Green Plan.* Gouvernement du Canada, Ottawa, 174 pages.
- MINSHULL, Roger (1975).** *An Introduction to Models in Geography.* Longman, Londres, Angleterre, 162 pages.
- MOLENAAR, M. (1991).** Status and Problems of Geographical Information Systems. The Necessity of a Geoinformation Theory. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 46, pp 85-103.
- MONDOU, Jean-François, CRAINIC, T., NGUYEN, S. (1991).** Shortest Path Algorithms: A Computational Study with the C Programming Language. *Computers and Operations Research*, vol 18, no 8, pp. 767-786.
- MOREHOUSE, Scott (1992).** The Arc/Info Geographic Information System. *Computers & Geosciences*, Vol. 18, No. 4, pp 435-441.
- MOULINS, Philippe de (1988).** *Les accidents technologiques. Les données de l'expérience analyses et recommandations.* Centre National de Prévention et de Protection, Paris, 137 pages.
- MOUNSEY, Helen, TOMLINSON, Roger F. (Editeurs) (1988).** *Building Databases for Global Science.* Taylor & Francis, Londres, Angleterre, 419 pages.
- MOYERS, Bill (1990).** *Global Dumping Ground. The International Traffic in Hazardous Waste.* Seven Locks Press, Washington, USA, 152 pages.
- NATIONAL RESEARCH BOARD (1986).** *Recent Advances in Hazardous Materials Transportation Research.* National Research Council, Washington, USA, 212 pages.
- NYSTUEN, John D., DACEY, Michael F. (1961).** A Graph Theory Interpretation of Nodal Regions. *Paper and Proceedings of the Regional Science Association*, No 7, pp. 29-42.
- OCDE (1988).** *Transport Routier de marchandises dangereuses.* Organisation de Coopération et de Développement Économique, France, 153 pages.
- OLIVIER, Marc (1995).** *Connaitre et contrôler les matières dangereuses.* Guérin Éditeur, Montréal, Canada, 417 pages.
- ONTARIO WASTE MANAGEMENT CORPORATION OWMC (1988).** *Environmental Assessment for a Waste Management System. Volume 1. The OWMC Undertaking.* Gouvernement de l'Ontario, 1988.
- OPENSHAW, Stan (1991).** A View on the GIS Crisis in Geography, or, Using GIS to Put Humpty-Dumpty Back Together Again. *Environment and Planning A*, vol 23, No 5, pp 621-628.
- ORE, Oystein (1970).** *Les graphes et leurs applications.* Dunod, Paris, France, 144 pages.

- ORCHARD, G.A.; PHILLIPS, W.A. (1991).** *Neural Computation. A Beginner's Guide.* Lawrence Erlbaum Associates Publishers, East Sussex, United Kingdom, 141 pages.
- PEARN, Wen Lea, LI, Mao Lin (1994).** Algorithms for the Windy Postman Problem. *Computers & Operations Research*, vol 21, no 6, pp. 641-651.
- PENA, Orlando; SANGUIN, André-Louis (1986).** *Concepts et méthodes de la géographie.* Guérin Éditeur, Montréal, Canada, 177 pages.
- PEUQUET, Donna J., MARBLE, Duane F. (Editeurs) (1990).** *Introductory Reading in Geographic Information Systems.* Taylor & Francis, Londres, Angleterre, 371 pages.
- PIJAWKA, D., FOOTE, S., SOESILO, A. (1985).** Risk Assessment of Transporting Hazardous Material: Route Analysis and Hazard Management. *Transportation Research Record*, no 1020, pp. 1-6.
- POTVIN, J.-Y., LAPALME, G., ROUSSEAU, J.-M. (1989).** ALTO: A Computer System for the Design of Vehicle Routing Algorithms. *Computer & Operations Research*, vol 16, no 5, pp. 451-470.
- RAGHUNATHAN, Srinivasan (1994).** An Application of Qualitative Reasoning to Derive Behavior from Structure of Quantitative Models. *Journal of Management Information Systems*, vol 11, no 1, pp. 73-100.
- RAPER, Jonathan F., MAGUIRE, David J. (1992).** Design Models and Functionality in GIS. *Computers & Geosciences*, Vol. 18, No. 4, pp 387-394.
- REVELLE, Charles, COHON Jared, SHOBRY, Donald (1991).** Simultaneous Siting and Routing in the Disposal of Hazardous Wastes. *Transportation Sciences*, Vol. 25, No. 2, pp 138-145.
- ROBERGE, Roger A. (1984).** Risk Analysis and Railway Crossing Safety: The Windsor to Quebec Corridor. *The Operational Geographer*, No. 4, pp 14-19.
- ROBERGE, Roger A., ROBERGE, Linda (1986).** *Railway Risk Management and the Spatial Allocation of Subsidies: A Causal Analysis.* Research Notes #54, Department of Geography, University of Ottawa, 17 pages.
- ROBINSON, Vincent B. (1985).** About Different Kinds of Uncertainty in Collections of Spatial Data. *Proceedings Auto-Carto 7*, Washington, pp 440-449.
- ROBINSON, Vincent B., FRANK, Adrew U. (1987).** Expert Systems for Geographic Information Systems. *Photogrammetric Engineering and Remote Sensing*, Vol. 53, No. 10, pp 1435-1441
- ROBINSON, Vincent B., FRANK, Adrew U. & Matthew A. BLAZE (1986).** Expert Systems Applied to Problems in Geographic Information Systems: Introduction, Review and Prospects. *Computer, Environment and Urban Systems*, vol 11, No 4, pp 161-173

- ROLSTON, David W. (1988).** *Principles of Artificial Intelligence and Expert Systems Development.* McGraw-Hill, New York, USA, 257 pages.
- ROUSSEAU, Jacques (1986).** *Le transport des marchandises dangereuses. Étude générale pour les parlementaires.* Gouvernement du Canada, Ottawa, 13 pages.
- ROWE, William D. (1994).** Understanding Uncertainty. *Risk Analysis*, vol 14, no 5, pp. 743-750.
- RUBISE, Patrick; GAUTIER, Yves (1995).** *Les risques technologiques.* Cité des Sciences et de l'Industrie, Pocket, France, 127 pages.
- RUSSELL, Eugene R. (1993).** Rating Countermeasures for Mitigation of Hazardous Materials Incidents. *Journal of Transportation Engineering*, vol 119, no 2, pp. 211-225.
- SACCOMANNO, F.F., SHORTREED, J.H. (1993).** HAZMAT Transport Risks: Societal and Individual Perspectives. *Journal of Transportation Engineering*, vol 119, no 2, pp. 177-188.
- SACCOMANNO, F.F., VAN AERDE, M., QUEEN, D. (1987).** *Interactive Selection of Minimum-Risk Routes for Dangerous Goods Shipments.* Transportation Research Record No 1148, Washington, USA, pp 9-17.
- SCANLON, Raymond D. (Editeur) (1987).** *Hazardous Materials Hazardous Waste: Local Management Options.* International City Management Association, Washington, USA, 223 pages.
- SCHENK, Tony, ZILBERSTEIN, Ofer (1990).** Experiments with a Rule-Based System for Interpreting Linear Map Features. *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 6, pp 911-917.
- SCHNEIDER, Jerry B. (1984).** A Review of a Decade of Applications of Computer Graphics Software in the Transportation Field. *Computer, Environment and Urban Systems*, Vol. 9, No. 1, pp 1-20.
- SCHOLTEN, Henk J., STILLWELL, John C. (1990).** *Geographical Information Systems for Urban and Regional Planning.* Kluwer Academic Publishers, Dordrecht, Pays-Bas, 261 pages.
- SELKIRK, K.E. (1982).** *Pattern and Places. An Introduction to the Mathematics of Geography.* Cambridge University Press, Cambridge, Angleterre, 203 pages.
- SHAPIRO, S.C. (ed.) (1987).** *Encyclopedia of Artificial Intelligence.* John Wiley and Sons, USA, 1219 pages.
- SHORTREED, J.H. (Editeur) (1985).** *Dangerous Goods Movements.* Proceedings of the 1984 Waterloo Workshop. University of Waterloo Press, Waterloo, Canada, 235 pages.

- SHORTREED, J.H. (Editeur) (1989).** *Risk Management for Dangerous Goods.* University of Waterloo Press, Waterloo, Canada, 176 pages.
- SIEKIERSKA, E.M. (1989).** Advantages of Expert Systems with the Examples of the Use of the VP.Expert System for Cartographic Applications. *Conférence Nationale SIG Défi des Années 90,* Ottawa, pp 710-717.
- SIMKOWITZ, Howard J. (1989).** GIS in The Transportation Community. *GIS World,* Vol. 2, No. 5, USA, pp 53-55.
- SIMKOWITZ, Howard J. (1990).** *Integrating Geographic Information System Technology and Transportation Models.* Transportation Research Record No 1271, Washington, USA, pp 44-47.
- SMITH, Terence R. (1984).** Artificial Intelligence and Its Applicability to Geographical Problem Solving. *The Professional Geographer,* Vol. 36, No. 2, USA, pp 147-158.
- SMYTH, Stephen (1992).** A Representational Framework for Route Planning in Space and Time. *Proceedings, 5th International Symposium on Spatial Data Handling,* vol 2, USA, pp. 692-701.
- STRAHLER, A.N. (1952).** Hypsometric Analysis of Erosional Topography. *Bulletin of the Geological Society of America,* No 63, pp. 1117-1142.
- SYAM, Sid S., COURTNEY, James F. (1994).** The Case for Research in Decision Support Systems. *European Journal of Operational Research,* vol 73, no 3, pp. 450-457.
- TAAFFE, Edward J., GAUTHIER, Howard L. (1973).** *Geography of Transportation.* Prentice-Hall, Englewood Cliffs, USA, 226 pages.
- TEODOROVIC, DUSAN. (1986).** *Transportation Networks. A Quantitative Treatment.* Gordon and Breach Science Publishers, Londres, Angleterre, 219 pages.
- TEODOROVIC, Dusan (1994).** Fuzzy Sets Theory Applications in Traffic and Transportation. *European Journal of Operational Research,* vol 74, no 3, pp. 379-390.
- TIKUNOV, Vladimir Sergeevic (1990).** Recherches dans le domaine de l'intelligence artificielle et systèmes experts en géographie. *Annales de Géographie,* No. 552, pp 129-140.
- TINKLER, K.J. (1977).** *Introduction to Graph Theoretical Methods in Geography.* Geo Abstract Ltd, Norvège, 56 pages.
- TOMLINSON, Roger F. (1989).** Presidential Address: Geographic Information Systems and Geographers in the 1990s. *The Canadian Geographer,* vol 33, no 4, pp 290-298.
- TRANSPORT CANADA (1986).** *Transport des marchandises dangereuses. Rapport Annuel 1985-1986.* Gouvernement du Canada, Ottawa, 18 pages.

- TRANSPORT CANADA (1990).** *Marchandises dangereuses. La classification et les indications de danger.* Gouvernement du Canada, Ottawa, 15 pages.
- TRANSPORT CANADA (1991).** *Transport des marchandises dangereuses. Rapport Annuel 1990-1991.* Gouvernement du Canada, Ottawa, 16 pages.
- VAN AERDE, M., SHORTREED, J., STEWART, A.M., MATTHEWS, M. (1989).** Assessing the Risks Associated with the Transport of Dangerous Goods by Truck and Rail Using the RISKMOD Model. *Canadian Journal of Civil Engineering*, vol 16, no 3, pp. 326-334.
- VAN ZIJL, Lynette, MITTON, Deon, CROSBY, Simon (1992).** A Tool for Graphical Network Modeling and Analysis. *IEEE Software*, January 1992, USA, pp 47-54.
- VINCENT, P.J. (1989).** Transport of Hazardous Waste and GIS. *Civil Protection*, Summer 1989, pp 5-6.
- VONDEROHE, A.P., TRAVIS, L., SMITH, R.L., TSAI, V. (1993).** *Adaptation of Geographic Information Systems for Transportation.* Transportation Research Board, Report 359, National Research Council, Washington, USA, 55 pages.
- WAGNER, Daniel F., SCOTT, M., POSEY, A. (1993).** Improving Best Path Solutions in Raster GIS. *Proceeding GIS/LIS '93 Annual Conference*, vol 2, Mineapolis, pp. 718-726.
- WAHLSTROM, Bjorn (1994).** Models, modelling and modellers: An Application to Risk Analysis. in *European Journal of Operational Research*, vol 75, no 3, pp. 477-487.
- WALTER, Volker, FRITSCH, Dieter (1994).** GIS Data Structures for Vehicle Navigation Systems. *Proceedings, The Canadian Conference on GIS*, vol 1, Ottawa, pp. 489-501.
- WARNTZ, W. (1961).** Trans-Atlantic Paths and Pressure Patterns. *Geographical Review*, no 51, pp. 187-212.
- WATERS, C.D.J. (1990).** Expert Systems for Vehicle Scheduling. *Journal of Operational Research Society*, vol 41, no 6, pp. 505-515.
- WATERS, Nigel M. (1989).** Do You Sincerely Want to Be a Gis Analyst ? *The Operational Geographer*, Vol. 7, No. 4, Canada, pp 30-35.
- WATERS, Nigel M. (1989b).** Expert Systems within a GIS: Knowledge Acquisition for Spatial Decision Support Systems. *Actes, Conférence Nationale SIG Défi des années 90*, Ottawa, pp. 740-759.
- WATERSTONE, Marvin (1989).** Risk Analysis and Management of Natural and Technological Hazards: A Social/Behavioral Science Perspective. in *Risk Analysis and Management of Natural and Man-Made Hazards*. National Science Foundation, USA, pp 72-80.

- WERNER, Christian (1985).** *Spatial Transportation Modeling*. Sage Publications, Beverly Hills, USA, 95 pages.
- WERTHNER, Hannes (1994).** *Qualitative Reasoning. Modeling and the Generation of Behavior*. Springer-Verlag, Allemagne, 180 pages.
- WHITE, D.J. (1982).** The Set of Efficient Solutions for Multiple Objective Shortest Path Problems. *Computers & Operations Research*, vol 9, no 2, pp. 101-107.
- WHITE, H.P., SENIOR, M.L. (1983).** *Transport Geography*. Longman, New York, USA, 224 pages.
- WHYTE, Anne V. (1983).** "Probabilities, Consequences and Values in the Perception of Risk." in *A Symposium on the Assessment and Perception of Risk to Human Health in Canada*. The Royal Society of Canada, 225 pages.
- WILSON, A.G. (1981).** *Geography and the Environment. Systems Analytical Methods*. John Wiley & Sons, New York, USA, 297 pages.
- WILSON, Frank, CANTIN, L., BISSON, B. (1990).** Emergency Response System for Dangerous Goods Movements by Highways. *Journal of Transportation Engineering*, vol 116, no 6, pp. 789-794.
- WILSON, Robin J., BEINEKE, Lowell W. (1979).** *Applications of Graph Theory*. Academic Press, Angleterre, 423 pages.
- WILSON, Robin J., WATKINS, John J. (1990).** *Graphs. An Introductory Approach*. John Wiley & Sons, New York, USA, 340 pages.
- WOLDENBERG, Michael J. (1968).** *Hierarchical Systems: Cities, Rivers, Alpine Glaciers, Bovine Livers, and Trees*. Harvard University, Cambridge, USA, 150 pages.
- WOLKOWITSCH, Maurice (1973).** *Géographie des transports*. Armand Collin Éditeur, Paris, France, 381 pages.
- WOLKOWITSCH, Maurice (1992).** *Géographie des transports*. Armand Collin Éditeur, Paris, France, 191 pages.
- WORLD COMMISSION ON ENVIRONMENT AND DEVELOPMENT (1987).** *Our Common Future. The World Commission on Environment and Development*. Oxford University Press, Grande Bretagne, 400 pages.
- WORRALL, Les (Editeur) (1990).** *Geographic Information Systems: Developments and Applications*. Belhaven Press, Londres, Angleterre, 251 pages.

Bibliographie

YAN, Joel Z., WITIUK, Sidney W. (1989). Innovative GIS Applications Using Statistics Canada Information System Products - An Update. *Proceedings of SaskGIS'89*, Regina, Saskatchewan, pp 149-170.

YEATES, M. (1963). Hinterland Delimitation: A Distance Minimizing Approach. *Professional Geographer*, No 15, pp. 7-10.

ZADEH, Lofti (1965). Fuzzy sets. in *Information and Control*, no 8, pp. 338-353.

Annexe 1

Fonctions de conversion et de création des fichiers NodArc

Note: Ce code source décrit la structure des fichiers NodArc 1.0 et en permet également la création. Il faut cependant prendre note que plusieurs sous-fonctions appelées par le module CONVERT ne sont pas incluses dans cette annexe (exemple: fonctions d'affichage).

```
/*
*****
* Le module CONVERT.C contient les fonctions permettant de convertir des
* fichiers de formats BNA (Atlas GIS) GNI (ARC/INFO) en format NDS - LNK
* propre a NODARC. Le module CONVERT.C permet egalement de filtrer des
* fichiers GNL d'ARC/INFO et d'en extraire des elements selon leur
* identificateur numerique.
*
* Programmation: Denis Leroux, fevrier 1994
*****
*/

#include <stdio.h>
#include <graphics.h>
#include <dos.h>
#include <dir.h>
#include <alloc.h>
#include <stdlib.h>
#include <time.h>
#include <conio.h>
#include <string.h>
#include <io.h>
#include <math.h>
#include <bios.h>
#include "system.h"
#include "vpfview.h"

/* Prototypes des fonctions du module CONVERT.C */

void bna_nds( void );
void bna_to_nds( char *name_in, char *name_out );
void arcinfo_bna( void );
```

```

void conv_arc_bna( char *name_in, char *name_out, long min, long max );
void arcinfo_filter( void );
void conv_arc_filt( char *name_in, char *name_out, long min, long max );
void conv_bna_lnk( void );
int conv_lnk( char *name_in_bna, char *name_in_nds, char *name_in_lnk );
void create_lat_file();

```

```

/*****
* Cette fonction fait la lecture d'un fichier *.BNA, en determine les
* coordonnees minimum et maximum des elements PONCTUELS seulement puis
* par la suite cree a partir des elements ponctuels un fichier *.NDS de
* la structure suivante:
*
*      Node_Id, Node_Type, X_Coord, Y_Coord, Nom
*
* Denis Leroux, decembre 1993
*****/
*/

```

```

void bna_to_nds( char *name_in, char *name_out )
{

```

```

FILE *fp_bna, *fp_nds;

```

```

float xmin, xmax, ymin, ymax, x_coord, y_coord;
char c;
char ident_1[65], ident_2[65], limchar[25];
window_type info;
int i, nbre_pts, coul_actuel;
int node_type, node_id, ligne, nbre_point, polygone;

```

```

/* Initialisation des variables */
xmin = 3400000000000000.0;
ymin = 3400000000000000.0;
xmax = -3400000000000000.0;
ymax = -3400000000000000.0;

```

```

nbre_point = 0;
ligne = 0;
polygone = 0;
node_type = 1;
node_id = 0;

```

```

/* Va chercher la couleur actuelle de tracage */
coul_actuel = getcolor();

```

```

if( (fp_bna = fopen( name_in,"r" ) ) != NULL )
{
  /* Affiche l'heure de debut de la procedure */
  temps_debut();

  info = info_window("Scanning BNA File...");
  time_cursor();

  while( !feof( fp_bna ) )
  {
    c = fgetc( fp_bna );          /* lire " */
    fscanf(fp_bna,"%[^\""]",ident_1); /* Lire ident_1 */
    c = fgetc( fp_bna );          /* Lire " */
    c = fgetc( fp_bna );          /* Lire , */
    c = fgetc( fp_bna );          /* Lire " */
    fscanf(fp_bna,"%[^\""]",ident_2); /* Lire ident_2 */
    c = fgetc( fp_bna );          /* Lire " */
    c = c;
    fscanf(fp_bna,"%i\n",&nbre_pts);

    if( nbre_pts < 0 ) ligne++;
    if( nbre_pts > 2 ) polygone++;
    if( nbre_pts == 1 ) nbre_point++;

    if( nbre_pts != 1 )
    {
      for( i=0; i < abs(nbre_pts); i++ )
      {
        fscanf(fp_bna,"%f,%f\n",&x_coord,&y_coord);
      }
    }
    if( nbre_pts == 1 )
    {
      fscanf(fp_bna,"%f,%f\n",&x_coord,&y_coord);
      if( x_coord < xmin ) xmin = x_coord;
      if( x_coord > xmax ) xmax = x_coord;
      if( y_coord < ymin ) ymin = y_coord;
      if( y_coord > ymax ) ymax = y_coord;
    }
  }

  /* Efface le message du scanning BNA file */
  close_window( &info );
  arrow_cursor();

  /* Place la couleur de tracage a blanc */

```

```

setcolor( WHITE );
hidemousecursor();

/* Affiche les resultats du scanning dans File Info */
outtextxy(462,155,"# Points :");
outtextxy(462,168,"# Lines :");
outtextxy(462,181,"# Polygons:");
itoa(nbre_point,limchar,10);
outtextxy(555,155,limchar);
itoa(ligne,limchar,10);
outtextxy(555,168,limchar);
itoa(polygone,limchar,10);
outtextxy(555,181,limchar);

if( nbre_point > 0 )
{
    outtextxy(462,95,"X min:");
    outtextxy(462,108,"Y min:");
    outtextxy(462,121,"X max:");
    outtextxy(462,134,"Y max:");
    sprintf(limchar,"%0.3f",xmin);
    outtextxy(510,95,limchar);
    sprintf(limchar,"%0.3f",ymin);
    outtextxy(510,108,limchar);
    sprintf(limchar,"%0.3f",xmax);
    outtextxy(510,121,limchar);
    sprintf(limchar,"%0.3f",ymax);
    outtextxy(510,134,limchar);
}

showmousecursor();

if( nbre_point > 0 )
{
    rewind(fp_bna);
    fp_nds = fopen(name_out,"w");

    info = info_window("Building NDS File...");
    time_cursor();

    /* Ecrire les limites de la carte */
    fprintf(fp_nds,"%f %f %f %f",xmin,ymin,xmax,ymax);

    while( !feof( fp_bna ) )
    {
        c = fgetc( fp_bna );          /* lire " */
    }
}

```

```

fscanf(fp_bna,"%[^\""]",ident_1); /* Lire ident_1 */
c = fgetc( fp_bna );           /* Lire " */
c = fgetc( fp_bna );           /* Lire , */
c = fgetc( fp_bna );           /* Lire " */
fscanf(fp_bna,"%[^\""]",ident_2); /* Lire ident_2 */
c = fgetc( fp_bna );           /* Lire " */
fscanf(fp_bna,"%i\n",&nbre_pts);

if( nbre_pts != 1 )
{
    for( i=0; i < abs(nbre_pts); i++ )
    {
        fscanf(fp_bna,"%f,%fn",&x_coord,&y_coord);
    }
}
if( nbre_pts == 1 )
{
    fscanf(fp_bna,"%f,%fn",&x_coord,&y_coord);
    fprintf(fp_nds,"\n%i %i %f %s",node_id,node_type,x_coord,y_coord,ident_1);
    node_id++;
}
}
/* Efface le message Building NDS */
close_window( &info );
arrow_cursor();

/* Place le mode couleur a blanc et cache la souris */
setcolor( WHITE );
hidemousecursor();

outtextxy(462,205,"# Features in :");
outtextxy(462,218,"# Features out:");
itoa(nbre_point+ligne+polygone,limchar,10);
outtextxy(580,205,limchar);
itoa(node_id,limchar,10);
outtextxy(580,218,limchar);

fclose( fp_bna );
fclose( fp_nds );

/* Affiche l'heure de fin de la procedure */
temps_fin();

setcolor( coul_actuel );
showmousecursor();
}

```

```

else
{
    display_message("No point in BNA file!");
    fclose(fp_bna);
}
}
}

/*****
*
* Cette fonction fait la lecture d'un fichier UNGEN d'ARC/INFO contenant *
* des elements lineaires et en fait la conversion en format .BNA. Le *
* fichier *.BNA obtenu renferme des elements lineaires seulement. Cette *
* fonction permet egalement de filtrer les identificateurs en ne *
* conservant que les elements dont les identificateurs sont compris *
* entre les valeurs minimum et maximum specifiees a l'entree. *
*
* Parametres: name_in = nom du fichier d'entree ARC/INFO *
*             name_out = nom du fichier de sortie *.BNA *
*             lim_min = limite minimum des identificateurs *
*             lim_max = limite maximum des identificateurs *
*
*
* Valeur possible des identificateurs : -32766 a 32766 *
* Nombre maximal de coordonnees par arc: 32767 *
* Nombre maximal d'arcs par fichier : 32000 *
*
* Denis Leroux, novembre 1993, fevrier 1994
*****/
*/

void conv_arc_bna( char *name_in, char *name_out, long min, long max )
{
FILE *fp_arc, *fp_bna;

char ligne[79];
char limchar[25];
window_type info;
int compteur;
int z;
int ver_end;
int *id_list;
int id_compt;
int transfert, traites;

```

```

int i, j;
int ident;
int coul_actuel;
float x, y;

/* On stock en memoire la couleur actuelle de tracage */
coul_actuel = getcolor();

id_list = (int *)vpmalloc(32000);

if( id_list != NULL )
{
    if( ( fp_arc = fopen( name_in,"r" ) ) != NULL )
    {
        /* Affiche l'heure de debut de la procedure */
        temps_debut();

        fp_bna = fopen( name_out,"w" );

        fgets(ligne,79,fp_arc);
        id_compt = 0;
        compteur = 1;
        transfert = 0;
        traites = 0;

        /* Affiche l'operation en cours */
        info = info_window("Scanning Input File...");
        time_cursor();

        while( !feof( fp_arc ) )
        {
            ver_end = 0;
            fgets(ligne,79,fp_arc);
            for (z=0; z<strlen(ligne); z++)
            {
                if(ligne[z] == 'D') ver_end = 1; /* La chaine END a ete */
            }
                /* trouvee */
            compteur++;
            if(ver_end == 1)
            {
                id_list[id_compt] = compteur-2;
                id_compt++;
                traites++;
                fgets(ligne,79,fp_arc);
                compteur = 1;
            }
        }
    }
}

```

```

}
rewind( fp_arc );

/* Efface le message scanning file */
close_window( &info );
arrow_cursor();

/* Affiche l'operation en cours */
info = info_window("Building Output File...");
time_cursor();

for( i=0; i < id_compt-1; i++ )
{
    fscanf(fp_arc,"%i\n",&ident);
    if(( ident >= min ) && (ident <= max))
    {
        fprintf(fp_bna, "\">%i\",">
        for( j=0; j < id_list[j]; j++ )
        {
            fscanf(fp_arc,"%f %f\n",&x,&y);
            fprintf(fp_bna, "%f,%f\n",x,y);
        }
        transfert++;
    }
    else
    {
        for( j=0; j < id_list[j]; j++ )
        {
            fscanf(fp_arc,"%f %f\n",&x,&y);
        }
    }
}

fgets(ligne,79,fp_arc);
}

/* Efface le message d'attente */
close_window( &info );
arrow_cursor();

/* Place la couleur de tracage a blanc */
setcolor( WHITE );
hidemousecursor();
outtextxy(462,100,"# Lines In :");
outtextxy(462,115,"# Lines Out:");
itoa(traites-1,limchar,10);
outtextxy(560,100,limchar);

```

```

    itoa(transfert,limchar,10);
    outtextxy(560,115,limchar);

    /* Remettre la couleur de tracage au default */
    setcolor(coul_actuel);
    showmousecursor();

    /* Affiche l'heure de fin de la procedure */
    temps_fin();

    fclose( fp_arc );
    fclose( fp_bna );

}

free(id_list);

}
else
{
/* Probleme d'allocation de memoire */
display_message("Insufficient Memory!");
}
}

/*****
* Cette fonction fait la lecture d'un fichier UNGEN d'ARC/INFO contenant
* des elements lineaires et en filtre les elements selon les bornes
* etablies par les valeurs min et max. Le fichier de sortie est un
* fichier UNGEN d'ARC/INFO ne contenant que les elements issus du filtre.
* Le fichier de sortie ne renferme que des elements lineaires.
*
* Parametres: name_in = nom du fichier d'entree ARC/INFO
*             name_out = nom du fichier de sortie *.ARC
*             lim_min = limite minimum des identificateurs
*             lim_max = limite maximum des identificateurs
*
*
* Valeur possible des identificateurs : -32768 a 32767
* Nombre maximal de coordonnees par arc: 32767
* Nombre maximal d'arcs par fichier : 32000
*
* Denis Leroux, novembre 1993
*****/

```

```

*/

void conv_arc_filt( char *name_in, char *name_out, long min, long max )
{
FILE *fp_arc, *fp_lnk;

char ligne[79];
char limchar[25];
window_type info;
int compteur;
int z;
int ver_end;
int *id_list;
int id_compt;
int transfert, traites;
int i, j;
int ident;
int coul_actuel;
float x, y;

/* On stock en memoire la couleur actuelle de tracage */
coul_actuel = getcolor();

id_list = (int *)malloc(32000);

if( id_list != NULL )
{
if( ( fp_arc = fopen( name_in,"r" ) ) != NULL )
{
/* Affiche l'heure de debut de la procedure */
temps_debut();
fp_lnk = fopen( name_out,"w" );

fgets(ligne,79,fp_arc);
id_compt = 0;
compteur = 1;
transfert = 0;
traites = 0;

/* Affiche l'operation en cours */
info = info_window("Scanning Input File...");
time_cursor();

while( !feof( fp_arc ) )

```

```

{
    ver_end = 0;
    fgets(ligne,79,fp_arc);
    for (z=0; z<strlen(ligne); z++)
    {
        if(ligne[z] == 'D') ver_end = 1; /* La chaine END a ete trouvee */
    }
    compteur++;
    if(ver_end == 1)
    {
        id_list[id_compt] = compteur-2;
        id_compt++;
        traies++;
        fgets(ligne,79,fp_arc);
        compteur = 1;
    }
}
rewind( fp_arc );
/* Efface le message scanning file */
close_window( &info );
arrow_cursor();

/* Affiche l'operation en cours */
info = info_window("Building Output File...");
time_cursor();

for( i=0; i < id_compt-1; i++ )
{
    fscanf(fp_arc,"%i\n",&ident);
    if(( ident >= min ) && ( ident <= max))
    {
        fprintf(fp_lnk," %i\n",ident);
        for( j=0; j < id_list[i]; j++ )
        {
            fscanf(fp_arc,"%f %f\n",&x,&y);
            fprintf(fp_lnk,"%f %f\n",x,y);
        }
        fprintf(fp_lnk,"END\n");
        transfert++;
    }
    else
    {
        for( j=0; j < id_list[i]; j++ )
        {
            fscanf(fp_arc,"%f %f\n",&x,&y);
        }
    }
}

```

```

    }

    fgets(ligne,79,fp_arc);
}
fprintf(fp_Ink,"END\n");

/* Efface le message d'attente */
close_window( &info );
arrow_cursor();

/* Place la couleur de tracage a BLANC */
setcolor( WHITE );
hidemousecursor();

/* Affiche les stats sur le fichier */
outtextxy(462,100,"# Lines In :");
outtextxy(462,115,"# Lines Out:");
itoa(traites-1,limchar,10);
outtextxy(560,100,limchar);
itoa(transfert,limchar,10);
outtextxy(560,115,limchar);

/* Remettre la couleur de tracage au default */
setcolor(coul_actuel);
showmousecursor();

/* Affiche l'heure de fin de la procedure */
temps_fin();

fclose( fp_arc );
fclose( fp_Ink );

}

free(id_list);

} /* Memory allocation */

}

/*****
* La fonction conv_Ink fait la conversion d'un fichier BNA contenant des
* elements lineaires en format LNK pouvant servir aux analyses de reseau
* dans NodArc. Cette fonction necessite a l'entree deux fichiers: un fichier
* BNA contenant les aretes du reseau et un fichier NDS (NodArc) contenant
* les sommets des aretes. La fonction CONV_LNK assignera le nom UNKNOWN
*****/

```

```

* aux noeuds des aretes si ces dernieres contiennent des noeuds non definis      *
* dans le fichier NDS (Unknown node in BNA file). La fonction CONV_LNK          *
* creer egalement, pour chaque noeud un node_id unique et sequentiel de 0 a    *
* n. Les node_id du fichier NDS et les node_id du fichier LNK sont donc        *
* differents. Cependant, les attributs des noeuds du fichier NDS sont          *
* associes correctement aux noeuds du fichier LNK (nom & type de noeud).      *
*                                                                              *
* Denis Leroux, mai 1994                                                    *
*.....*
*/

int conv_lnk( char *name_in_bna, char *name_in_nds, char *name_in_lnk )
{
FILE *fp_bna;
FILE *fp_nds;      /* Pointeur vers le fichier NDS temporaire */
FILE *fp_lnk;
FILE *fp_tmp;      /* Pointeur vers le fichier FROM__TO temporaire */
FILE *fp_nds_org;  /* Pointeur vers le fichier NDS original */
FILE *fp_doc;      /* Pointeur vers le fichier documentation */

char c;
char ident_1[66], ident_2[66], node_name[66], from_node_name[66], to_node_name[66];
char nom_dum[66], ligne_minmax[129], limchar[25], nom_doc[128];
int nbre_pts, ligne, polygone, point, i, arc_id, arc_type, type_dum, nbre_node;
int node_id, node_type, from_node_id, from_node_type, flag_from, coul_actuel;
int to_node_id, to_node_type, flag_to, node_max, node_val, unknown_type;
int noeud_id, trouvee;
float x_coord, y_coord, xmin, xmax, ymin, ymax, xdum, ydum;
float from_xcoord, from_ycoord, to_xcoord, to_ycoord;
window_type info;

/* Initialisation des variables */
xmin = 3400000000000000.0;
ymin = 3400000000000000.0;
xmax = -3400000000000000.0;
ymax = -3400000000000000.0;
point = 0;
ligne = 0;
polygone = 0;
arc_id = 0;
arc_type = 1;
flag_from = 0;
flag_to = 0;
node_max = 0;

```

```

unknown_type = 1;
nbre_node = 0;
noeud_id = 0;
trouvee = 0;

/* On va chercher la valeur de la couleur de tracage */
coul_actuel = getcolor();

if( ((fp_bna = fopen(name_in_bna,"r")) != NULL) && ((fp_nds_org = fopen(name_in_nds,"r")) !=
NULL) )
{
/* Ouvre des fichiers temporaires pour le traitement */
fp_nds = fopen("NDS_LIST.TMP","w+"); /* Fichier temporaire */
fp_tmp = fopen("FROM__TO.TMP","w+"); /* Fichier temporaire */

/* Affichage de l'heure de debut de la procedure */
temps_debut();

/* Message a l'utilisateur */
info = info_window("Scanning BNA file...");
time_cursor();

while( !feof( fp_bna ) )
{
c = fgetc( fp_bna ); /* lire " */
fscanf(fp_bna, "%[^\""]",ident_1); /* Lire ident_1 */
c = fgetc( fp_bna ); /* Lire " */
c = fgetc( fp_bna ); /* Lire , */
c = fgetc( fp_bna ); /* Lire " */
fscanf(fp_bna, "%[^\""]",ident_2); /* Lire ident_2 */
c = fgetc( fp_bna ); /* Lire " */
c = c;
fscanf(fp_bna, "%i\n",&nbre_pts);

if( nbre_pts < 0 ) ligne++;
if( nbre_pts > 2 ) polygone++;
if( nbre_pts == 1 ) point++;

if( nbre_pts < 1 ) /* Element lineaire (arete) */
{
/* Ecriture de l'arc ID et des coords FROM dans FROM__TO.TMP */
fscanf(fp_bna, "%f,%f\n",&x_coord,&y_coord);
fprintf(fp_tmp, "%i %f %f ",arc_id,x_coord,y_coord);
arc_id+=1;

for( i=1; i < abs(nbre_pts); i++ )

```

```

    {
        fscanf(fp_bna,"%f,%f\n",&x_coord,&y_coord);
        if( x_coord < xmin ) xmin = x_coord;
        if( x_coord > xmax ) xmax = x_coord;
        if( y_coord < ymin ) ymin = y_coord;
        if( y_coord > ymax ) ymax = y_coord;
    }
    /* Ecriture des coords TO dans le fichier FROM__TO.TMP */
    fprintf(fp_tmp,"%f %f\n",x_coord,y_coord);

}
if( nbre_pts == 1 ) /* Elomet ponctuel */
{
    fscanf(fp_bna,"%f,%f\n",&x_coord,&y_coord);
}
if( nbre_pts > 2 ) /* Element polygonal */
{
    for( i=0; i < nbre_pts; i++ )
    {
        fscanf(fp_bna,"%f,%f\n",&x_coord,&y_coord);
    }
}
}
rewind(fp_bna);
rewind(fp_tmp);

/* Efface le message Scanning BNA file */
close_window( &info );
arrow_cursor();

/* On affiche le resultat du scan dans File Info */
hidemousecursor();
setcolor( WHITE );
outtextxy(462,110,"BNA content:");
line(462,111,550,111);
outtextxy(462,128,"# Points :");
outtextxy(462,141,"# Lines :");
outtextxy(462,154,"# Polygones:");
itoa(point,limchar,10);
outtextxy(560,128,limchar);
itoa(ligne,limchar,10);
outtextxy(560,141,limchar);
itoa(polygone,limchar,10);
outtextxy(560,154,limchar);
setcolor( coul_actuel );
showmousecursor();

```

```

if( ligne > 0 )
{
    hidemousecursor();
    setcolor( WHITE );
    outtextxy(462,167,"X min:");
    outtextxy(462,180,"Y min:");
    outtextxy(462,193,"X max:");
    outtextxy(462,206,"Y max:");
    sprintf(limchar,"%f",xmin);
    outtextxy(510,167,limchar);
    sprintf(limchar,"%f",ymin);
    outtextxy(510,180,limchar);
    sprintf(limchar,"%f",xmax);
    outtextxy(510,193,limchar);
    sprintf(limchar,"%f",ymax);
    outtextxy(510,206,limchar);
    setcolor( coul_actuel );
    showmousecursor();
}
else
{
    temps_fin();
    display_message("No lines in BNA file! Aborting...");
    fclose(fp_bna);
    fclose(fp_nds);
    fclose(fp_nds_org);
    fclose(fp_tmp);
    return(0);
}

/* Message Scanning NDS file */
info = info_window("Scanning NDS file & Assigning sequential node Id ...");
time_cursor();

/* Scanning du fichier NDS pour determiner la valeur max de node_id */
fgets(ligne_minmax,128,fp_nds_org);
fprintf(fp_nds,"%s",ligne_minmax);
while( !feof(fp_nds_org) )
{
    fscanf(fp_nds_org,"%i %i %f %f",&node_val,&type_dum,&xdum,&y dum);
    fgets(nom_dum,65,fp_nds_org);
    while( !feof(fp_tmp) )
    {
        fscanf(fp_tmp,"%i %f %f %f
%f\n",&arc_id,&from_xcoord,&from_ycoord,&to_xcoord,&to_ycoord);
    }
}

```

```

        if(((xdum == from_xcoord) && (ydum == from_ycoord)) || ((xdum == to_xcoord) &&
(ydum == to_ycoord)))
        {
            if( trouvee != 1 )
            {
                fprintf(fp_nds,"%i %i %f %f %s",noeud_id,type_dum,xdum,ydum,nom_dum);
                trouvee = 1;
                noeud_id +=1;
            }
        }
    }
    rewind(fp_tmp);
    trouvee = 0;

    if( noeud_id > node_max )
    {
        node_max = noeud_id;
    }
    nbre_node+=1;
}
/* Ajout d'un new line aditionnel a la fin du fichier temporaire */
if( noeud_id != 0)
{
    fprintf(fp_nds,"\n");
}

rewind(fp_nds);

/* Efface le message Scanning NDS file */
close_window( &info );
arrow_cursor();

/* Affichage du contenu du fichier NDS */
setcolor( WHITE );
hidemousecursor();
outtextxy(462,235,"NDS content:");
line(462,236,550,236);
outtextxy(462,253,"# Nodes:");
itoa(nbre_node,limchar,10);
outtextxy(530,253,limchar);
setcolor( coul_actuel );
showmousecursor();

/* Ouverture du fichier LNK */
fp_lnk = fopen(name_in_lnk,"w");
fprintf(fp_lnk,"%f %f %f %f\n",xmin,ymin,xmax,ymax);

```

```

/* On indique a l'usager que l'on match les nodes et les arcs */
info = info_window("Matching arcs & nodes. Wait!...");
time_cursor();
while( !feof(fp_bna) )
{
    c = fgetc( fp_bna );          /* lire " */
    fscanf(fp_bna,"%^[^\""]",ident_1); /* Lire ident_1 */
    c = fgetc( fp_bna );          /* Lire " */
    c = fgetc( fp_bna );          /* Lire , */
    c = fgetc( fp_bna );          /* Lire " */
    fscanf(fp_bna,"%^[^\""]",ident_2); /* Lire ident_2 */
    c = fgetc( fp_bna );          /* Lire " */
    c = c;
    fscanf(fp_bna,"%i\n",&nbre_pts);

    if( nbre_pts == 1 )
    {
        fscanf(fp_bna,"%f,%f\n",&x_coord,&y_coord);
    }
    if( nbre_pts > 2 )
    {
        for(i=0; i<nbre_pts; i++)
        {
            fscanf(fp_bna,"%f,%f\n",&x_coord,&y_coord);
        }
    }
    if( nbre_pts < 0 )
    {
        fscanf(fp_tmp,"%i %f %f %f
%f\n",&arc_id,&from_xcoord,&from_ycoord,&to_xcoord,&to_ycoord);
        fprintf(fp_lnk,"%i %i %i %s\n",arc_id,arc_type,abs(nbre_pts),ident_1);

        /* Scanning du fichier NDS pour trouver les info sur les nodes */
        fgets(ligne_minmax,128,fp_nds);
        while( !feof(fp_nds) )
        {
            fscanf(fp_nds,"%i %i %f %f",&node_id,&node_type,&x_coord,&y_coord);
            fgets(node_name,65,fp_nds);
            if( (x_coord == from_xcoord) && (y_coord == from_ycoord) )
            {
                from_node_id = node_id;
                from_node_type = node_type;
                strcpy(from_node_name,node_name);
                from_node_name[strcspn(from_node_name,"\n")] = '\0';
                flag_from = 1;
            }
        }
    }
}

```

```

    if( (x_coord == to_xcoord) && (y_coord == to_ycoord) )
    {
        to_node_id = node_id;
        to_node_type = node_type;
        strcpy(to_node_name,node_name);
        to_node_name[strcspn(to_node_name,"\n")] = '\0';
        flag_to = 1;
    }
}

if(flag_from == 0)
{
    from_node_id = node_max;
    from_node_type = unknown_type;
    strcpy(from_node_name,"UNKNOWN");
    fprintf(fp_nds,"%i %i %f %f
%s\n",from_node_id,from_node_type,from_xcoord,from_ycoord,from_node_name);
    node_max+=1;
}

if(flag_to == 0)
{
    to_node_id = node_max;
    to_node_type = unknown_type;
    strcpy(to_node_name,"UNKNOWN");
    fprintf(fp_nds,"%i %i %f %f
%s\n",to_node_id,to_node_type,to_xcoord,to_ycoord,to_node_name);
    node_max+=1;
}
rewind(fp_nds);
flag_from = 0;
flag_to = 0;

/* Ecrire les nodes de depart et de destination dans LNK */
fprintf(fp_lnk,"%i %i %f %f
%s\n",from_node_id,from_node_type,from_xcoord,from_ycoord,from_node_name);
fprintf(fp_lnk,"%i %i %f %f
%s\n",to_node_id,to_node_type,to_xcoord,to_ycoord,to_node_name);

for(i=0; i< abs(nbre_pts); i++)
{
    fscanf(fp_bna,"%f,%f\n",&x_coord,&y_coord);
    fprintf(fp_lnk,"%f,%f\n",x_coord,y_coord);
}
}
}

```

```

/* Efface le message matching nodes & arcs */
close_window( &info );
arrow_cursor();

/* Affiche le message on construit le fichier documentation */
info = info_window("Writing documentation file. Wait!...");
time_cursor();

/* On ouvre le fichier de documentation */
strcpy(nom_doc,name_in_lnk);
nom_doc[(strlen(nom_doc)-3)] = '\0';
strcat(nom_doc,"DOC");
fp_doc = fopen(nom_doc,"w");
rewind(fp_nds);

/* Lire la premiere ligne avec limites de la carte */
fgets(ligne_minmax,128,fp_nds);

/* Ecriture de l'entete du fichier de documentation */
fprintf(fp_doc,"-----\n");
fprintf(fp_doc,"      NodArc: LNK Documentation File      \n");
fprintf(fp_doc,"-----\n");
fprintf(fp_doc,"Data file: ");
fprintf(fp_doc,"%s\n\n",name_in_lnk);

while( !feof(fp_nds) )
{
    fscanf(fp_nds,"%i %i %f %f",&to_node_id,&to_node_type,&to_xcoord,&to_ycoord);
    fgets(to_node_name,65,fp_nds);
    fprintf(fp_doc,"%4i %s",to_node_id,to_node_name);
}

/* Efface le message writing documentation file */
close_window( &info );
arrow_cursor();

/* Fermeture des fichiers */
fclose(fp_bna);
fclose(fp_nds);
fclose(fp_tmp);
fclose(fp_lnk);
fclose(fp_nds_org);
fclose(fp_doc);

/* On efface les fichiers temporaires */

```

```

unlink("NDS_LIST.TMP");
unlink("FROM__TO.TMP");

/* On affiche le temps de fin de l'operation */
temps_fin();
}
else
{
    display_message("File not found!");
}
return(1);
}

/*****
* La fonction create_lat_file permet de creer un fichier d'attributs a
* partir d'un fichier LNK. Cette fonction fait la lecture du fichier et
* associe un poids par defaut a chaque arc du reseau. L'utilisateur peut
* par la suite modifier ce fichier pour y inclure ses propres donnees.
*
* Denis Leroux, juin 1994
*****/

*/

void create_lat_file()
{

char nom_lnk[128], nom_lat[128], nom_l[13], nom_n[13], chemin[50];
char *ptr;
int coul_actuel, arc_id, arc_type, nbre_pts, from_id, from_type;
int to_id, to_type, nbre_arc, i;
float xminim, yminim, xmaxim, ymaxim, from_xcoord, from_ycoord, to_xcoord, to_ycoord;
float xcoord, ycoord;
double poids;
char poids_str[10], arc_name[66], from_name[66], to_name[66];
window_type info;
FILE *fp_lnk;
FILE *fp_lat;

/* Stockage en memoire de la couleur actuelle */
coul_actuel = getcolor();
nbre_arc = 0;

/* Determination du repertoire courant */
ptr = currentdir();
strcpy(chemin,ptr);
free(ptr);

```

```

strcpy(nom_l,pickfile("*.LNK",0,chemin));
strcpy(nom_lnk,chemin);
strcat(nom_lnk,"\\");
strcat(nom_lnk,nom_l);

if(strlen(nom_l) != 0)
{
  /* On demande a l'utilisateur de specifier un fichier de sortie */
  strcpy(nom_n,"");
  if(getfilename(nom_n))
  {
    if(strcspn(nom_n," ") != 0)
    {
      /* Reconstitue le chemin d'acces au fichier */
      strcpy(nom_lat,chemin);
      strcat(nom_lat,"\\");
      strcat(nom_lat,nom_n,strcspn(nom_n," "));
      strcat(nom_lat,".LAT");
      /* On verifie si le fichier existe deja */
      if(access(nom_lat,0) != 0)
      {
        strcpy(poids_str,"1.0");
        if(get_input_string("Default arc weight:",poids_str))
        {
          poids = atof(poids_str);
          if(poids > 0 )
          {
            /* Affichage des noms de fichiers dans file info */
            efface_file_info();
            setcolor(WHITE);
            hidemousecursor();
            outtextxy(462,55,"Input :");
            outtextxy(525,55,nom_l);
            outtextxy(462,70,"Output:");
            nom_n[strcspn(nom_n," ")] = '\0';
            strcat(nom_n,".LAT");
            outtextxy(525,70,nom_n);
            setcolor( coul_actuel );
            showmousecursor();

            /* Debut de la procedure de creation du LAT */
            temps_debut();

            if(( fp_lnk = fopen(nom_lnk,"r")) != NULL)
            {
              fp_lat = fopen(nom_lat,"w");
            }
          }
        }
      }
    }
  }
}

```

```

/* Message a l'usager */
info = info_window("Processing. Please wait...!");
time_cursor();

fscanf(fp_Ink,"%f %f %f %f\n",&xminim,&yminim,&xmaxim,&ymaxim);

while( !feof(fp_Ink) )
{
    fscanf(fp_Ink,"%i %i %i",&arc_id,&arc_type,&nbre_pts);
    fgets(arc_name,65,fp_Ink);
    fscanf(fp_Ink,"%i %i %f
%f",&from_id,&from_type,&from_xcoord,&from_ycoord);
    fgets(from_name,65,fp_Ink);
    fscanf(fp_Ink,"%i %i %f %f",&to_id,&to_type,&to_xcoord,&to_ycoord);
    fgets(to_name,65,fp_Ink);

    if( arc_type == 1 )
    {
        fprintf(fp_lat,"%f %5i %5i %5i
%s",poids,from_id,to_id,arc_id,arc_name);
        fprintf(fp_lat,"%f %5i %5i %5i
%s",poids,to_id,from_id,arc_id,arc_name);
    }
    if( arc_type == 0 )
    {
        fprintf(fp_lat,"%f %5i %5i %5i
%s",poids,from_id,to_id,arc_id,arc_name);
    }

    for(i=0;i<nbre_pts;i++)
    {
        fscanf(fp_Ink,"%f,%f\n",&xcoord,&ycoord);
    }

    /* On compte les arcs */
    nbre_arc+=1;
}

/* On ferme les fichier */
fclose(fp_Ink);
fclose(fp_lat);

/* On efface le message a l'usager et on affiche le temps */
close_window( &info );
arrow_cursor();

```

```
        temps_fin();
    }
    else
    {
        display_message("Arc weight must be > 0.0 !");
    }
}
else
{
    display_message("File already exist!");
}
else
{
    display_message("Invalid file name!");
}
}
/* Cancel du getfilename */
}
}
```

Annexe 2

Algorithme du plus court chemin

Note: Cette annexe renferme la *version modifiée* de l'algorithme de Dijkstra calculant le plus court chemin simultanément à partir du noeud d'origine et du noeud de destination. C'est à partir de cet algorithme de base que les autres modèles de recherche d'itinéraires ont été développés.

```
/******  
* La fonction SHORT_PATH effectue le calcul du plus court chemin base sur *  
* un fichier de reseau LNK. Il prend a l'entree le nom du fichier LNK ainsi *  
* que les nodes_id des noeuds de depart et d'arrivee. Cette fonction creer *  
* un fichier output (aussi en parametre d'entree) contenant le plus court *  
* itineraire (ou le plus securitaire) (de noeud en noeud) ainsi que la longueur *  
* du trajet. Cette fonction affiche egalement l'itineraire trouvee par un appel a *  
* une fonction d'affichage. *  
* *  
* Denis Leroux, mai 1994 *  
*****  
*/
```

```
int short_path( char *name_in, char *name_out, int pt_dep, int pt_fin )  
{
```

```
int isorce, isink, arc_id, arc_type, nbre_pts, from_id, from_type, to_id, to_type;  
int nbre_arc, i, cpt_lien, dep_found, fin_found, node_id_max, coul_actuel;  
int ipt1, ipt2, n, z, index, mnode, num1, num2, numnod, trouvee, sortie;  
float key1, key2, minln1, minln2, minln3, large, distj, idist, lensum, lnpath;  
int huge *from_arete, huge *to_arete, huge *path, huge *point1, huge *point2;  
float huge *vect_dist, huge *vect_lig, huge *vect_col, huge *len1, huge *len2;  
int huge *stack1, huge *stack2;  
float xminim, yminim, xmaxim, ymaxim, from_xcoord, from_ycoord, to_xcoord;  
float to_ycoord, xcoord_1, ycoord_1, xcoord_2, ycoord_2;  
char arc_name[66], from_name[66], to_name[66], dep_nom[66], fin_nom[66];  
char limchar[25];  
double long_arc;  
window_type info;  
FILE *fp_lnk;  
FILE *fp_out;
```

```

FILE *fp_id;

/* Initialisation des variables */
nbre_pts = 0;
nbre_arc = 0;
cpt_lien = 0;
dep_found = 0;
fin_found = 0;
node_id_max = 0;
large = 34000000000000.0;
key1 = 0.0;
key2 = 0.0;
isorce = pt_dep;
isink = pt_fin;

/* Allocation de memoire pour les vecteurs */
from_arete = (int huge *) farcalloc(8000,sizeof(int));
to_arete = (int huge *) farcalloc(8000,sizeof(int));
vect_dist = (float huge *) farcalloc(8000,sizeof(float));
vect_lig = (float huge *) farcalloc(2000,sizeof(float));
vect_col = (float huge *) farcalloc(2000,sizeof(float));
path = (int huge *) farcalloc(2000,sizeof(int));
point1 = (int huge *) farcalloc(2000,sizeof(int));
point2 = (int huge *) farcalloc(2000,sizeof(int));
len1 = (float huge *) farcalloc(2000,sizeof(float));
len2 = (float huge *) farcalloc(2000,sizeof(float));
stack1 = (int huge *) farcalloc(2000,sizeof(int));
stack2 = (int huge *) farcalloc(2000,sizeof(int));

/* verification de la memoire */
if( from_arete == NULL )
{
    display_message("Not enough memory. Aborting...");
    if( from_arete != NULL ) farfree(from_arete);
    if( to_arete != NULL ) farfree(to_arete);
    if( vect_dist != NULL ) farfree(vect_dist);
    if( vect_lig != NULL ) farfree(vect_lig);
    if( vect_col != NULL ) farfree(vect_col);
    if( path != NULL ) farfree(path);
    if( point1 != NULL ) farfree(point1);
    if( point2 != NULL ) farfree(point2);
    if( len1 != NULL ) farfree(len1);
    if( len2 != NULL ) farfree(len2);
    if( stack1 != NULL ) farfree(stack1);
    if( stack2 != NULL ) farfree(stack2);
}

```

```

    return(0);
}

if( to_arete == NULL )
{
    display_message("Not enough memory. Aborting...");
    if( from_arete != NULL ) farfree(from_arete);
    if( to_arete != NULL ) farfree(to_arete);
    if( vect_dist != NULL ) farfree(vect_dist);
    if( vect_lig != NULL ) farfree(vect_lig);
    if( vect_col != NULL ) farfree(vect_col);
    if( path != NULL ) farfree(path);
    if( point1 != NULL ) farfree(point1);
    if( point2 != NULL ) farfree(point2);
    if( len1 != NULL ) farfree(len1);
    if( len2 != NULL ) farfree(len2);
    if( stack1 != NULL ) farfree(stack1);
    if( stack2 != NULL ) farfree(stack2);
    return(0);
}

if( vect_dist == NULL )
{
    display_message("Not enough memory. Aborting...");
    if( from_arete != NULL ) farfree(from_arete);
    if( to_arete != NULL ) farfree(to_arete);
    if( vect_dist != NULL ) farfree(vect_dist);
    if( vect_lig != NULL ) farfree(vect_lig);
    if( vect_col != NULL ) farfree(vect_col);
    if( path != NULL ) farfree(path);
    if( point1 != NULL ) farfree(point1);
    if( point2 != NULL ) farfree(point2);
    if( len1 != NULL ) farfree(len1);
    if( len2 != NULL ) farfree(len2);
    if( stack1 != NULL ) farfree(stack1);
    if( stack2 != NULL ) farfree(stack2);
    return(0);
}

if( vect_lig == NULL )
{
    display_message("Not enough memory. Aborting...");
    if( from_arete != NULL ) farfree(from_arete);
    if( to_arete != NULL ) farfree(to_arete);
    if( vect_dist != NULL ) farfree(vect_dist);
    if( vect_lig != NULL ) farfree(vect_lig);
    if( vect_col != NULL ) farfree(vect_col);
}

```

```

if( path != NULL ) farfree(path);
if( point1 != NULL ) farfree(point1);
if( point2 != NULL ) farfree(point2);
if( len1 != NULL ) farfree(len1);
if( len2 != NULL ) farfree(len2);
if( stack1 != NULL ) farfree(stack1);
if( stack2 != NULL ) farfree(stack2);
return(0);
}
if( vect_col == NULL )
{
display_message("Not enough memory. Aborting...");
if( from_arete != NULL ) farfree(from_arete);
if( to_arete != NULL ) farfree(to_arete);
if( vect_dist != NULL ) farfree(vect_dist);
if( vect_lig != NULL ) farfree(vect_lig);
if( vect_col != NULL ) farfree(vect_col);
if( path != NULL ) farfree(path);
if( point1 != NULL ) farfree(point1);
if( point2 != NULL ) farfree(point2);
if( len1 != NULL ) farfree(len1);
if( len2 != NULL ) farfree(len2);
if( stack1 != NULL ) farfree(stack1);
if( stack2 != NULL ) farfree(stack2);
return(0);
}

if( path == NULL )
{
display_message("Not enough memory. Aborting...");
if( from_arete != NULL ) farfree(from_arete);
if( to_arete != NULL ) farfree(to_arete);
if( vect_dist != NULL ) farfree(vect_dist);
if( vect_lig != NULL ) farfree(vect_lig);
if( vect_col != NULL ) farfree(vect_col);
if( path != NULL ) farfree(path);
if( point1 != NULL ) farfree(point1);
if( point2 != NULL ) farfree(point2);
if( len1 != NULL ) farfree(len1);
if( len2 != NULL ) farfree(len2);
if( stack1 != NULL ) farfree(stack1);
if( stack2 != NULL ) farfree(stack2);
return(0);
}

if( point1 == NULL )

```

```

{
  display_message("Not enough memory. Aborting...");
  if( from_arete != NULL ) farfree(from_arete);
  if( to_arete != NULL ) farfree(to_arete);
  if( vect_dist != NULL ) farfree(vect_dist);
  if( vect_lig != NULL ) farfree(vect_lig);
  if( vect_col != NULL ) farfree(vect_col);
  if( path != NULL ) farfree(path);
  if( point1 != NULL ) farfree(point1);
  if( point2 != NULL ) farfree(point2);
  if( len1 != NULL ) farfree(len1);
  if( len2 != NULL ) farfree(len2);
  if( stack1 != NULL ) farfree(stack1);
  if( stack2 != NULL ) farfree(stack2);
  return(0);
}

if( point2 == NULL )
{
  display_message("Not enough memory. Aborting...");
  if( from_arete != NULL ) farfree(from_arete);
  if( to_arete != NULL ) farfree(to_arete);
  if( vect_dist != NULL ) farfree(vect_dist);
  if( vect_lig != NULL ) farfree(vect_lig);
  if( vect_col != NULL ) farfree(vect_col);
  if( path != NULL ) farfree(path);
  if( point1 != NULL ) farfree(point1);
  if( point2 != NULL ) farfree(point2);
  if( len1 != NULL ) farfree(len1);
  if( len2 != NULL ) farfree(len2);
  if( stack1 != NULL ) farfree(stack1);
  if( stack2 != NULL ) farfree(stack2);
  return(0);
}

if( len1 == NULL )
{
  display_message("Not enough memory. Aborting...");
  if( from_arete != NULL ) farfree(from_arete);
  if( to_arete != NULL ) farfree(to_arete);
  if( vect_dist != NULL ) farfree(vect_dist);
  if( vect_lig != NULL ) farfree(vect_lig);
  if( vect_col != NULL ) farfree(vect_col);
  if( path != NULL ) farfree(path);
  if( point1 != NULL ) farfree(point1);
  if( point2 != NULL ) farfree(point2);
  if( len1 != NULL ) farfree(len1);
}

```

```

    if( len2 != NULL ) farfree(len2);
    if( stack1 != NULL ) farfree(stack1);
    if( stack2 != NULL ) farfree(stack2);
    return(0);
}

if( len2 == NULL )
{
    display_message("Not enough memory. Aborting...");
    if( from_arete != NULL ) farfree(from_arete);
    if( to_arete != NULL ) farfree(to_arete);
    if( vect_dist != NULL ) farfree(vect_dist);
    if( vect_lig != NULL ) farfree(vect_lig);
    if( vect_col != NULL ) farfree(vect_col);
    if( path != NULL ) farfree(path);
    if( point1 != NULL ) farfree(point1);
    if( point2 != NULL ) farfree(point2);
    if( len1 != NULL ) farfree(len1);
    if( len2 != NULL ) farfree(len2);
    if( stack1 != NULL ) farfree(stack1);
    if( stack2 != NULL ) farfree(stack2);
    return(0);
}

if( stack1 == NULL )
{
    display_message("Not enough memory. Aborting...");
    if( from_arete != NULL ) farfree(from_arete);
    if( to_arete != NULL ) farfree(to_arete);
    if( vect_dist != NULL ) farfree(vect_dist);
    if( vect_lig != NULL ) farfree(vect_lig);
    if( vect_col != NULL ) farfree(vect_col);
    if( path != NULL ) farfree(path);
    if( point1 != NULL ) farfree(point1);
    if( point2 != NULL ) farfree(point2);
    if( len1 != NULL ) farfree(len1);
    if( len2 != NULL ) farfree(len2);
    if( stack1 != NULL ) farfree(stack1);
    if( stack2 != NULL ) farfree(stack2);
    return(0);
}

if( stack2 == NULL )
{
    display_message("Not enough memory. Aborting...");
    if( from_arete != NULL ) farfree(from_arete);
    if( to_arete != NULL ) farfree(to_arete);
}

```

```

if( vect_dist != NULL ) farfree(vect_dist);
if( vect_lig != NULL ) farfree(vect_lig);
if( vect_col != NULL ) farfree(vect_col);
if( path != NULL ) farfree(path);
if( point1 != NULL ) farfree(point1);
if( point2 != NULL ) farfree(point2);
if( len1 != NULL ) farfree(len1);
if( len2 != NULL ) farfree(len2);
if( stack1 != NULL ) farfree(stack1);
if( stack2 != NULL ) farfree(stack2);
return(0);
}
/* Affichage de l'heure du debut de la procedure */
temps_debut();

if(( fp_lnk = fopen(name_in,"r")) != NULL)
{
  /* Message a l'utilisateur */
  info = info_window("Scanning LNK file...");
  time_cursor();

  /* On scanne le fichier LNK pour verification les parametres */
  fscanf(fp_lnk,"%f %f %f %f\n",&xminim,&yminim,&xmaxim,&ymaxim);

  while( !feof(fp_lnk) )
  {
    fscanf(fp_lnk,"%i %i %i",&arc_id,&arc_type,&nbre_pts);
    fgets(arc_name,65,fp_lnk);
    fscanf(fp_lnk,"%i %i %f %f",&from_id,&from_type,&from_xcoord,&from_ycoord);
    fgets(from_name,65,fp_lnk);
    fscanf(fp_lnk,"%i %i %f %f",&to_id,&to_type,&to_xcoord,&to_ycoord);
    fgets(to_name,65,fp_lnk);

    if( from_id == pt_dep )
    {
      strcpy(dep_nom,from_name);
      dep_found = 1;
    }
    if( to_id == pt_dep )
    {
      strcpy(dep_nom,to_name);
      dep_found = 1;
    }
    if( from_id == pt_fin )
    {
      strcpy(fin_nom,from_name);

```

```

        fin_found = 1;
    }
    if( to_id == pt_fin )
    {
        strcpy(fin_nom,to_name);
        fin_found = 1;
    }

    if( from_id > node_id_max )
    {
        node_id_max = from_id;
    }
    if( to_id > node_id_max )
    {
        node_id_max = to_id;
    }

    for( i=0; i < nbre_pts; i++ )
    {
        fscanf(fp_lnk,"%f,%fn",&xcoord_2,&ycoord_2);
    }
    /* On compte les arcs */
    nbre_arc+=1;
}
rewind(fp_lnk);

/* Efface le message a l'usager */
close_window( &info );
arrow_cursor();

/* On test les parametres obtenus lors du scan */
if( dep_found == 0 )
{
    display_message("Starting node not found! Aborting...");
    farfree(from_arete);
    farfree(to_arete);
    farfree(vect_dist);
    farfree(vect_lig);
    farfree(vect_col);
    farfree(path);
    farfree(point1);
    farfree(point2);
    farfree(len1);
    farfree(len2);
    farfree(stack1);
    farfree(stack2);
}

```

```

    return(0);
}
if( fin_found == 0 )
{
    display_message("Ending node not found! Aborting...");
    farfree(from_arete);
    farfree(to_arete);
    farfree(vect_dist);
    farfree(vect_lig);
    farfree(vect_col);
    farfree(path);
    farfree(point1);
    farfree(point2);
    farfree(len1);
    farfree(len2);
    farfree(stack1);
    farfree(stack2);
    return(0);
}
if( node_id_max > 1999 )
{
    display_message("Node ID greater than 1999 or more than 2000 nodes. Aborting...");
    farfree(from_arete);
    farfree(to_arete);
    farfree(vect_dist);
    farfree(vect_lig);
    farfree(vect_col);
    farfree(path);
    farfree(point1);
    farfree(point2);
    farfree(len1);
    farfree(len2);
    farfree(stack1);
    farfree(stack2);
    return(0);
}
/* On donne a n la valeur de node_id_max plus 1 */
n = node_id_max + 1;

/* On affiche a l'ecran le nombre de noeuds et d'arcs */
coul_actuel = getcolor();
setcolor( WHITE );
hidemousecursor();
outtextxy(462,100,"# Nodes:");
outtextxy(462,113,"# Arcs :");
itoa(n,limchar,10);

```

```

outtextxy(535,100,limchar);
itoa(nbre_arc,limchar,10);
outtextxy(535,113,limchar);
setcolor( coul_actuel );
showmousecursor();

/* Ouverture du fichier output et ecriture de l'entete du fichier */
fp_out = fopen(name_out,"w");
fprintf(fp_out,"-----\n");
fprintf(fp_out,"      Output from NodArc: SHORTEST PATH ANALYSIS  \n");
fprintf(fp_out,"-----\n\n");
fprintf(fp_out,"Data File Name: %s\n\n",name_in);
fprintf(fp_out,"Number of arcs in data file : %i\n",nbre_arc);
fprintf(fp_out,"Number of nodes in data file: %i\n\n",n);
fprintf(fp_out,"Requested shortest path:\n");
fprintf(fp_out,"      From: (%i) %s",pt_dep,dep_nom);
fprintf(fp_out,"      To: (%i) %s\n",pt_fin,fin_nom);

/* On ouvre un fichier temporaire pour le stockage des nodes_name */
fp_id = fopen("NODE_ID.TMP","w+");

/* Message a l'utilisateur */
info = info_window("Creating adjacency & distance matrices...");
time_cursor();

/* On creer les vecteurs d'adjacence et de distance */
fscanf(fp_Ink,"%f %f %f %f\n",&xminim,&yminim,&xmaxim,&ymaxim);

/* Initialise a zero la distance de chaque noeud avec lui-meme */
for(i=0;i<n;i++)
{
  from_arete[i] = i;
  to_arete[i] = i;
  vect_dist[i] = 0.0;
}
cpt_lien = n;

while( !feof(fp_Ink) )
{
  long_arc = 0;
  fscanf(fp_Ink,"%i %i %i",&arc_id,&arc_type,&nbre_pts);
  fgets(arc_name,65,fp_Ink);
  fscanf(fp_Ink,"%i %i %f %f",&from_id,&from_type,&from_xcoord,&from_ycoord);
  fgets(from_name,65,fp_Ink);
  fscanf(fp_Ink,"%i %i %f %f",&to_id,&to_type,&to_xcoord,&to_ycoord);
  fgets(to_name,65,fp_Ink);
}

```

```

/* On ecrit dans le fichier temporaire le nom de chaque noeud */
fprintf(fp_id,"%i %s",from_id,from_name);
fprintf(fp_id,"%i %s",to_id,to_name);

/* Lecture des vertex de l'arc et calcul de la distance*/
fscanf(fp_lnk,"%f,%fn",&xcoord_1,&ycoord_1);

for( i=1; i < nbre_pts; i++ )
{
    fscanf(fp_lnk,"%f,%fn",&xcoord_2,&ycoord_2);
    long_arc = long_arc + sqrtl(((xcoord_1 - xcoord_2) * (xcoord_1 - xcoord_2)) +
    ((ycoord_1 - ycoord_2) * (ycoord_1 - ycoord_2)));
    xcoord_1 = xcoord_2;
    ycoord_1 = ycoord_2;
}

/* On compte les arcs */
nbre_arc+=1;

/* On batie le vecteur d'adjacence */
if(arc_type == 1)
{
    from_arete[cpt_lien] = from_id;
    to_arete[cpt_lien] = to_id;
    vect_dist[cpt_lien] = long_arc;
    cpt_lien++;
    from_arete[cpt_lien] = to_id;
    to_arete[cpt_lien] = from_id;
    vect_dist[cpt_lien] = long_arc;
    cpt_lien++;
}
if(arc_type == 0)
{
    from_arete[cpt_lien] = from_id;
    to_arete[cpt_lien] = to_id;
    vect_dist[cpt_lien] = long_arc;
    cpt_lien++;
}
}

/* On replace au debut du fichier le pointeur de fp_id */
rewind(fp_id);

/* Efface le message a l'usager */
close_window( &info );

```

```

    arrow_cursor();

    /***** Debut de la routine de calcul du plus court chemin *****/

    /* Message a l'utilisateur */
    info = info_window("Finding shortest path. Please wait...");
    time_cursor();

    /* On emplis les vecteur ligne et les vecteurs colonnes de valeur infinie */
    for(i=0; i<n; i++)
    {
        vect_lig[i] = large;
        vect_col[i] = large;
    }

    for(i=0; i<cpt_lien; i++)
    {
        if(from_arete[i] == isorce)
        {
            vect_lig[to_arete[i]] = vect_dist[i];
        }
        if(to_arete[i] == isink)
        {
            vect_col[from_arete[i]] = vect_dist[i];
        }
    }

    for(i=0; i<n; i++)
    {
        len1[i] = vect_lig[i];
        len2[i] = vect_col[i];
        point1[i] = isorce;
        point2[i] = isink;
    }

    ipt1 = 0;
    ipt2 = 0;
    minln2 = large;
    minln1 = large;

    for(i=0; i<n; i++)
    {
        distj = len1[i];
        if(distj > key1)
        {
            if(distj < minln1)

```

```

    {
        ipt1 = 0;
        minln1 = distj;
        stack1[ipt1] = i;
    }
    else
    {
        if(distj == minln1)
        {
            ipt1 = ipt1 + 1;
            stack1[ipt1] = i;
        }
    }
}

distj = len2[i];
if(distj > key2)
{
    if(distj < minln2)
    {
        ipt2 = 0;
        minln2 = distj;
        stack2[ipt2] = i;
    }
    else
    {
        if(distj == minln2)
        {
            ipt2 = ipt2 + 1;
            stack2[ipt2] = i;
        }
    }
}
}

do
{
    if(minln1 <= minln2)
    {
        key1 = minln1;
        while(ipt1 >= 0)
        {
            index = stack1[ipt1];

            /* On emplit le vecteur ligne de valeur infinie */
            for(i=0; i<n; i++)

```

```

{
  vect_lig[i] = large;
}

for(i=0; i<cpt_lien; i++)
{
  if(from_arete[i] == index)
  {
    vect_lig[to_arete[i]] = vect_dist[i];
  }
}

for(i=0; i<n; i++)
{
  idist = vect_lig[i];
  lensum = minln1 + idist;
  if(len1[i] > lensum)
  {
    len1[i] = lensum;
    point1[i] = index;
  }
}
ipt1 = ipt1 - 1;
}
minln1 = large;
ipt1 = 0;

for(i=0; i<n; i++)
{
  distj = len1[i];
  if(distj > key1)
  {
    if(distj < minln1)
    {
      ipt1 = 0;
      minln1 = distj;
      stack1[ipt1] = i;
    }
    else
    {
      if(distj == minln1)
      {
        ipt1 = ipt1 + 1;
        stack1[ipt1] = i;
      }
    }
  }
}

```

```

    }
}
else
{
    key2 = minln2;
    while(ipt2 >= 0)
    {
        index = stack2[ipt2];

        /* On emplis le vecteur colonnes de valeur infinie */
        for(i=0; i<n; i++)
        {
            vect_col[i] = large;
        }

        for(i=0; i<cpt_lien; i++)
        {
            if(to_arete[i] == index)
            {
                vect_col[from_arete[i]] = vect_dist[i];
            }
        }

        for(i=0; i<n; i++)
        {
            idist = vect_col[i];
            lensum = minln2 + idist;
            if(len2[i] > lensum)
            {
                len2[i] = lensum;
                point2[i] = index;
            }
        }
        ipt2 = ipt2 - 1;
    }
    minln2 = large;
    ipt2 = 0;
    for(i=0; i<n; i++)
    {
        distj = len2[i];
        if(distj > key2)
        {
            if(distj < minln2)
            {
                ipt2 = 0;
            }
        }
    }
}

```

```

        minln2 = distj;
        stack2[ipt2] = i;
    }
    else
    {
        if(distj == minln2)
        {
            ipt2 = ipt2 + 1;
            stack2[ipt2] = i;
        }
    }
}
}
}
minln3 = large;
for(i=0; i<n; i++)
{
    lensum = len1[i] + len2[i];
    if(lensum < minln3)
    {
        minln3 = lensum;
        mnode = i;
    }
}
} while(minln3 > (minln1 + minln2)); /* Fin du goto 30 */

num1 = mnode;
path[node_id_max] = mnode;
if(mnode != isorce)
{
    numnod = node_id_max - 1;
    sortie = 0;
    while( sortie == 0 )
    {
        num2 = point1[num1];
        if(num2 != isorce)
        {
            num1 = num2;
            path[numnod] = num2;
            numnod = numnod - 1;
        }
        else
        {
            sortie = 1;
        }
    }
}
}

```

```

}
else
{
    numnod = node_id_max;
}

path[0] = isorce;
num1 = numnod + 1;
numnod = 1;

while(num1 <= node_id_max)
{
    path[numnod] = path[num1];
    numnod = numnod + 1;
    num1 = num1 + 1;
}

if(mnode != isink)
{
    num1 = mnode;
    sortie = 0;
    while( sortie == 0 )
    {
        num2 = point2[num1];
        if(num2 != isink)
        {
            num1 = num2;
            path[numnod] = num2;
            numnod = numnod + 1;
        }
        else
        {
            sortie = 1;
        }
    }
    path[numnod] = isink;
}

Inpath = len1[mnode] + len2[mnode];

if( Inpath < large )
{
    fprintf(fp_out,"Length of shortest path = %.3f\n\n",Inpath);

    fprintf(fp_out,"Routing table:\n");
    fprintf(fp_out,"_____ \n\n");
}

```

```

z = 0;
fprintf(fp_out,"From: (%i) %s",path[z],dep_nom);

if(isorce != isink)
{
do
{
z++;
trouvee = 0;
while( trouvee == 0 )
{
fscanf(fp_id,"%i",&to_id);
fgets(to_name,65,fp_id);
if( to_id == path[z])
{
fprintf(fp_out," To: (%i) %s",path[z],to_name);
trouvee = 1;
rewind(fp_id);
}
}
} while(path[z] != isink);
}
}
else
{
fprintf(fp_out,"NO POSSIBLE PATH!\n");
}

/* Efface le message a l'usager */
close_window( &info );
arrow_cursor();

/***** Fin de la routine de calcul de plus court chemin *****/

fclose(fp_lnk);
fclose(fp_out);
fclose(fp_id);
unlink("NODE_ID.TMP");

}
else
{
display_message("File not found!");
}

farfree(from_arete);

```

```
farfree(to_arete);
farfree(vect_dist);
farfree(vect_lig);
farfree(vect_col);
farfree(point1);
farfree(point2);
farfree(len1);
farfree(len2);
farfree(stack1);
farfree(stack2);

/* Affiche le temps de fin de l'operation */
temps_fin();

/* Appel de la routine de tracage du plus court chemin */
if( lnpath < large )
{
    voir_court_chemin( name_in, path, isink);
}

farfree(path);

/* Indique a l'utilisateur d'utiliser l'option Display output file */
display_message("Use DISPLAY OUTPUT FILE for Results");

return(0);
}
```