



**HAL**  
open science

# Quelques outils infographiques pour l'analyse structurale des systèmes

Michel Delarche

► **To cite this version:**

Michel Delarche. Quelques outils infographiques pour l'analyse structurale des systèmes. Modélisation et simulation. Université Joseph-Fourier - Grenoble I; Institut National Polytechnique de Grenoble - INPG, 1979. Français. NNT : . tel-00289234

**HAL Id: tel-00289234**

**<https://theses.hal.science/tel-00289234>**

Submitted on 20 Jun 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE

*présentée à*

**Université Scientifique et Médicale de Grenoble  
Institut National Polytechnique de Grenoble**

*pour obtenir le grade de*  
**DOCTEUR - INGENIEUR**  
Informatique

*par*

**DELARCHE Michel**



**QUELQUES OUTILS INFOGRAPHIQUES POUR  
L'ANALYSE STRUCTURALE DE SYSTEMES.**



**Thèse soutenue le 29 Juin 1979 devant la commission d'examen**

<b>G. VEILLON</b>	<b>Président</b>
<b>C. BENZAKEN</b>	
<b>M. LUCAS</b>	<b>Examineurs</b>
<b>B. MUNIER</b>	
<b>J.M. TREILLE</b>	
<b>F. CHATELIN</b>	<b>Rapporteur</b>



# UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

Monsieur Gabriel CAU : Président

Monsieur Joseph KLEIN : Vice-Président

---

## MEMBRES DU CORPS ENSEIGNANT DE L'U.S.M.G.

### PROFESSEURS TITULAIRES

MM.	AMBLARD Pierre	Clinique de dermatologie
	ARNAUD Paul	Chimie
	ARVIEU Robert	I.S.N.
	AUBERT Guy	Physique
	AYANT Yves	Physique approfondie
Mme	BARBIER Marie-Jeanne	Electrochimie
MM.	BARBIER Jean-Claude	Physique expérimentale
	BARBIER Reynold	Géologie appliquée
	BARJON Robert	Physique nucléaire
	BARNOUD Fernand	Biosynthèse de la cellulose
	BARRA Jean-René	Statistiques
	BARRIE Joseph	Clinique chirurgicale A
	BEAUDOING André	Clinique de pédiatrie et puériculture
	BELORIZKY Elie	Physique
	BARNARD Alain	Mathématiques pures
Mme	BERTRANDIAS Françoise	Mathématiques pures
MM.	BERTRANDIAS Jean-Paul	Mathématiques pures
	BEZES Henri	Clinique chirurgicale et traumatologie
	BLAMBERT Maurice	Mathématiques pures
	BOLLIET Louis	Informatique (I.U.T. B)
	BONNET Jean-Louis	Clinique ophtalmologie
	BONNET-EYMARD Joseph	Clinique hépto-gastro-entérologie
Mme	BONNIER Marie-Jeanne	Chimie générale
MM.	BOUCHERLE André	Chimie et toxicologie
	BOUCHEZ Robert	Physique nucléaire
	BOUSSARD Jean-Claude	Mathématiques appliquées
	BOUTET DE MONVEL Louis	Mathématiques pures
	BRAVARD Yves	Géographie
	CABANEL Guy	Clinique rhumatologique et hydrologique
	CALAS François	Anatomie
	CARLIER Georges	Biologie végétale
	CARRAZ Gilbert	Biologie animale et pharmacodynamie

.../...

MM.	CAU Gabriel	Médecine légale et toxicologie
	CAUQUIS Georges	Chimie organique
	CHABAUTY Claude	Mathématiques pures
	CHARACHON Robert	Clinique ot-rhino-laryngologique
	CHATEAU Robert	Clinique de neurologie
	CHIBON Pierre	Biologie animale
	COEUR André	Pharmacie chimique et chimie analytique
	COUDERC Pierre	Anatomie pathologique
	DEBELMAS Jacques	Géologie générale
	DEGRANGE Charles	Zoologie
	DELORMAS Pierre	Pneumophtisiologie
	DEPORTES Charles	Chimie minérale
	DESRE Pierre	Métallurgie
	DODU Jacques	Mécanique appliquée (I.U.T. I)
	DOLIQUE Jean-Michel	Physique des plasmas
	DREYFUS Bernard	Thermodynamique
	DUCROS Pierre	Cristallographie
	FONTAINE Jean-Marc	Mathématiques pures
	GAGNAIRE Didier	Chimie physique
	GALVANI Octave	Mathématiques pures
	GASTINEL Noël	Analyse numérique
	GAVEND Michel	Pharmacologie
	GEINDRE Michel	Electroradiologie
	GERBER Robert	Mathématiques pures
	GERMAIN Jean-Pierre	Mécanique
	GIRAUD Pierre	Géologie
	JANIN Bernard	Géographie
	KAHANE André	Physique générale
	KLEIN Joseph	Mathématiques pures
	KOSZUL Jean-Louis	Mathématiques pures
	KRAVTCHENKO Julien	Mécanique
	LACAZE Albert	Thermodynamique
	LACHARME Jean	Biologie végétale
Mme	LAJZEROWICZ Janine	Physique
MM.	LAJZEROWICZ Joseph	Physique
	LATREILLE René	Chirurgie générale
	LATURAZE Jean	Biochimie pharmaceutique
	LAURENT Pierre	Mathématiques appliquées
	LEDRU Jean	Clinique médicale B
	LE ROY Philippe	Mécanique (I.U.T. I)

MM.	LLIBOUTRY Louis	Géophysique
	LOISEAUX Jean-Marie	Sciences nucléaires
	LONGQUEUE Jean-Pierre	Physique nucléaire
	LOUP Jean	Géographie
Mlle	LUTZ Elisabeth	Mathématiques pures
MM.	MALINAS Yves	Clinique obstétricale
	MARTIN-NOEL Pierre	Clinique cardiologique
	MAYNARD Roger	Physique du solide
	MAZARE Yves	Clinique Médicale A
	MICHEL Robert	Minéralogie et pétrographie
	MICOUD Max	Clinique maladies infectieuses
	MOURIQUAND Claude	Histologie
	MOUSSA André	Chimie nucléaire
	NEGRE Robert	Mécanique
	NOZIERES Philippe	Spectrométrie physique
	OZENDA Paul	Botanique
	PAYAN Jean-Jacques	Mathématiques pures
	PEBAY-PEYROULA Jean-Claude	Physique
	PERRET Jean	Séméiologie médicale (neurologie)
	RASSAT André	Chimie systématique
	RENARD Michel	Thermodynamique
	REVOL Michel	Urologie
	RINALDI Renaud	Physique
	DE ROUGEMONT Jacques	Neuro-Chirurgie
	SARRAZIN Roger	Clinique chirurgicale B
	SEIGNEURIN Raymond	Microbiologie et hygiène
	SENGEL Philippe	Zoologie
	SIBILLE Robert	Construction mécanique (I.U.T. I)
	SOUTIF Michel	Physique générale
	TANCHE Maurice	Physiologie
	VAILLANT François	Zoologie
	VALENTIN Jacques	Physique nucléaire
Mme	VERAIN Alice	Pharmacie galénique
MM.	VERAIN André	Physique biophysique
	VEYRET Paul	Géographie
	VIGNAIS Pierre	Biochimie médicale

**PROFESSEURS ASSOCIES**

MM.	CRABBE Pierre	CERMO
	SUNIER Jules	Physique

**PROFESSEURS SANS CHAIRE**

Mlle	AGNIUS-DELORS Claudine	Physique pharmaceutique
	ALARY Josette	Chimie analytique
MM.	AMBROISE-THOMAS Pierre	Parasitologie
	ARMAND Gilbert	Géographie
	BENZAKEN Claude	Mathématiques appliquées
	BIAREZ Jean-Pierre	Mécanique
	BILLET Jean	Géographie
	BOUCHET Yves	Anatomie
	BRUGEL Lucien	Energétique (I.U.T. I)
	BUISSON René	Physique (I.U.T. I)
	BUTEL Jean	Orthopédie
	COHEN-ADDAD Jean-Pierre	Spectrométrie physique
	COLOMB Maurice	Biochimie médicale
	CONTE René	Physique (I.U.T. I)
	DELOBEL Claude	M.I.A.G.
	DEPASSEL Roger	Mécanique des fluides
	GAUTRON René	Chimie
	GIDON Paul	Géologie et minéralogie
	GLENAT René	Chimie organique
	GROULADE Joseph	Biochimie médicale
	HACQUES Gérard	Calcul numérique
	HOLLARD Daniel	Hématologie
	HUGONOT Robert	Hygiène et médecine préventive
	IDELMAN Simon	Physiologie animale
	JOLY Jean-René	Mathématiques pures
	JULLIEN Pierre	Mathématiques appliquées
Mme	KAHANE Josette	Physique
MM.	KRAKOWIACK Sacha	Mathématiques appliquées
	KUHN Gérard	Physique (I.U.T. I)
	LUU DUC Cuong	Chimie organique - pharmacie
	MICHOULIER Jean	Physique (I.U.T. I)
Mme	MINIER Colette	Physique (I.U.T. I)

MM.	PELMONT Jean	Biochimie
	PERRIAUX Jean-Jacques	Géologie et minéralogie
	PFISTER Jean-Claude	Physique du solide
Mlle	PIERY Yvette	Physiologie animale
MM.	RAYNAUD Hervé	M.I.A.G.
	REBECQ Jacques	Biologie (CUS)
	REYMOND Jean-Charles	Chirurgie générale
	RICHARD Lucien	Biologie végétale
Mme	RINAUDO Marguerite	Chimie macromoléculaire
MM.	SARROT-REYNAULD Jean	Géologie
	SIROT Louis	Chirurgie générale
Mme	SOUTIF Jeanne	Physique générale
MM.	STIEGLITZ Paul	Anesthésiologie
	VIALON Pierre	Géologie
	VAN CUTSEM Bernard	Mathématiques appliquées

#### MAITRES DE CONFERENCES ET MAITRES DE CONFERENCES AGREGES

MM.	ARMAND Yves	Chimie (I.U.T. I)
	BACHELOT Yvan	Endocrinologie
	BARGE Michel	Neuro-chirurgie
	BEGUIN Claude	Chimie organique
Mme	BERIEL Hélène	Pharmacodynamie
MM.	BOST Michel	Pédiatrie
	BOUCHARLAT Jacques	Psychiatrie adultes
Mme	BOUCHE Liane	Mathématiques (CUS)
MM.	BRODEAU François	Mathématiques (I.U.T. B) (Personne étrangère habilitée à être directeur de thèse)
	BERNARD Pierre	Gynécologie
	CHAMBAZ Edmond	Biochimie médicale
	CHAMPETIER Jean	Anatomie et organogénèse
	CHARDON Michel	Géographie
	CHERADAME Hervé	Chimie papetière
	CHIAVERINA Jean	Biologie appliquée (EFP)
	COLIN DE VERDIERE Yves	Mathématiques pures
	CONTAMIN Charles	Chirurgie thoracique et cardio-vasculaire
	CORDONNER Daniel	Néphrologie
	COULOMB Max	Radiologie
	CROUZET Guy	Radiologie



MM.	CYROT Michel	Physique du solide
	DENIS Bernard	Cardiologie
	DOUCE Roland	Physiologie végétale
	DUSSAUD René	Mathématiques (CUS)
Mme	ETERRADOSSI Jacqueline	Physiologie
MM.	FAURE Jacques	Médecine légale
	FAURE Gilbert	Urologie
	GAUTIER Robert	Chirurgie générale
	GIDON Maurice	Géologie
	GROS Yves	Physique (I.U.T. I)
	GUIGNIER Michel	Thérapeutique
	GUITTON Jacques	Chimie
	HICTER Pierre	Chimie
	JALBERT Pierre	Histologie
	JUNIEN-LAVILLAVROY Claude	O.R.L.
	KOLODIE Lucien	Hématologie
	LE NOC Pierre	Bactériologie-virologie
	MACHE Régis	Physiologie végétale
	MAGNIN Robert	Hygiène et médecine préventive
	MALLION Jean-Michel	Médecine du travail
	MARECHAL Jean	Mécanique (I.U.T. I)
	MARTIN-BOUYER Michel	Chimie (CUS)
	MASSOT Christian	Médecine interne
	NEMOZ Alain	Thermodynamique
	NOUGARET Marcel	Automatique (I.U.T. I)
	PARAMELLE Bernard	Pneumologie
	PECCOUD François	Analyse (I.U.T. B) (Personnalité étrangère habilitée à être directeur de thèse)
	PEFFEN René	Métallurgie (I.U.T. I)
	PERRIER Guy	Géophysique-glaciologie
	PHELIP Xavier	Rhumatologie
	RACHALL Michel	Médecine interne
	RACINET Claude	Gynécologie et obstétrique
	RAMBAUD Pierre	Pédiatrie
	RAPHAEL Bernard	Stomatologie
Mme	RENAUDET Jacqueline	Bactériologie (pharmacie)
MM.	ROBERT Jean-Bernard	Chimie-physique
	ROMIER Guy	Mathématiques (I.U.T. B) (Personnalité étrangère habilitée à être directeur de thèse)
	SAKAROVITCH Michel	Mathématiques appliquées

MM. SCHAERER René	Cancérologie
Mme SEIGLE-MURANDI Françoise	Cryptogamie
MM. STOEBCNER Pierre	Anatomie pathologie
STUTZ Pierre	Mécanique
VROUSOS Constantin	Radiologie

#### MAITRES DE CONFERENCES ASSOCIES

MM. DEVINE Roderick	Spectro Physique
KANEKO Akira	Mathématiques pures
JOHNSON Thomas	Mathématiques appliquées
RAY Tuhina	Physique

#### MAITRE DE CONFERENCES DELEGUE

M. ROCHAT Jacques	Hygiène et hydrologie (pharmacie)
-------------------	-----------------------------------

Fait à Saint Martin d'Hères, novembre 1977



# INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

Année universitaire 1979-1980

Président : M. Philippe TRAYNARD  
Vice-Présidents : M. Georges LESPINARD  
M. René PAUTHENET

---

## PROFESSEURS DES UNIVERSITES

MM.	ANCEAU François	Informatique fondamentale et appliquée
	BENOIT Jean	Radioélectricité
	BESSON Jean	Chimie Minérale
	BLIMAN Samuel	Electronique
	BLOCH Daniel	Physique du Solide - Cristallographie
	BOIS Philippe	Mécanique
	BONNETAIN Lucien	Génie Chimique
	BONNIER Etienne	Métallurgie
	BOUVARD Maurice	Génie Mécanique
	BRISSONNEAU Pierre	Physique des Matériaux
	BUYLE-BODIN Maurice	Electronique
	CHARTIER Germain	Electronique
	CHERADAME Hervé	Chimie Physique Macromoléculaires
Mme	CHERUY Arlette	Automatique
MM.	CHIAVERINA Jean	Biologie, Biochimie, Agronomie
	COHEN Joseph	Electronique
	COUMES André	Electronique
	DURAND Francis	Métallurgie
	DURAND Jean-Louis	Physique Nucléaire et Corpusculaire
	FELICI Noël	Electrotechnique
	FOULARD Claude	Automatique
	GUYOT Pierre	Métallurgie Physique
	IVANES Marcel	Electrotechnique
	JOUBERT Jean-Claude	Physique du Solide - Cristallographie
	LACOUME Jean-Louis	Géographie - Traitement du Signal
	LANCIA Roland	Electronique - Automatique
	LESIEUR Marcel	Mécanique
	LESPINARD Georges	Mécanique
	LONGEQUEUE Jean-Pierre	Physique Nucléaire Corpusculaire
	MOREAU René	Mécanique
	MORET Roger	Physique Nucléaire Corpusculaire
	PARIAUD Jean-Charles	Chimie - Physique
	PAUTHENET René	Physique du Solide - Cristallographie
	PERRET René	Automatique

.../...

MM.	PERRET Robert	Electrotechnique
	PIAU Jean-Michel	Mécanique
	PIERRARD Jean-Marie	Mécanique
	POLOUJADOFF Michel	Electrotechnique
	POUPOT Christian	Electronique - Automatique
	RAMEAU Jean-Jacques	Chimie
	ROBERT André	Chimie Appliquée et des matériaux
	ROBERT François	Analyse numérique
	SABONNADIÈRE Jean-Claude	Electrotechnique
Mme	SAUCIER Gabrielle	Informatique fondamentale et appliquée
M.	SOHM Jean-Claude	Chimie - Physique
Mme	SCHLENKER Claire	Physique du Solide - Cristallographie
MM.	TRAYNARD Philippe	Chimie - Physique
	VEILLON Gérard	Informatique fondamentale et appliquée
	ZADWORNÝ François	Electronique

#### CHERCHEURS DU C.N.R.S. (Directeur et Maître de Recherche)

M.	FRUCHART Robert	Directeur de Recherche
MM.	ANSARA Ibrahim	Maître de Recherche
	BRONOEL Guy	Maître de Recherche
	CARRE René	Maître de Recherche
	DAVID René	Maître de Recherche
	DRIOLE Jean	Maître de Recherche
	KAMARINOS Georges	Maître de Recherche
	KLEITZ Michel	Maître de Recherche
	LANDAU Ioan-Doré	Maître de Recherche
	MERMET Jean	Maître de Recherche
	MUNIER Jacques	Maître de Recherche

#### Personnalités habilitées à diriger des travaux de recherche (décision du Conseil Scientifique)

##### E.N.S.E.E.G.

MM.	ALLIBERT Michel
	BERNARD Claude
	CAILLET Marcel
Mme	CHATILLON Catherine
MM.	COULON Michel
	HAMMOU Abdelkader
	JOUD Jean-Charles
	RAVAINE Denis
	SAINFORT

C.E.N.G.

.../...

MM. SARRAZIN Pierre  
 SOUQUET Jean-Louis  
 TOUZAIN Philippe  
 URBAIN Georges

Laboratoire des Ultra-Réfractaires ODEILLO

**E.N.S.M.E.E.**

MM. BISCONDI Michel  
 BOOS Jean-Yves  
 GUILHOT Bernard  
 KOBILANSKI André  
 LALAUZE René  
 LANCELOT François  
 LE COZE Jean  
 LESBATS Pierre  
 SOUSTELLE Michel  
 THEVENOT François  
 THOMAS Gérard  
 TRAN MINH Canh  
 DRIVER Julian  
 RIEU Jean

**E.N.S.E.R.G.**

MM. BOREL Joseph  
 CHEHIKIAN Alain  
 VIKTOROVITCH Pierre

**E.N.S.I.E.G.**

MM. BORNARD Guy  
 DESCHIZEAUX Pierre  
 GLANGEAUD François  
 JAUSSAUD Pierre  
 Mme JOURDAIN Geneviève  
 MM. LEJEUNE Gérard  
 PERARD Jacques

**E.N.S.H.G.**

M. DELHAYE Jean-Marc

**E.N.S.I.M.A.G.**

MM. COURTIN Jacques  
 LATOMBE Jean-Claude  
 LUCAS Michel  
 VERDILLON André



QUELQUES OUTILS INFOGRAPHIQUES  
POUR L'ANALYSE STRUCTURALE DE SYSTEMES

Ce travail, financé par la DGRST et l'IMAG, a été coordonné avec une recherche menée au GAPSET, sur contrat IRIA, sous la direction de M.J. CARPANO, chargée de recherche au GAPSET.





REMERCIEMENTS

---

A Monsieur G.VEILLON, professeur à l'I.N.P.G., président du Jury, à Messieurs Cl. BENZAKEN, professeur à l'U.S.M.G., M. LUCAS, maître de conférences à l'Université de Nantes, B. MUNIER, professeur à l'Université d'Aix-Marseille, et J.M. TREILLE, directeur du GAPSET, pour l'intérêt qu'ils ont témoigné à mon travail et l'honneur que me vaut leur présence dans le jury de cette thèse.

A Madame F. CHATELIN, professeur à l'U.S.S.G., rapporteur de cette thèse, pour la sympathie, les critiques et les encouragements prodigués au cours de mon travail et pour avoir mené à bien le déchiffrement de mon manuscrit initial...

A M.J. CARPANO et A. LEDUC-LEBALLEUR, respectivement chargée de recherche et ingénieur au GAPSET, pour les fructueux échanges de vues que nous avons eus à toutes les phases de ce travail.

A l'équipe Techniques Graphiques Interactives de l'IMAG pour ne s'être jamais rendue coupable du délit de non-assistance à utilisateur en danger ...

A Madame P. SOUILLARD, secrétaire à l'IMAG, et à l'équipe de reprographie de l'IMAG dont la diligence et l'efficacité ont permis d'effectuer la frappe et le tirage de ce texte dans des délais fort courts.

*POUR PLUS DE PRECISION*

*Il en revenait toujours là. Un et un  
- reprenait-il simplement pour se convaincre lui-même - égale deux.*

*Deux et deux : quatre. Mais pourtant  
le temps de faire cette addition, un intervalle infime  
s'intercalait, un souffle, et il ressentait le besoin  
de l'ajouter celui-là aussi, pour plus de précision - un souffle,  
et derrière cette haleine on distinguait le même chiffre  
comme inscrit en haut d'une porte ou sur le flanc d'un navire  
ou comme une étoile derrière un brouillard très fin, à cette heure vacillante,  
à cette heure si belle, entre le crépuscule et la nuit.*



SOMMAIRE

---

<u>CHAPITRE 1 : Introduction</u>	6
1. <u>L'ANALYSE DE SYSTEME</u>	6
1.1. Une méthodologie	6
1.2. L'émergence d'un concept	8
1.3. Le "paradigme Système"	9
2. <u>Présentation du travail</u>	14
2.1. Axes de travail	14
2.2. Problèmes étudiés	15
<u>CHAPITRE 2 : Graphes hiérarchisés</u>	18
1. <u>Représentation matricielle d'un système</u>	18
1.1. Modélisation matricielle	18
1.2. Graphe associé	19
2. <u>Graphes hiérarchisés</u>	20
3. <u>La représentation réduite éclatée</u>	22
3.1. Réduction	22
3.2. Hiérarchie des rangs	23
3.3. Hiérarchie serrée	26
3.3.1. Définitions	26
3.3.2. Algorithme de raccourcissement pour la hiérarchie directe	28
3.4. Représentation réduite éclatée d'un graphe orienté	35
4. <u>Représentations centrifuges</u>	38
4.1. Graphes non orientés connexes	38
4.2. Hiérarchie des écartements	40
4.3. Graphes orientés fortement connexes	41
4.4. Hiérarchies des écartements directes et inverses.	43

4.5. Représentation centrifuge standard	44
4.6. Exemple d'application	45

CHAPITRE 3 : Dessin automatique des graphes hiérarchisés 47

1. <u>Présentation des problèmes</u>	47
1.1. Le problème général de la planarisation	49
1.2. Problèmes spécifiques des graphes hiérarchisés	50
2. <u>Minimisation du nombre de croisements d'arêtes pour des graphes bipartis</u>	52
2.1. Dénombrement des croisements d'arêtes	53
2.2. La méthode des degrés relatifs : principe heuristique	56
2.3. Formalisation itérative	58
2.4. Propriétés des matrices d'itération	61
2.5. Convergence en ordre de la méthode des degrés relatifs	66
2.5.1. Définition de la convergence en ordre	66
2.5.2. Caractérisation	67
2.5.3. Un exemple d'optimalité	72
2.6. Etude expérimentale	74
2.6.1. Notations	74
2.6.2. Méthode de travail	74
2.6.3. Résultats	75
2.7. Un problème voisin : la bande diagonalisation des matrices	78
3. <u>Les graphes multipartis</u>	80
3.1. Le problème des croisements	81
3.2. Extension de la méthode des degrés relatifs	82
4. <u>Forme réduite éclatée : problèmes spécifiques</u>	84
4.1. Les shunts	84
4.2. La hiérarchie serrée	86
4.3. Les composantes fortement connexes	86
4.4. Heuristique des degrés relatifs.	86
5. <u>Forme centrifuge : l'heuristique des angles polaires</u>	91

<u>CHAPITRE 4</u> : Traitement préparatoire des tableaux de données : le seuillage	94
1. <u>Le seuillage</u>	95
2. <u>Diverses conditions de seuil</u>	96
3. <u>Problèmes d'interprétation</u>	98
<u>CHAPITRE 5</u> : Description des logiciels	103
1. <u>Traitement des graphes bipartis</u>	103
2. <u>Les logiciels LØVIGREC et CØPERNIC</u>	106
2.1. Infrastructure	106
2.2. Description externe	106
2.2.1. Les entrées	107
2.2.2. Structure du dialogue	110
2.2.3. Les sorties	111
2.3. Description interne	123
2.3.1. La structure de données	123
2.3.2. La structure graphique	128
2.3.3. Typologie des sous-programmes	129
<u>CHAPITRE 6</u> : Expérimentation et conclusions critiques	131
1. <u>Utilisation des logiciels</u>	131
1.1. Exemple d'utilisation de LØVIGREC	146
1.2. Exemple d'utilisation indépendante de CØPERNIC	159
2. <u>Bilan critique pour de nouvelles directions de recherche</u>	159
2.1. Hiérarchies dans les graphes	159
2.2. Croisements d'arêtes et problèmes connexes	161
2.2.1. Dessin automatique	162
2.2.2. Bande diagonalisation	162
2.2.3. Classification automatique interactive.	162
2.3. Problèmes de seuillage	167
2.3.1. Changement qualitatif dans une structure hiérarchique	168
2.3.2. Ressemblance entre deux graphes hiérarchisés	169

2.4. Fonctionnement des logiciels	170
2.4.1. Souplesse d'utilisation	170
2.4.2. Accessibilité	170
2.4.3. Vitesse de fonctionnement.	171
2.5. Acquis Méthodologiques et extensions envisagées.	172
3. <u>Conclusions générales</u>	174
Références.	176

## CHAPITRE I

### INTRODUCTION

Nous allons présenter rapidement le contexte global de notre travail : l'analyse structurale de systèmes.

Ensuite nous essayerons de situer dans cet ensemble très vaste l'objet de cette thèse, ainsi que ses caractères spécifiques.

#### 1. L'ANALYSE DE SYSTEMES

##### 1.1. Une méthodologie

Au départ, l'analyse de systèmes peut être considérée comme une nébuleuse de techniques assez disparates, transplantées du domaine des mathématiques vers plusieurs champs d'applications pratiques, avec des résultats variables. Au contact des problèmes pratiques de telle ou telle discipline, ces outils empiriques se sont à la fois diversifiés et unifiés jusqu'à constituer une sorte de méthodologie générale de la modélisation.

Ces techniques variées vont des structures logiques et ensemblistes les plus simples (treillis, groupes, monoïdes) utilisées en linguistique, en psychologie, en anthropologie (cf. les travaux de Chomsky en grammaire, de Piaget en psychogénèse, et les célèbres études de Levi-Strauss sur les réseaux de parentés et la "pensée sauvage" etc ...) à des outils très sophistiqués du point de vue mathématique : optimisation linéaire ou non, calcul matriciel, équations aux dérivées partielles etc ... utilisés maintenant très couramment en économie, en automatique, en biologie (cf. Bertalanffy [2]).



On peut classer, en première approximation, les méthodes utilisées en trois groupes :

- \* les structures *discrètes* qui permettent d'analyser un système d'un point de vue *statique* : structures ensemblistes, théorie des graphes, calcul matriciel algébrique.

- \* les structures *continues* qui rendent possible l'étude *dynamique* (souvent localement : petites variations autour d'une zone d'équilibre) : équations différentielles.

- \* les structures *aléatoires* qui permettent l'introduction de l'incertitude dans le fonctionnement du système : par exemple pour l'étude des files d'attente, ou le traitement du signal.

Cette répartition est purement descriptive. En fait, elle correspond plutôt à un idéal d'approche d'un système :

On commence en général par se donner un ensemble d'éléments sur lesquels se noue tout un réseau d'interactions, qu'on cherche à apprécier quantitativement : ici la théorie des graphes et le calcul matriciel sont des outils privilégiés pour effectuer un découpage qualitatif du système, ou pour calculer une répartition dans un état d'équilibre spatio-temporel (graphes hiérarchisés, méthodes de la R.O. pour l'optimisation dans les réseaux, résolution numérique de systèmes d'équations ...).

Une fois cernés les principaux mécanismes, on peut chercher à dynamiser le système : étude de la stabilité, action des conditions extérieures etc ...

Les relations et flux entre éléments sont décrits comme des fonctions du temps, et non plus comme des constantes ou des agrégats.

On travaille alors sur des équations aux dérivées partielles (spatiales ou temporelles) ou sur des fonctions de transition (systèmes à états) qui forment un "pont" entre le discret et le continu.

Finalement, on peut essayer d'affiner le modèle construit en tenant compte des incertitudes qui entachent les données de base et introduire ainsi de petites variations aléatoires qui permettent de préciser la sensibilité du système aux perturbations etc ...

Suivant le type de données dont on dispose, et le problème que l'on cherche à résoudre, l'une ou l'autre de ces trois étapes sera prépondérante dans la démarche effectuée : si les données sont intrinsèquement statiques, seule la première étape sera parcourue.

Si au contraire on dispose de variables aléatoires (pour décrire le système de façon non déterministe), la troisième étape sera essentielle dans la construction d'un modèle.

## 1.2. L'émergence d'un concept

Nous avons parlé jusqu'ici de système, de structure, de modèle, sans préciser la signification de ces termes.

Parler de système, c'est vouloir évoquer une "totalité" du point de vue du "jeu" de ses "parties" et de ses rapports à son "environnement".

On entend en général par structure l'organisation interne d'un système, l'ensemble des lois qui régissent le système (Mais J. Piaget [17] définit une structure comme "système de transformations"!) et par modèle la représentation formelle d'un système (cf. B. Walliser [22]). Ces distinctions ne servent en fait pas à grand-chose pour clarifier la notion de système : que serait un système non structuré, sinon précisément un non-système ? Quant à la distinction système / modèle, elle fait oublier qu'un système n'est pas plus "concret" que le modèle qui le représente. Définir un système, c'est déjà construire une abstraction mentale, et non pas découper un morceau de la "Nature". Entre système et modèle la différence n'est pas d'*essence* (qualitative) mais de *niveau* d'abstraction, de *degré* de formalisation (quantitative).

Il est préférable, pour mieux cerner le contenu du concept de système de laisser s'enchaîner les associations d'idées qu'il provoque. La première notion ainsi repérable est celle de "totalité". Un système est donc une entité *isolable* dans un espace plus vaste qui lui sert d'environnement. Une deuxième notion apparaît immédiatement derrière cette isolabilité, c'est celle de *stabilité*. Pour isoler un système sur le fond de décor de son environnement, il faut pouvoir lui prêter un certain nombre d'*invariants*, qui garantissent son identité par une certaine permanence dans l'espace, dans le temps ou à travers certaines règles de transformations. La troisième notion essentielle pour préciser le concept de système est celle de *régulation*. On entend par là l'ensemble des mécanismes qui permettent d'assurer la permanence du système malgré (et grâce à) des transformations de son environnement ou des changements internes.

A travers ces trois notions : totalité, stabilité et régulation, le concept de système se déploie dans le champ de la pensée comme un moyen de mettre un nom sur tout ce qui constitue une organisation plus ou moins autonome d'éléments maintenus en équilibre dynamique par des échanges avec le milieu extérieur et/ou des restructurations internes.

### 1.3. Le "paradigme<sup>\*</sup> Système"

Séduits par la richesse et la souplesse du concept de système, de nombreux auteurs (dont le premier fut Bertalanffy [2]) ont pensé que cet outil d'une grande généralité d'emploi annonçait une nouvelle révolution scientifique, comparable à ce que furent, en leur temps, la géométrie d'Euclide, l'astronomie de Copernic ou la mécanique relativiste d'Einstein.

---

\* En linguistique, les paradigmes sont les unités sémantiques (par opposition aux unités syntaxiques : les syntagmes).

En épistémologie, depuis l'ouvrage de Th. S. Kuhn "La structure des révolutions scientifiques", ce mot a servi à désigner un concept (ou un ensemble de concepts) jouant un rôle-clé dans le développement de la connaissance à une époque donnée.

Le succès des termes de "système" et de "structure" est dû à plusieurs facteurs, en particulier :

1) Un renouvellement indéniable de la méthodologie des sciences sociales et des sciences humaines grâce à l'approche structuraliste.

2) Une démarche globalisante et dialectique face à une parcellisation et une spécialisation de plus en plus poussées du savoir scientifique.

En France, une mode intellectuelle baptisée "structuralisme" et inspirée dans les années soixante par les travaux d'auteurs renommés (Althusser, Barthes, Derrida, Foucault, Goldmann, Lacan, Lévi-Strauss) a cru voir dans ces deux facteurs l'amorce d'une "méta-science", d'une théorie générale de l'élaboration de la connaissance scientifique, dont toutes les disciplines classiques seraient destinées à devenir des applications locales particulières.

Le structuralisme suscita dans l'intelligentsia un engouement comparable à l'intérêt éveillé par le paratonnerre de Franklin et le baquet de Mesmer dans les salons parisiens du XVIII<sup>e</sup> siècle.

Au delà du phénomène de mode (cf. plus récemment les "nouveaux philosophes") trois facteurs idéologiques nous paraissent liés à ce succès, en plus des deux points positifs précédemment cités :

3) d'un point de vue historique à court terme, la prospérité économique ininterrompue des années 50 et 60, appuyée sur les recettes keynésiennes de régulation étatique du capitalisme "justifiait" l'approche systémique dans le domaine économique et social grâce à l'absence de tensions sociales fortes. (Depuis, le caractère conjoncturel de cette situation est apparu nettement ..).

4) le concept de système impose une réconciliation dialectique de la Structure et de l'Histoire, de la synchronie et de la diachronie, de la Continuité et du Changement.

Cependant, aucune formalisation mathématique (faute de pouvoir modéliser les "sauts qualitatifs") ne peut être effectuée sans privilégier les invariances qui donnent son identité au système. Le Système, c'est donc "le changement dans la continuité", avec la certitude réconfortante que la continuité structurale donne les limites que le changement ne peut dépasser, sous peine de détruire les invariants de référence.

En période de stabilité socio-politique, ces invariants de référence prennent l'aspect d'incontestables évidences éternelles d'où l'"utilité" de l'approche systémique dans le domaine économique (régulation keynésienne) et politique (sociologie fonctionnaliste et "théorie du consensus" ; cf. par exemple les travaux de David Easton).

5) la liaison étroite entre l'analyse de systèmes et la biologie permet de considérer les organismes vivants comme des systèmes particuliers, mais surtout, elle amène à envisager des systèmes comme des êtres vivants (cf. le célèbre automate nommé "life") mis en équations.

Sur le plan historique, on peut ainsi remarquer que la conception moderne de la régulation (comme invariance dynamique) apparaît avec la notion d'homéostasie en biologie.

De même, la vision fonctionnaliste des systèmes sociaux comme des ensembles d'"institutions" associées à des "fonctions sociales" doit beaucoup à la conception schématique des êtres vivants comme assemblages d'organes (chaque organe ayant une fonction et une seule) à la manière d'un mouvement d'horlogerie.

Ce biologisme plus ou moins conscient a trois conséquences : il donne l'impression valorisante d'être capable de manipuler la complexité extraordinaire que recèle un organisme vivant, il incite souvent paradoxalement à refuser (aux autres) le droit d'intervenir dans le fonctionnement du système vu la fragilité des systèmes complexes, et enfin, dans la mesure où il met en correspondance biunivoque des organes et des fonctions, il évacue toutes les contradictions qui pourraient se manifester si un même organe pouvait entrer en conflit de compétence avec d'autres pour assurer une fonction donnée ...

Sur le plan pratique, ces biais idéologiques ont des conséquences contradictoires qui se traduisent par un comportement qu'on pourrait qualifier de "progressiste conservateur" du fait d'un mélange de volontarisme technocratique (cf. facteur 5 : sentiment de capacité à maîtriser la complexité) et d'une perspective historique privilégiant les invariances (facteurs 3 et 4).

Cette dérive idéologique fut vigoureusement attaquée par Henri Lefebvre (cf. [14]) qui considérait le structuralisme comme l'idéologie propre de la caste des technocrates.

Des critiques similaires, plus ou moins élaborées théoriquement, trouvèrent un large écho dans les mouvements étudiants des années 67-72 ; citons en particulier : H. Marcuse : "L'homme unidimensionnel" Ed. Minuit (1968) J. Habermas, "La technique et la science comme idéologie" trad. Franç. Galimard (1975).

Avec quelques années de recul, et la crise économique aidant, ni l'espoir de certains structuralistes d'arriver à bâtir le "Système de tous les systèmes", ni la prétention des technocrates à réguler l'ensemble de la société suivant la stricte rationalité "scientifique", ni non plus le romantisme révolutionnaire de leurs opposants n'ont été justifiés.

De part et d'autre, on a redécouvert l'irréductibilité de certaines contradictions.

Désormais plus conscients de leurs limites, les théoriciens de l'analyse de systèmes avancent à pas comptés, en évitant les extrapolations hâtives et les prétentions à l'infailibilité.

Sur le plan théorique, le débat de fond reste ouvert, et tourne autour du concept de "rationalité".

Le propre de la mentalité technocratique est de croire à la réductibilité de toute décision à des critères techniques.

Si cela était, il n'y aurait plus de choix à effectuer, mais simplement une solution optimale à découvrir et à appliquer.

Transposée dans le cadre du "paradigme Système" cette attitude signifie la croyance en l'identité de la rationalité interne et de la rationalité externe pour les éléments d'un système, c'est-à-dire qu'une "partie" du "tout" ne peut pas poursuivre un but propre (rationalité interne) qui aille à l'encontre de la conservation du "tout" (rationalité externe).

Cette pétition de principe, qui va de soi pour des systèmes physiques, est insoutenable pour les systèmes sociaux, du fait que les éléments (individus ou groupes sociaux) y possèdent des capacités de représentation consciente de la réalité et d'élaboration d'un *projet* qui peut (consciemment ou non) être contradictoire avec la structure du système, même si celle-ci détermine dans une large mesure la nature même des représentations mentales des individus ou des groupes.

## 2. PRESENTATION DU TRAVAIL

En regard des vastes problèmes théoriques et pratiques qui viennent d'être abordés, notre travail a été fort modeste.

### 2.1. Axes de travail

1) Notre but a été avant tout pratique : il s'est agi d'utiliser, par la production de logiciels effectivement opérationnels, les *techniques graphiques interactives* dans le domaine de l'analyse structurale de système.

Plus précisément, nous nous sommes orientés, à l'initiative du G.A.P.S.E.T.\* , vers une catégorie précise de représentation des systèmes : les graphes. Selon la formule de A.Kaufmann , un graphe c'est "des points et des flèches". Représenter un système par un graphe fournit donc immédiatement un support visuel à la représentation. Nous avons donc envisagé la réalisation de logiciels permettant d'automatiser :

a) la détermination d'une structure hiérarchisée dans un graphe, c'est-à-dire d'une partition de l'ensemble des sommets en niveaux suivant une certaine relation d'ordre.

b) la représentation du graphe hiérarchisé ainsi obtenu sous la forme d'un dessin aussi clair que possible, faisant bien ressortir la structure hiérarchique, avec des arcs rectilignes entre les sommets et peu de croisements d'arêtes.

En effet, d'une part, un graphe hiérarchisé communique instantanément une vision globale de la structure du système qu'il représente (filières, relations indirectes, composantes fortement connexes etc ...), vision qui n'existe pas lorsque l'on examine seulement sa matrice d'incidence ou la liste de ses arcs. D'autre part, les méthodes développées permettent de mettre en évidence des sous-systèmes particuliers et des relations entre ces sous-systèmes.

---

\* Groupe d'Analyse et de Prospective des Systèmes Economiques et Technologiques.



2) L'intérêt de logiciels *interactifs* réside essentiellement dans leur souplesse du point de vue de l'utilisation :

- \* possibilité de corriger manuellement les dessins obtenus,
- \* grande facilité pour modifier les données de départ.

Cette souplesse est particulièrement attrayante si on envisage de faire de la simulation sur des modèles qu'on souhaite pouvoir modifier au vu des résultats obtenus ou bien si l'utilisateur a besoin d'une assez grande latitude de choix pour sélectionner les résultats les plus intéressants.

Il ressort de la classification esquissée en 1.1. que ce choix implique une approche a priori *statique* des systèmes, l'ensemble des éléments du système étant montré en *interaction figée*, selon la structure du graphe associé.

Malgré cela, l'aspect dynamique reste sous-jacent de deux manières :

- \* une dynamisation temporelle discrète est toujours envisageable, à partir d'une succession de données sur le système prises pour des dates différentes.

- \* dans une représentation sous forme de graphe, les arcs du graphe pourront souvent représenter un cumul de flux entre sous-systèmes pendant une période de temps donnée.

## 2.2. Problèmes étudiés

Pour passer automatiquement d'un système donné sous une forme quelconque à sa représentation par le dessin d'un graphe hiérarchisé, trois problèmes sont à résoudre, que nous présentons dans l'ordre suivant lequel ils apparaissent au cours de la modélisation :

2.2.1. la préparation des données : il s'agit essentiellement de mettre le système sous une forme permettant le passage à un graphe. La forme mathématique ultime pour cette phase de travail est une matrice booléenne décrivant les relations entre tous les éléments du système pris deux à deux. Cette forme n'implique pas de restrictions sur le plan de la modélisation des interactions du système avec l'environnement : il suffit pour cela de construire un sur-système englobant le système et les facteurs extérieurs susceptibles d'agir sur lui, puis de mener l'étude avec ce sur-système.

Le problème général de la modélisation sous forme matricielle, ne sera pas abordé ici ; le problème de la préparation des données sera étudié au chapitre 4 dans le cas particulier du passage d'une matrice (à coefficients réels ou entiers) à une matrice booléenne par des techniques de *seuillage*.

2.2.2. La structuration du système : on doit définir en termes de théorie des graphes les types de structures que l'on veut mettre en évidence visuellement, afin que la vision globale apportée par le dessin du graphe soit porteuse d'informations intéressantes sur le système étudié.

Les structures hiérarchiques que nous avons étudiées seront présentées au chapitre 2, à partir de la notion générale de *graphe hiérarchisé*. Nous disposons au départ de la méthode classique de hiérarchisation par les rangs sur le graphe réduit, et d'une forme de représentation définie par M.J. Carpano [4] à partir de travaux de Warfield [23] [24]. Nous avons précisé cette représentation en vue du dessin automatique et nous avons défini au chapitre 2, une deuxième méthode de hiérarchisation, complémentaire de la première, et à laquelle nous avons également associé une forme standard de représentation.

2.2.3. le dessin automatique : pour que les informations véhiculées par les formes standard de représentation soient exploitables, il faut que les dessins soient *lisibles*.

Pour chacune des deux représentations adoptées, nous définirons au chapitre 3 une heuristique de placement des sommets qui vise à obtenir peu de croisements d'arcs, tout en respectant les contraintes de disposition imposées par les formes standard .

Nous avons pu donner de l'une de ces heuristiques une formalisation itérative intéressante dans le cas particulier de graphes bipartis. Cette formalisation garantit la convergence de la méthode et justifie partiellement son efficacité pratique par des propriétés d'optimalité établies dans des cas simples. Dans le cas général, les possibilités de modifications interactives des dessins sur écran graphique permettent à l'utilisateur d'améliorer la lisibilité des dessins obtenus.

Finalement, les deux méthodes de hiérarchisation de graphes qui ont été retenues ont donné lieu à l'écriture de deux logiciels interactifs de visualisation pouvant être utilisés indépendamment ou complémentirement pour étudier des systèmes mis sous forme matricielle.

Nous présenterons ces logiciels au chapitre 5, en insistant sur le point de vue de l'utilisateur (structure du dialogue interactif).

Au chapitre 6, nous décrirons quelques exemples d'utilisation des logiciels réalisés sur des données matricielles (tableaux économiques) fournies par le GAPSET.

Nous essaierons ensuite de tirer un bilan critique d'ensemble et de suggérer quelques prolongements envisageables du travail effectué.

## CHAPITRE 2

### LES GRAPHES HIERARCHISES

#### 1. REPRESENTATION MATRICIELLE D'UN SYSTEME

##### 1.1. Modélisation matricielle

La modélisation mathématique d'un système peut être effectuée de la façon suivante :

- on définit l'ensemble des *éléments* du système, qui sont considérés comme homogènes entre eux d'un certain point de vue.

- sur cet ensemble sont définies des relations qui traduisent toutes les *interactions* pertinentes entre les éléments du système au niveau de modélisation envisagé.

Nous ne considérons ici que la structure la plus pauvre qui puisse être associée à un système : une relation  $R(x,y)$  sur les éléments du système,  $x R y$  voulant dire : "il y a une relation dans le système de l'élément  $x$  à l'élément  $y$ ".

## 1.2. Graphe associé

a) Si la relation  $R$  est *qualitative*, elle est entièrement déterminée par la donnée d'une matrice carrée booléenne  $M$  définie comme suit :

Soit  $X = \{x_i, i = 1, 2, \dots, n\}$  l'ensemble des éléments du système,

$M = (M_{ij})$   $i, j = 1, 2, \dots, n$  est telle que :

$$(M_{ij} = 1) \iff x_i R x_j \quad i, j = 1, 2, \dots, n$$

$$(M_{ij} = 0) \iff x_i \neg R x_j \quad i, j = 1, 2, \dots, n$$

$M$  peut alors être considérée comme la matrice d'incidence aux sommets d'un graphe  $G = (X, U)$  dont chaque arc  $u_{ij} \in U$  symbolise une relation de  $x_i$  à  $x_j$  dans le système, chaque élément du système devant être représenté par un sommet du graphe  $G$ .

b) Si la relation est *quantitative*, on la suppose définie par une matrice  $M$  à coefficients entiers ou réels, en général ;  $G$  est alors un graphe *valué* : à chaque arc  $u_{ij} = (x_i, x_j)$  est associé le coefficient  $M_{ij} \neq 0$ .

Remarque : si  $x R y$  et  $y R x$  ont même signification dans le système, la matrice  $M$  est symétrique, et le graphe  $G$  associé est non orienté.

Pour la formalisation simple qui vient d'être décrite, on voit qu'il suffit de définir sur un système une liste des relations entre ses éléments pour pouvoir le représenter par un graphe.

Dans chaque système étudié de cette manière, on peut espérer donner un sens à un certain nombre de notions de la théorie des graphes, et en extraire ainsi des informations sur la structure interne du système.

Remarque épistémologique :

Cette approche recèle un danger évident : le réductionnisme. Un système *n'est pas* un graphe : représenter un système par son graphe, c'est privilégier systématiquement dans les relations entre éléments les contenus aux dépens des formes, privilégier l'existence même de la structure par rapport aux modes de cette existence (cf. Lefebvre [14]). Par exemple, il peut être tentant de simplifier à outrance la relation binaire  $R$  pour obtenir un graphe lisible. On aura toujours une structure, mais elle sera trop schématique pour être utile à la compréhension du système. Réciproquement, il serait absurde de vouloir transcrire toutes les notions de la théorie des graphes dans les termes du modèle. Seules certaines d'entre elles seront intéressantes. Ainsi, le problème de la pertinence des correspondances système/graphe se pose dans les deux sens.

Il nous a paru d'autant plus important de le rappeler ici que ce point ne sera pas examiné dans cet exposé, bien que ce type de problème soit très important du point de vue de la validité des méthodes d'analyse structurale.

2. GRAPHES HIERARCHISES

Définition :

On appelle *hiérarchie*  $H$  sur un graphe  $G = (X, U)$  toute application de l'ensemble des sommets  $x = \{x_i, i = 1, 2, \dots, n\}$  dans une partie  $I = \{1, 2, \dots, k\}$  de  $N$  associant à tout sommet  $x_i \in X$  une valeur  $n_i \in I$ . On dit que le sommet  $x_i$  est au niveau  $n_i$  dans la hiérarchie  $H$ , et  $k$  est le nombre de niveaux de la hiérarchie  $H$ .

En pratique,  $H$  est définie à partir d'une relation d'ordre sur les sommets : une hiérarchie est significative par rapport à la structure du graphe si l'ordre des entiers naturels entre les niveaux est compatible avec une relation d'ordre  $P$  sur l'ensemble des sommets,  $P$  étant dérivée

de telle ou telle notion de la théorie des graphes :

$$x_i P x_j \implies n_i \leq n_j$$

Mettre un graphe G sous une certaine forme hiérarchisée H présente l'intérêt de faire apparaître dans le système de départ une structure d'ordonnement entre éléments ou groupes d'éléments induite par la relation d'ordre P.

Nous allons définir et *dessiner* deux types de hiérarchies dans les graphes.

La première est dérivée de la classique hiérarchie des rangs. Celle-ci est liée à l'analyse des propriétés de connexité forte dans les graphes orientés. C'est la hiérarchie la plus naturelle pour l'analyse de systèmes, car elle permet un découpage en sous-systèmes sans rétroactions entre eux. (Lorsque l'on parle de "systèmes hiérarchisés", c'est à cette hiérarchie des rangs qu'il est généralement fait référence. [5] [18] [23]).

Nous modifierons la hiérarchie des rangs en introduisant la notion de hiérarchie "serrée", pour définir ensuite la représentation "réduite éclatée" d'un graphe orienté quelconque. Cette représentation sera la forme standard sous laquelle nous dessinerons les graphes ainsi hiérarchisés.

Nous définirons, à partir de la notion de centre d'un graphe, une deuxième méthode de hiérarchisation. La hiérarchie obtenue (dite "des écartements") est complémentaire de la première deux points de vue :

1. - elle permet de hiérarchiser des graphes *non orientés*
2. - elle peut aussi donner une représentation des graphes *orientés fortement connexes* pour lesquels la première méthode est sans intérêt.

### 3. LA REPRESENTATION REDUITE ECLATEE

#### 3.1. Réduction

a) Définition du graphe réduit : on appelle *graphe réduit*

$G_r = (X_r, U_r)$  d'un graphe *orienté*  $G = (X, U)$ , le graphe défini de la façon suivante :

\* les éléments de  $X_r$  représentent les composantes fortement connexes (notées c.f.c.) de  $G$  (c'est-à-dire les ensembles de sommets mutuellement accessibles de  $G$ ).

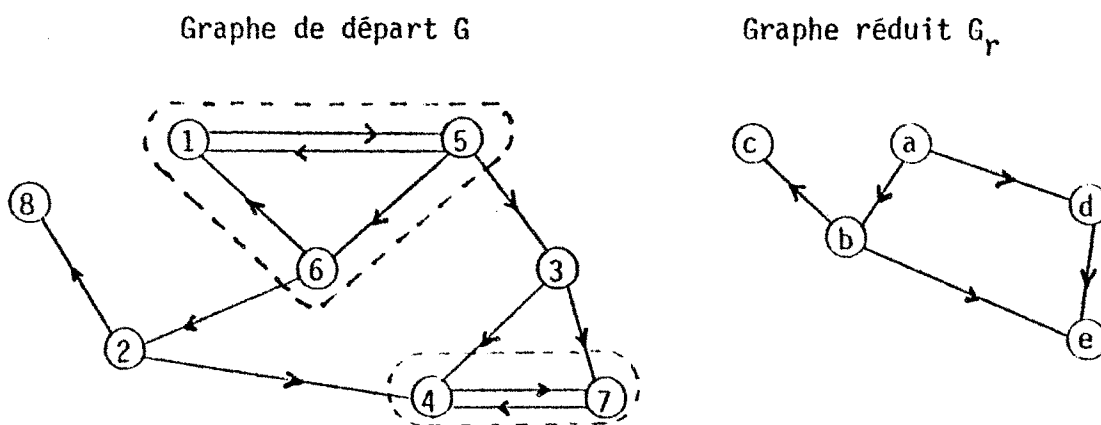
\* les éléments de  $U_r$  sont les arcs reliant les c.f.c. entre elles, c'est-à-dire que  $(a,b)$  existe dans  $U_r$  si, et seulement si, il y a au moins un arc reliant dans  $G$  un des sommets de la c.f.c.  $a$  à un des sommets de la c.f.c.  $b$ .

b) Propriétés

\*  $G_r$  est sans circuit, puisque tous les circuits sont internes aux c.f.c.

\* si  $G$  est sans circuit,  $G_r$  est identique à  $G$ , toutes les c.f.c. de  $G$  sont réduites à un seul sommet.

Exemple :



- Figure 1 -

$a = (1,5,6)$  et  $e = (4,7)$ , les autres c.f.c. sont des sommets isolés.



Les algorithmes de passage de  $G$  à  $G_r$  sont bien connus : méthode de double marquage des sommets de  $G$ , ou calcul de la matrice de fermeture réflexo-transitive (cf. Roy [20]). (Nous avons adopté l'algorithme de double marquage de Roy).

L'intérêt du graphe réduit pour notre problème est qu'il ne contient aucun circuit. En effet, la méthode de hiérarchisation par le rang qui va maintenant être décrite ne s'applique qu'au cas des graphes sans circuit.

### 3.2. Hiérarchie des rangs

Définition du rang : dans un graphe orienté sans circuit, un sommet  $s$  est dit de *rang*  $r$  si et seulement si  $r$  est le nombre maximal d'arcs de tout chemin d'extrémité terminale  $s$ . (cf. Roy [20], Berge [1]). (Un sommet sans antécédent est donc de rang zéro).

On voit tout de suite que le rang ne peut pas être défini s'il y a des circuits : en faisant autant de fois qu'on veut le tour d'un circuit, on obtiendrait un rang aussi grand qu'on voudrait. Autrement dit, tout sommet ayant des successeurs formant un circuit serait de rang infini.

Soit la relation d'ordre  $P_r$  dite *relation de précédence* :

$x P_r y \iff$  "il existe un chemin de  $x$  à  $y$  dans le graphe  $G$ ".

L'application  $H_r : x_i \xrightarrow{H_r} n_i = r_i + 1$  est la *hiérarchie des rangs*.

$H_r$  est compatible avec  $P_r$ .

Le calcul des niveaux se fait par l'algorithme suivant :

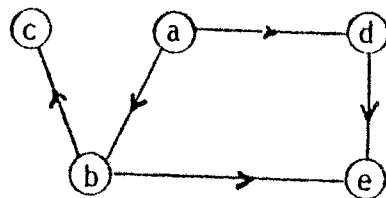
On commence par marquer 1 tous les sommets sans antécédents (il en existe au moins un, puisque le graphe est sans circuit). Ensuite on marque 2 tous les suivants des sommets marqués 1, puis on marque 3 les suivants des sommets marqués 2 etc ... On arrête le marquage lorsqu'il n'y a plus de marques nouvelles à ajouter.

Comme le graphe est sans circuit et qu'il comporte un nombre fini d'arcs, l'algorithme est fini.

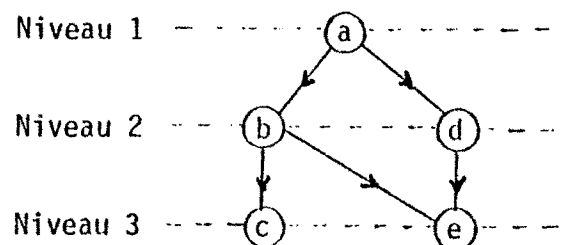
Le nombre de marquages effectués étant égal au nombre d'arcs, si  $n$  est le nombre de sommets du graphe, l'algorithme est en  $n^2$ . On peut hiérarchiser le graphe en sens inverse, en commençant par marquer les sommets sans successeur (cf. figure 2).

Une fois effectuée la détermination des niveaux, on peut adopter une convention de représentation pour le dessin du graphe hiérarchisé :

Graphe orienté sans circuit



Graphe hiérarchisé par les rangs



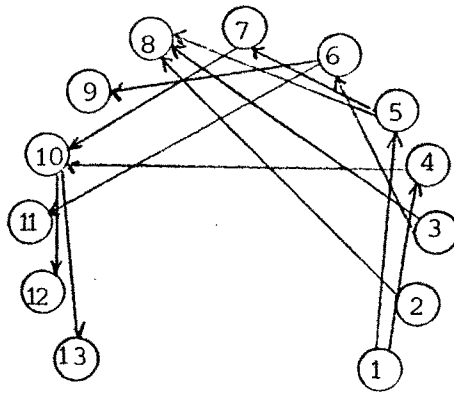
- Figure 2 -

Ici, les niveaux sont symbolisés par des lignes horizontales parallèles entre elles et équidistantes. Cette convention sera conservée dans la suite pour définir la représentation standard. L'intérêt d'une telle représentation est évident, même sur un graphe aussi simple.

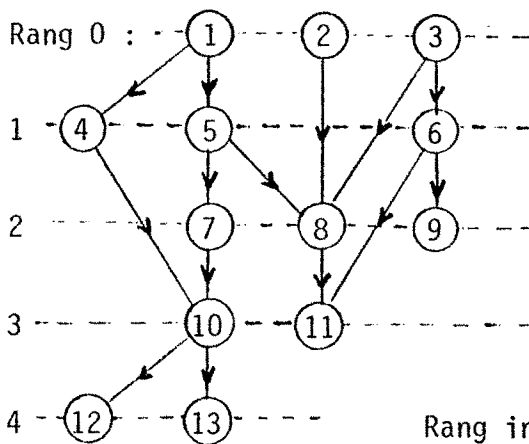
Considérée telle quelle, la hiérarchisation par le rang a deux inconvénients, l'un d'ordre théorique, l'autre pratique.

L'inconvénient théorique est le suivant : le résultat obtenu en déterminant la hiérarchie inverse des rangs (c'est-à-dire en calculant les niveaux sur le même graphe, mais avec tous les arcs changés de sens) n'est pas équivalent à la hiérarchisation directe. Ainsi, suivant les auteurs et les algorithmes utilisés, on obtient donc des représentations différentes : Roy [20] construit la hiérarchie directe, Warfield [24] la hiérarchie inverse.

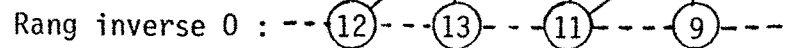
Exemple : Graphe non hiérarchisé



Hiérarchie directe



Hiérarchie inverse



- Figure 3 -

Si les deux hiérarchies donnaient la même représentation, on aurait pour tout sommet  $s$  :  $\text{rang inverse}(s) = L - \text{rang}(s)$  où  $L$  désigne le rang le plus élevé, qui est le même dans les deux hiérarchies, puisque c'est aussi la longueur du plus long chemin dans le graphe. (Ici  $L = 4$ ).

L'inconvénient pratique apparaît sur la figure 3 : certains arcs sont inutilement longs (2 - 8 à gauche, 5 - 8 à droite). Pour le dessin automatique de ces graphes, il est préférable de ne pas avoir d'arcs plus longs que nécessaire, afin de ne pas augmenter les risques de croisements d'arcs.

Nous allons donc définir une nouvelle hiérarchie qui éliminera ces deux inconvénients tout en respectant la condition de compatibilité avec la relation de précédence  $P_r$  des graphes sans circuit :  $x_i P_r x_j \iff n_i \leq n_j$ . Cette hiérarchie aura donc le même intérêt que la hiérarchie des rangs par rapport à la structure du graphe.

### 3.3. Hiérarchie serrée

#### 3.3.1. Définitions :

Par référence à la terminologie de Warfield [25], nous appelons *portée* d'un arc  $(x_i, x_j)$  la différence  $n_j - n_i$  des niveaux de ses extrémités.

Un *chemin maximum* est un chemin contenant autant de sommets qu'il y a de niveaux dans la hiérarchie.

Un *embranchement descendant* (respectivement *ascendant*) est un chemin dont le premier sommet (respectivement le dernier) appartient à un chemin maximum.

Un *sommet flottant* est un sommet intermédiaire d'un chemin qui est à la fois embranchement descendant et embranchement ascendant.

Par exemple, sur la figure 3 :

- les chemins maxima sont 1-5-7-10-12 et 1-5-7-10-13 ;
- le chemin 5-8-11 est un embranchement descendant ;

- le chemin 1-4-10 est un embranchement à la fois descendant et ascendant ;
- le sommet 4 est un sommet flottant.

Propriété : Pour tout graphe sans circuit, il existe une hiérarchie *unique* (modulo l'affectation des sommets flottants), qui soit compatible avec l'ordre des rangs et qui conduise à une représentation telle que les portées de tous les arcs soient minimales.

Cette hiérarchie est appelée la *hiérarchie serrée*. La propriété énoncée sera établie de manière constructive au paragraphe 3.3.2. suivant, où est décrit l'algorithme de raccourcissement. Observons d'abord que la condition d'équivalence des hiérarchies directes et inverses (et simultanément la condition de portée minimale des arcs afin de minimiser la longueur des arcs sur le dessin) est vérifiée pour tous les sommets faisant partie d'un chemin maximum.

En effet, dans ce cas, un sommet de niveau  $k$  strictement compris entre 1 et  $L - 1$  est encadré par un précédent de niveau  $k - 1$ , et par un suivant de niveau  $k + 1$ , et tous les arcs du chemin sont de portée 1, donc minimale. De même, dans la hiérarchie inverse, d'où l'équivalence.

Le problème va donc consister à redéfinir les positions des sommets n'appartenant pas aux chemins maxima : les difficultés proviennent en fait des chemins partant de certains sommets de rang 0 qui conduisent à des arcs de portée trop grande au moment où ces chemins rejoignent les chemins maxima ou leurs embranchements descendants, car les embranchements descendants sont construits "dans le sens de la marche" par la hiérarchie directe et ont donc des arcs de portée 1, qui n'ont pas à être raccourcis.

Un raisonnement symétrique sur la hiérarchie inverse amènerait à déplacer vers le haut certains sommets qui ne sont pas sur les chemins maxima, ni sur les embranchements ascendants.

### 3.3.2. Algorithme de raccourcissement pour la hiérarchie directe

Pour raccourcir les arcs en modifiant les positions de certains sommets, nous allons procéder de la manière suivante :

1 \* on calculera pour chaque sommet  $s$  le nombre maximum de niveaux dont on peut le descendre sans enfreindre la compatibilité entre les niveaux et l'orientation des arcs.

2 \* Soit  $DES(s)$  le tableau contenant les résultats de ce calcul pour tous les sommets, nous modifierons  $DES(s)$  pour ne jamais descendre un sommet plus qu'un de ses précédents, ce qui allongerait au moins un arc.

3 \* Une fois faites toutes les modifications, on affectera chaque sommet à son niveau définitif :

$$NIV(s) = NIV(s) + DES(s).$$

L'algorithme que nous présentons ici est celui qui a été effectivement programmé en Fortran. Il est dérivé de l'algorithme doublement récursif présenté dans Delarche [6].

#### 1) Calcul du tableau DES

Ce calcul s'effectue simplement si on parcourt le graphe par niveaux en partant du bas.

- Si un sommet  $s$  n'a pas de suivant on peut le descendre jusqu'au dernier niveau  $L$  :  $DES(s) = L - NIV(s)$  ( $E_1$ ).

- Sinon, on peut le descendre jusqu'à ce qu'il soit juste au dessus d'un (au moins) de ses suivants :

$$DES(s) = \min_{s' \in \text{Suiv}(s)} [DES(s') + NIV(s') - NIV(s) - 1] \quad (E_2)$$

Avec le parcours par niveaux, on initialise  $DES(s)$  à zéro pour les sommets du dernier niveau, qui n'ont pas de suivant, toutes les autres valeurs de  $DES$  étant à  $+\infty$ .

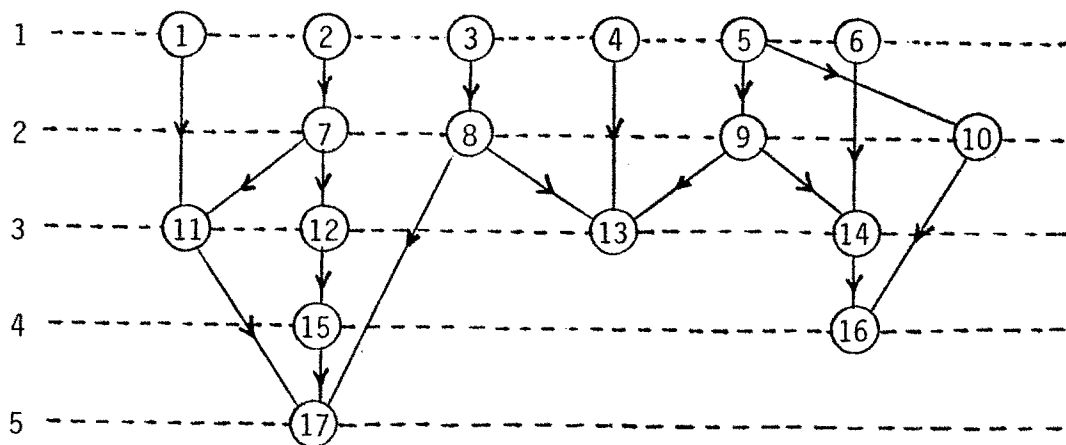
Au niveau  $k$ , si un sommet  $s$  n'a pas de suivant on lui applique  $(E_1)$ , sinon  $DES(s)$  est déjà fixé.

Ensuite, on change éventuellement les valeurs  $DES(s')$  de ses prédécesseurs :

$$DES(s') = \min [DES(s'), DES(s) + NIV(s) - NIV(s') - 1]$$

Si  $DES(s')$  est déjà fixé à une valeur plus basse, on ne la change pas, sinon, quand tous les sommets  $s$  des niveaux  $\geq k$  ont été parcourus,  $DES(s')$  contient la valeur qui correspond à la formule  $(E_2)$ .

Exemple : pour le graphe suivant, on a 17 sommets sur 5 niveaux :



- Figure 4 -

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
DES	2	0	2	3	1	2	0	2	1	2	1	0	2	1	0	1	0

Les sommets 2, 7, 12, 15 et 17 sont sur le chemin maximum, ils sont donc déjà en place, et pour eux,  $DES(s) = 0$ .

## 2) Modifications du tableau DES

D'après les valeurs actuellement présentes dans le tableau, certains sommets pourraient descendre d'un plus grand nombre de niveaux qu'un de leurs précédents, ce qui allongerait des arcs.

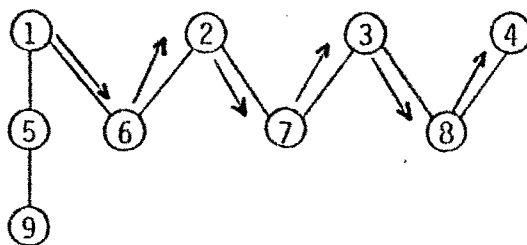
Par exemple :  $DES(9) = 1 < DES(13) = 2$ .

Pour éviter celà, il faut refaire un parcours par niveau en partant du haut et aligner  $DES(s)$  sur :

$$\text{Min}_{s' \in \text{Préc}(s)} [DES(s')]$$

Mais celà ne suffit pas. En effet, si on impose  $DES(13) = 1$ , par exemple, il faudra transmettre la modification aux précédents de 13 autres que 9, soit :  $DES(4) = 3 - 1 = 2$  et  $DES(8) = 2 - 1 = 1$ .

On voit ainsi apparaître la nécessité de parcourir tous les "zigzags" du graphe pour ajuster DES au minimum :



- Figure 5 -



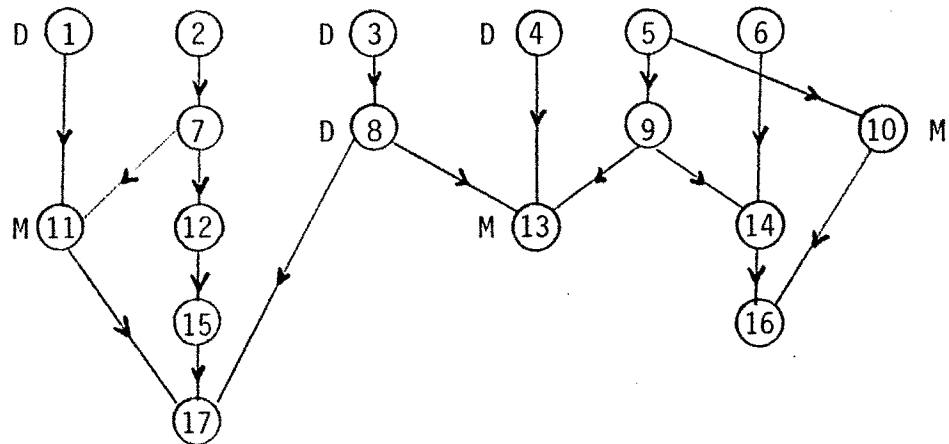
Sur le graphe figure 5 , on a :  $DES(1) = DES(5) = DES(9) = 0$ , toutes les autres valeurs étant à 1.

Il va donc falloir successivement imposer :

$DES(6) = 0$  puisque  $DES(1) = 0$ , puis transmettre la modification le long du "zigzag" :  $DES(2) = 0$ ,  $DES(7) = 0$  etc ... jusqu'à :  $DES(4) = 0$ .

Cette modification se fera par une succession de parcours par niveaux alternativement descendants et ascendants, en marquant à chaque parcours tous les sommets pour lesquels une modification de DES a eu lieu, jusqu'à ce que toutes les modifications nécessaires aient été faites.

Exemple : sur le graphe de la figure 4, on a, en notant D les marques du parcours descendant, et M les marques du parcours montant :



- Figure 6 -

Au niveau 1 : on n'a aucune modification à faire (pas de précédent )

Au niveau 2 :  $DES(10) = 2$ , or,  $DES(5) = 1$ , on modifie donc :  $DES(10) = 1$   
et on marque 10 avec M (pour transmettre la modification vers le haut).

Comme  $DES(16) = 1$ , la transmission vers le bas s'arrête.

On a aussi :  $DES(17) = 0$  et  $DES(11) = 1 \Rightarrow DES(11) = 0$

$DES(9) = 1$  et  $DES(13) = 2 \Rightarrow DES(13) = 1$

A la fin du premier parcours descendant, on a donc 3 sommets s marqués M à partir desquels on va recalculer  $DES(s')$  pour leurs précédents s' par la formule utilisée à l'étape 1 :

$$DES(s') = \text{Min} [DES(s'), DES(s) + NIV(s) - NIV(s') - 1]$$

Si une modification est effectivement réalisée par ce nouveau calcul, on marque D le sommet s' au cours du parcours ascendant pour transmettre ensuite la modification vers le bas etc ...

Ici, le parcours ascendant impose :  $DES(1) = 1$ ,  $DES(3) = DES(8) = 1$   
et  $DES(4) = 2$ .

Le parcours descendant suivant ne provoquant plus de modifications, on peut arrêter l'algorithme et calculer les positions définitives.

On a :

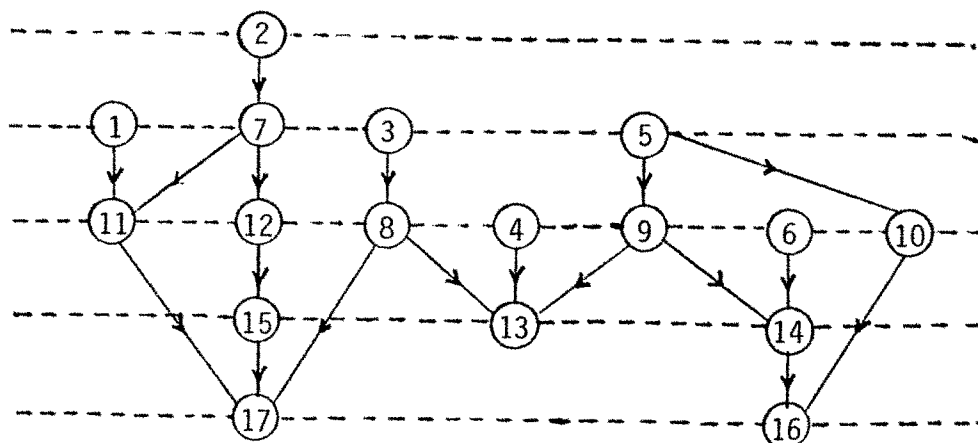
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
DES	1	0	1	2	1	2	0	1	1	1	0	0	1	1	0	1	0

### 3) Positions définitives

En effectuant pour tout sommet  $s$  l'opération  $NIV(s) := NIV(s) + DES(s)$  on obtient le tableau des niveaux :

2	1	2	3	2	3	2	3	3	3	3	3	4	4	4	5	5
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Soit finalement le dessin :



- Figure 7 -

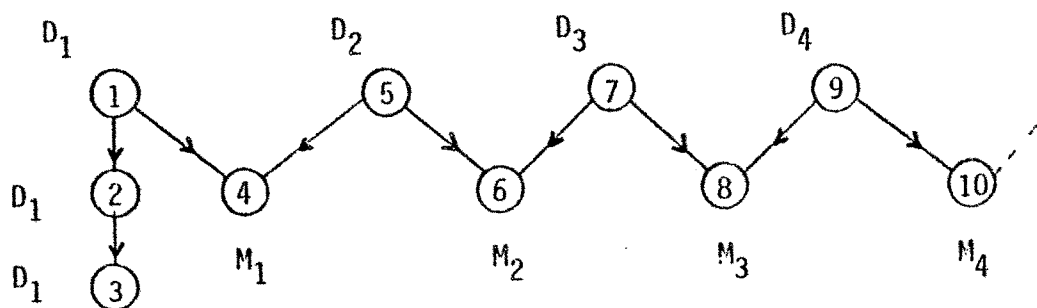
Les propriétés énoncées plus haut découlent du fonctionnement de l'algorithme : on déplace au maximum vers le bas tout sommet  $i$ , sauf quand il a un précédent  $j$  qui ne peut pas descendre autant. En effet, dans ce cas on allongerait l'arc  $(j,i)$ . Toutes les arêtes ont donc une portée minimale, sauf celles qui sont issues des sommets flottants, pour la raison qui vient d'être donnée : raccourcir l'arc  $(11,17)$  sur la figure 7 allongerait l'arc  $(7, 11)$ . Ceci est également valable pour les sommets

qui sont "flottants" par rapport à un chemin non maximum, comme le sommet 10 par rapport au chemin 5 - 9 - 14 - 16. De même, une fois que l'on a bloqué les sommets flottants le plus haut possible, toutes les positions des sommets sont définies sans ambiguïté, d'où l'unicité de la hiérarchie serrée, les sommets flottants (par rapport à un chemin maximum ou non) étant les mêmes dans la hiérarchie inverse que dans la hiérarchie directe. Pour construire la hiérarchie serrée à partir de la hiérarchie inverse, on utiliserait un algorithme du même genre, modifié de façon à remonter au maximum les sommets flottants.

L'algorithme que nous venons de décrire est en  $n^3$ , comme l'algorithme récursif donné dans Delarche [6].

En effet, chaque parcours par niveau considère au plus une fois chaque arc, que ce soit en montant ou en descendant. Les parcours par niveaux sont donc en  $n^2$ .

Or, on a à faire moins de  $n$  parcours, puisque le nombre de parcours est déterminé par le plus long "zigzag" existant dans le graphe, le cas le plus défavorable étant le graphe de la figure 8 :



- Figure 8 -

Les marques D et M ont été indicées par le numéro de parcours correspondant (descendant ou montant).

Le nombre de parcours à effectuer est donc ici  $(n - 3)$ , ce qui donne finalement une borne  $n^3 - 3n^2$  pour l'algorithme complet.

Cependant, ce type de graphe est très académique, dans les exemples réels que nous avons traités, le nombre de parcours à effectuer n'a jamais dépassé 3.

Remarque : Le graphe de la Figure 8 est au contraire idéal pour un parcours doublement récursif, puisque tout le zigzag 1 - 4 - 5 - 6 - 7 ... est parcouru d'un coup.

#### 3.4. Représentation réduite éclatée d'un graphe orienté

Nous allons étendre la hiérarchie serrée aux graphes non fortement connexes de la manière suivante :

\* Les sommets du graphe réduit qui représentent des sommets isolés sont laissés tels quels.

\* Dans une c.f.c. non réduite à un seul sommet, tous les sommets sont au même niveau. Ce niveau est le niveau auquel appartient, (dans la hiérarchie serrée du graphe réduit) le sommet qui représente la c.f.c. considérée. La c.f.c. est "éclatée" et cet éclatement est symbolisé par un trait horizontal reliant entre eux les sommets de la c.f.c. sur le niveau commun.

Cette représentation a l'intérêt de bien schématiser visuellement l'existence des arcs internes à chaque c.f.c. (voir [4]). Elle implique

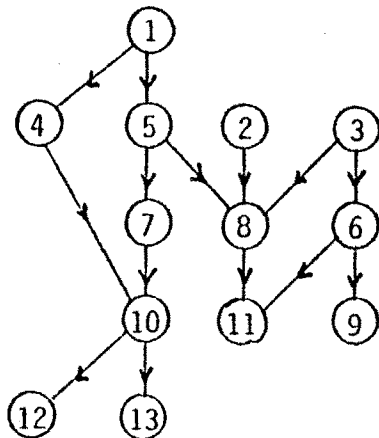
qu'aucun sommet extérieur à la c.f.c. ne soit intercalé parmi les sommets de cette c.f.c., sous peine d'être pris dans la chaîne horizontale représentant les arcs internes.

Nous appellerons *représentation réduite éclatée* le dessin hiérarchisé ayant ces deux caractéristiques :

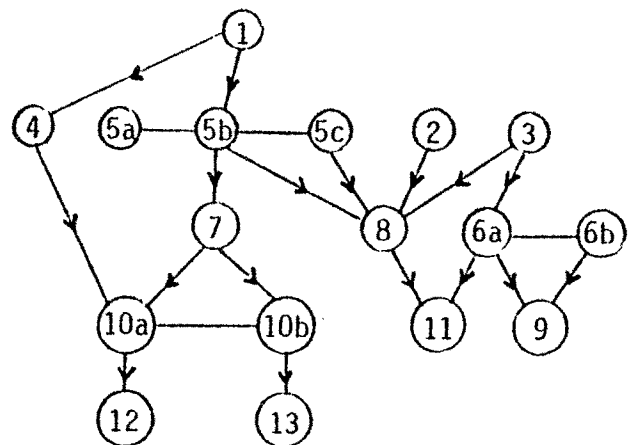
- la hiérarchie adoptée est la hiérarchie serrée sur le graphe réduit,
- les c.f.c. sont symbolisées à chaque niveau par un trait horizontal unissant tous les sommets d'une même c.f.c.

Sur cette représentation standard (*unique* pour un graphe orienté donné) sera définie au chapitre suivant une méthode heuristique de placement des sommets pour le dessin automatique.

Exemple : Soit le graphe de la figure 3 considéré comme graphe réduit d'un graphe G où le sommet 5 représente une c.f.c. à 3 sommets et les sommets 6 et 10 des c.f.c. à 2 sommets :



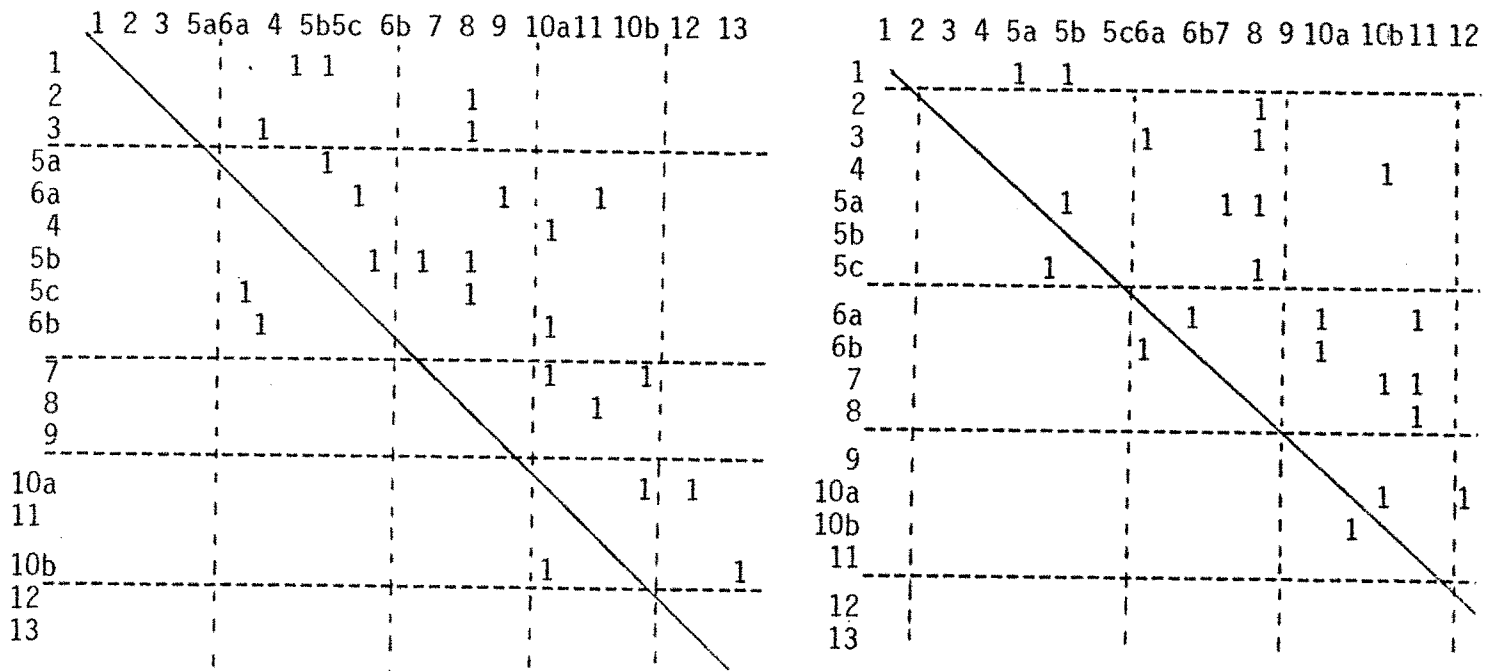
Graphe réduit en hiérarchie serrée



Représentation réduite éclatée

Nous pouvons interpréter la représentation réduite éclatée sur la matrice du graphe ; si on classe les c.f.c. (et donc les sommets) par ordre de rang croissant et si on transmute la matrice (permutation simultanée des lignes et des colonnes correspondantes), on obtient une matrice bloc-triangulaire supérieure puisque l'ordre des rangs suit l'orientation des arcs :  $\text{rang}(s_i) > \text{rang}(s_j) \Rightarrow$  pas d'arcs de  $i$  vers  $j$ .

On donne ci-dessous les deux matrices suivantes : à gauche la matrice (avec un rangement des lignes et des colonnes respectant l'ordre des rangs) du graphe de la figure 9 (en considérant les c.f.c. comme des circuits) ; et à droite, la même matrice avec un rangement qui tient compte de la représentation réduite éclatée (les cases blanches sont des zéros).



- Figure 10 -

La représentation réduite éclatée correspond à une amélioration de la concentration des 1 autour de la diagonale : le raccourcissement des arcs fait glisser des 1 vers le bas et vers la droite dans le triangle sur-diagonal (par exemple l'arc (6a, 11)), ce qui les rapproche de la diagonale. Le regroupement des sommets par c.f.c. fait apparaître des sous-blocs diagonaux où sont réunis des 1, qui, auparavant, pouvaient être n'importe où dans les blocs diagonaux (par exemple l'arc (6b, 6a)).

#### 4. REPRESENTATIONS CENTRIFUGES

Nous allons maintenant définir une deuxième méthode de hiérarchisation et de représentation pour des graphes *non orientés* connexes. Au lieu du rang et des composantes fortement connexes qui caractérisaient la représentation réduite éclatée, les notions de base sont ici l'écartement et les centres du graphe, d'où les noms de hiérarchie des écartements et représentation centrifuge.

Nous étudierons d'abord le cas des graphes non orientés, puis nous construirons une hiérarchie directe et une hiérarchie inverse dans le cas de graphes orientés *fortement connexes*.

##### 4.1. Graphes non orientés connexes

###### a) Définitions

Soit  $G = (X, U)$  un graphe non orienté connexe à  $n$  sommets, on appelle :

- *Distance du sommet  $x$  au sommet  $y$*  : le nombre d'arêtes d'une plus courte chaîne de  $x$  à  $y$ .

On la note :  $d(x, y)$ , et on pose :  $d(x, x) = 0$ .

- *Ecartement du sommet  $x$*  :  $e(x) = \text{Max}_{y \in X} d(x, y)$



- *Centre de G* : tout sommet d'écartement minimum  $r$  ;  $r$  est appelé le *rayon* du graphe.
- *Diamètre de G* : (noté  $\text{diam}(G)$ ) : la longueur de la plus longue des courtes chaînes dans  $G$ .  $\text{diam}(G) = \text{Max}_{x \in X} (e(x))$ .
- *Sommet périphérique* : tout sommet  $p$  tel que :  $e(p) = \text{diam}(G)$ . Comme  $G$  est connexe,  $r$  et  $\text{diam}(G)$  existent, puisque toutes les distances  $d(x,y)$  existent.

Dans le cas contraire, on poserait  $d(x,y) = +\infty$  pour des sommets  $x$  et  $y$  appartenant à des composantes connexes différentes.

#### b) Propriétés

##### 1. $d(x,y)$ est une distance :

En effet :

$$\begin{cases} d(x,x) = 0 \\ d(x,y) = d(y,x) \\ d(x,y) \leq d(x,z) + d(z,y) \end{cases}$$

(l'égalité étant atteinte pour les sommets  $z$  où passent les plus courtes chaînes de  $x$  à  $y$ ),

##### 2. Un sommet périphérique n'est jamais point d'articulation

Démonstration : Soit  $p$  un tel sommet, s'il est point d'articulation de  $G$ , le graphe  $G' = (X',U')$  où  $X' = X - \{p\}$  est formé de deux composantes connexes  $G_1$  et  $G_2$  (au moins).

L'une de ces deux composantes contient au moins un sommet  $s'$  tel que  $d(s',p) = \text{diam}(G)$ .

Alors il existe un sommet  $s''$  dans l'autre composante tel que :  $d(p,s'') \geq 1$ .

$$\text{D'où, } d(s', s'') = d(s', p) + d(p, s'') \geq \text{diam}(G) + 1$$

Ceci est en contradiction avec l'hypothèse que  $p$  est d'écart maximum.  
(Pour d'autres propriétés classiques liées à ces notions, voir Roy [20] et Harary [81]).

#### 4.2. Hiérarchie des écartements

a) Définition : Soit  $P_e$  la relation d'ordre sur les sommets du graphe  $G$

$$x_i P_e x_j \iff e(x_i) \leq e(x_j)$$

Alors l'application  $H_e$  qui à tout  $x_i$  fait correspondre l'entier  $n_i$

$$x_i \xrightarrow{H_e} n_i = e(x_i) - r + 1$$

est la hiérarchie associée à la relation d'ordre  $P_e$  sur les écartements des sommets du graphe  $G$ .

Cette hiérarchie est dite *hiérarchie des écartements*.

#### b) Propriétés

1.  $H_e$  est unique
2. Les sommets de niveau 1 sont les centres du graphe  $G$ .
3. Le nombre de niveaux est  $\text{diam}(G) - r + 1$ , et les sommets du dernier niveau sont les sommets périphériques
4. Toutes les arêtes sont de portée au plus égale à 1 dans la hiérarchie

Démonstration : s'il existait une arête  $(a, b)$  de portée 2, on aurait (par exemple)  $n_a = k$  et  $n_b = k + 2$ .

$$\text{soit : } e(a) - r + 1 = k$$

$$\text{et : } e(b) - r + 1 = k + 2$$

soit encore :  $e(b) - e(a) = 2$  alors que pour tout  $s$  de  $X$ ,  
 $d(b,s) \leq d(b,a) + d(a,s) = 1 + d(a,s)$  donc :  $e(b) \leq 1 + e(a)$  et on  
aboutit à une contradiction.

On voit ainsi qu'il n'y a pas à envisager de "resserrage" de la hiérarchie des écartements non orientée.

Par ailleurs, il peut exister des arêtes entre des sommets situés au même niveau dans cette hiérarchie.

#### 4.3. Graphes orientés fortement connexes

##### a) Définitions

Soit  $G = (X,V)$  un graphe orienté à  $n$  sommets, on appelle :

- *Ecart du sommet  $x$  au sommet  $y$*  : le nombre d'arcs d'un plus court chemin de  $x$  à  $y$ .

On le note ;  $e(x,y)$  et on pose  $e(x,x) = 0$ . S'il n'y a pas de chemin de  $x$  à  $y$ ,  $e(x,y) = +\infty$ . Généralement,  $e(x,y)$  et  $e(y,x)$  auront des valeurs *différentes*.

Dans le cas où  $G$  est *fortement connexe*, les valeurs  $e(x,y)$  sont définies pour tous les couples  $(x,y)$  possibles, on va donc définir deux sortes d'écartement pour les sommets de  $G$ , suivant que l'on considère un sommet donné  $s$  du point de vue de l'*émission* ou de la *réception*, c'est-à-dire du point de vue des plus courts chemins *entrants* ou *sortants* ; on les indicera par  $E$  (émission) ou  $R$  (réception), ainsi on appellera :

$E$  - écartement du sommet  $x$  :  $E_E(x) = \text{Max}_{y \in X} e(x,y)$

R - écartement du sommet  $x$  :  $E_R(x) = \text{Max}_{y \in X} e(y,x)$

E - centre de  $G$  : tout sommet  $s$  de E-écartement minimum  $R_E$ ,  
 $s$  est un E-centre de  $G \iff E_E(s) = \text{Min}_{x \in X} [\text{Max}_{y \in X} e(x,y)]$ .

$R_E$  sera le E-rayon de  $G$ .

R - centre de  $G$  : tout sommet de  $s$  de R-écartement minimum  $R_R$

$s$  est un R-centre de  $G \iff E_R(s) = \text{Min}_{x \in X} [\text{Max}_{y \in X} e(x,y)]$ .

$R_R$  sera le R-rayon de  $G$ .

Un R-centre est un E-centre du même graphe  $G$  où tous les arcs ont été changés.

Ces définitions correspondent aux termes connus d'écartement et anti-écartement, centre et anticentre, mais nous préférons réserver les appellations de centre et d'écartement au cas des graphes *non-orientés* (§ 1). Les écartements maxima donnent le E-diamètre et le R-diamètre de  $G$ .

#### b) Propriétés

1. L'écart n'est pas une distance :  $e(x,y) \neq e(y,x)$

2. Sur la matrice des plus courts chemins de  $G$ ,  $E_E(s)$  est le max des éléments de la  $s^{\text{ème}}$  ligne,  $E_R(s)$  le max des éléments de la  $s^{\text{ème}}$  colonne.

3. Un centre d'un graphe connexe non orienté peut être considéré comme étant à la fois un E-centre et R-centre du graphe orienté obtenu en remplaçant chaque arête par une paire d'arcs opposés : la matrice d'incidence est alors symétrique, la matrice des plus courts chemins l'est aussi et devient la matrice des plus courtes chaînes.

4. Le E-diamètre et le R-diamètre sont égaux, puisqu'ils sont tous deux égaux à la plus grande valeur contenue dans la matrice des plus courts chemins.

5. Le E-rayon et le R-rayon sont généralement différents.
6. Un sommet E-périphérique ou R-périphérique n'est jamais point d'articulation (démonstration semblable au cas des graphes non orientés).

#### 4.4. Hiérarchies des écartements directes et inverses

##### a) Définition

Dans un graphe orienté fortement connexe, on peut construire deux hiérarchies des écartements orientées : l'une à partir des E-centres et des E-écartements, l'autre à partir des R-centres et R-écartements

Par analogie avec la hiérarchie des rangs, et du fait de la correspondance entre ces notions par changement d'orientation des arcs (cf. Supra) nous les appellerons respectivement hiérarchie directe et hiérarchie inverse plutôt que hiérarchie des E-écartements et des R-écartements.

On a donc deux formules de calcul des niveaux :

Pour la hiérarchie directe :

$$x_i \xrightarrow{H_E} n_i = E_E(x_i) - R_E + 1$$

Pour la hiérarchie inverse :

$$x_i \xrightarrow{H_R} n_i = E_R(x_i) - R_R + 1$$

Les relations d'ordre associées sont respectivement  $P_E$  et  $P_R$ , relations d'ordre entre les E-écartements et R-écartements :

$$x_i P_E x_j \iff E_E(x_i) \leq E_E(x_j)$$

$$x_i P_R x_j \iff E_R(x_i) \leq E_R(x_j)$$

### b) Propriétés

1. Dans la hiérarchie directe (resp. inverse) les sommets de niveau 1 sont les E-centres (resp. R-centres) de G.
2. Comme en général :  $R_E(G) \neq R_R(G)$ , le nombre des niveaux est différent dans les deux hiérarchies. Il s'agit donc bien de deux hiérarchies *distinctes* et on ne peut envisager de passer à une "hiérarchie serrée des écartements" compatible avec chacune d'elles.
3. Dans la hiérarchie directe (resp. inverse) tous les arcs entre des niveaux non consécutifs (shunts) sont centrifuges (resp. centripètes). Ceci est nécessaire pour respecter l'ordre des écarts.

Dans la hiérarchie directe, supposons qu'il existe un shunt centripète allant du sommet a du niveau i au sommet b du niveau i - 2 ; alors  $E(a,b) : 1$  et comme  $E_E(b) = i - 2 + (R_E - 1)$ , on a :

$E_E(a) = e(a,b) + E_E(b) = i - 1 + (R_E - 1)$  ce qui est contradictoire avec l'appartenance du sommet a au niveau i.

#### 4.5. Représentation centrifuge standard

Il nous faut choisir une forme de représentation des graphes en hiérarchie des écartements qui fasse clairement apparaître la structure du graphe dans sa décomposition en niveaux, comme le tracé des droites parallèles pour la hiérarchie des rangs.

On ne peut adopter le même dessin pour les hiérarchies centrifuges : d'une part, il n'y a pas ici de liens directs entre la position des sommets sur les niveaux et l'éventuelle orientation des arcs ; or c'était cette caractéristique essentielle de la hiérarchie des rangs qui nous avait amené naturellement à une représentation "en cascade" des graphes sans circuit.

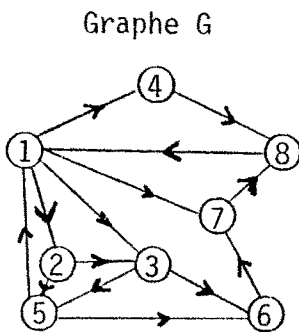
D'autre part, on a intérêt à choisir un type de dessin différent, pour éviter toute confusion en cas d'utilisation simultanée des deux méthodes de hiérarchisation (voir le chapitre 6).

Nous avons donc dessiné les hiérarchies des écartements en prolongeant la métaphore géométrique (centres, rayon, diamètre, ...) . Les niveaux sont symbolisés par des *cercles concentriques*, le plus petit portant les centres du graphedessiné, d'où le nom de *représentation centrifuge*.

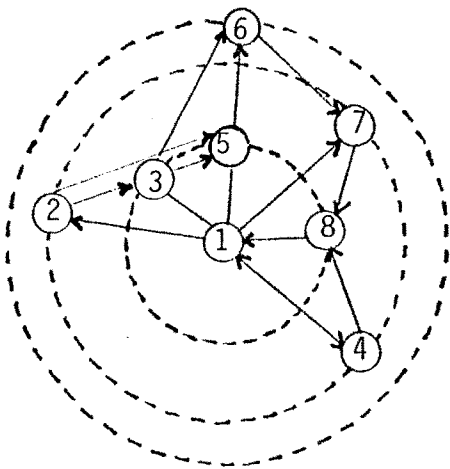
Si un seul sommet est un centre, on le place ... au centre du dessin.

#### 4.6. Exemple d'application

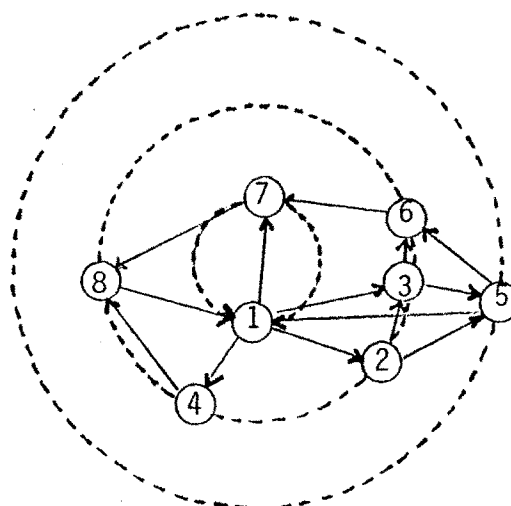
Nous allons dessiner les trois hiérarchies centrifuges pour un graphe fortement connexe à 8 sommets, G.



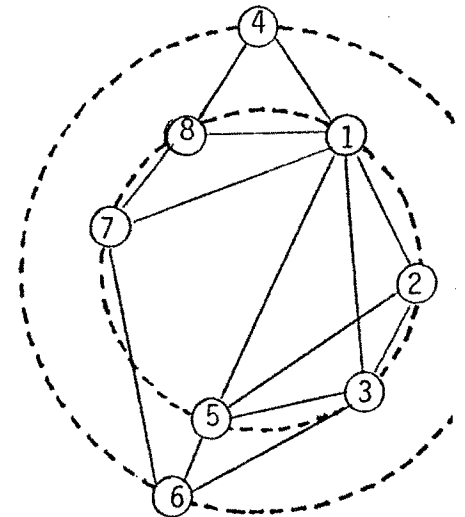
Matrice des plus courts chemins		$E_E$	Matrice des plus courtes chaînes															
0	1	1	1	2	2	1	1	1	2	1	1							
2	0	1	3	1	2	3	4	4	1	0	1	2	1	2	2			
2	3	0	3	1	1	2	3	3	1	1	0	2	1	1	2	2		
2	3	3	0	4	4	3	1	4	1	2	2	0	2	3	2	1		
1	2	2	2	0	1	2	3	3	1	1	1	2	0	1	2	2		
3	4	4	4	5	0	1	2	5	2	2	2	1	3	1	0	1	2	
2	3	3	3	4	4	0	1	4	1	1	2	2	2	2	1	0	1	
1	2	2	2	3	3	2	0	3	0	1	2	2	1	2	2	1	0	
$E_R$	3	4	4	4	5	4	3	4		$E$	2	2	2	3	2	3	2	2



Hiérarchie directe des écartements



Hiérarchie inverse des écartements



Hiérarchie non orienté des écartements

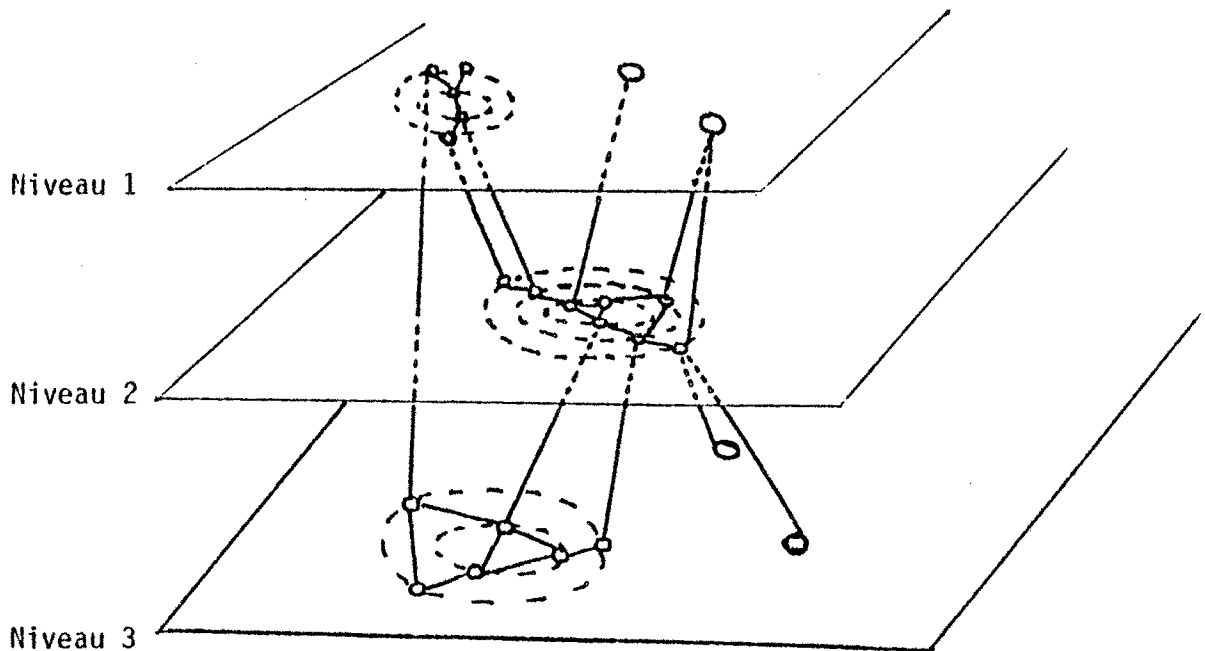
On a :  $R = R_E = 2$  ;  $R_R = 3$  ;  $\text{diam}(G) = 3$  ;  $E\text{-diam}(G) = R\text{-diam}(G) = 5$  :

La hiérarchie directe a 4 niveaux, la hiérarchie inverse 3 niveaux et la hiérarchie non orientée seulement 2 niveaux.

Suivant l'intérêt qu'elles présentent pour chaque graphe considéré, on pourra choisir telle ou telle de ces trois hiérarchies.

Remarque :

On peut illustrer la complémentarité des deux méthodes de hiérarchisation en considérant la représentation réduite éclatée et les représentations centrifuges des c.f.c. respectivement comme la projection sur un plan vertical du graphe global et les projections sur des plans horizontaux de chacune des c.f.c., l'ensemble étant construit en trois dimensions:



- Figure 12 -





## CHAPITRE III

### DESSIN AUTOMATIQUE DES GRAPHES HIERARCHISES

#### 1. PRESENTATION DES PROBLEMES

Au chapitre précédent nous avons défini la forme sous laquelle nous allons représenter chacune des deux structures hiérarchiques utilisées.

Pour dessiner automatiquement sur écran graphique ou sur papier les graphes ainsi hiérarchisés, il faut disposer d'un logiciel qui calcule pour chaque sommet un couple  $(x,y)$  de coordonnées cartésiennes dans le plan.

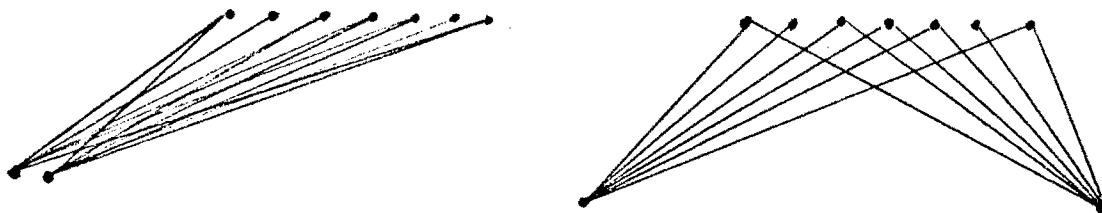
D'une part, il faut que ce placement des sommets respecte les contraintes de disposition inhérentes aux formes standard de représentation retenues : la forme réduite éclatée ou la forme centrifuge. D'autre part, le dessin obtenu doit être aussi satisfaisant que possible sur le plan de la lisibilité.

Généralement, on aura un grand nombre de solutions compatibles avec les contraintes hiérarchiques. Le programme devra nous en fournir une qui soit visuellement agréable, c'est-à-dire :

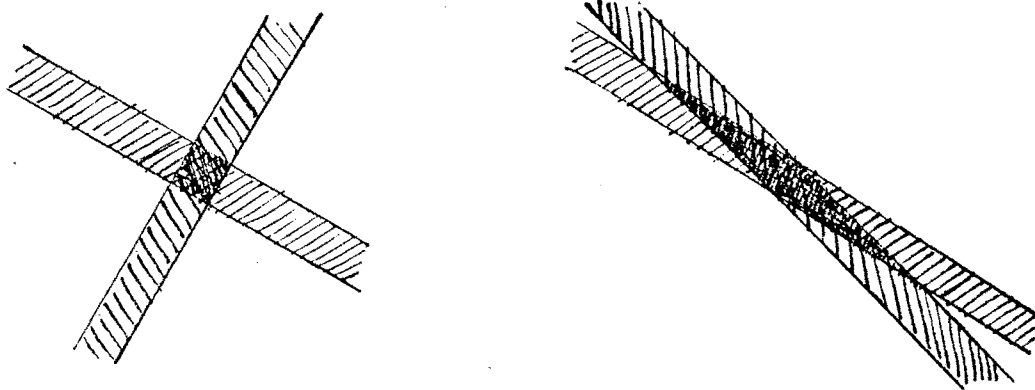
1\* donnant une répartition homogène des sommets dans l'espace disponible.

2\* comportant peu de croisements d'arcs.

Il n'y a pas de définition objective du "meilleur dessin" car la minimisation du nombre de croisements n'est pas un critère absolu du fait que la perception visuelle des croisements est d'autant plus néfaste pour la lisibilité du graphe que les arcs qui se croisent ont des directions voisines :



Il y a deux raisons complémentaires à celà : on perçoit moins bien la différence d'orientation des faisceaux d'arcs presque parallèles ; et du fait que les arcs ont en pratique une certaine épaisseur, la surface du croisement augmente, ce qui accroît son importance visuelle :



Comme d'autre part le dessin ayant le moins de croisements ne pourrait être trouvé qu'en analysant un nombre prohibitif de configurations, on devra se contenter de définir une heuristique de placement à la fois efficace et rapide.

Pour chacune des hiérarchisations précédemment définie nous avons essayé de construire des méthodes heuristiques adaptées aux représentations standard et qui permettent d'obtenir en des temps acceptables des images suffisamment lisibles pour qu'on puisse envisager de les corriger en déplaçant quelques sommets du graphe sur l'écran de visualisation. (mise en oeuvre de l'interactivité offerte par ce type de terminal).

### 1.1. Le problème général de la planarisation

Savoir si un graphe est planaire, c'est-à-dire susceptible d'être dessiné dans le plan sans que ses arêtes se coupent est un problème bien connu de la théorie des graphes.

Hopcroft & Tarjan [9] ont donné, en 1974, un algorithme optimal pour tester si un graphe est planaire ou non.

Si oui, on peut déduire de l'algorithme une représentation planaire du graphe, mais avec des arêtes non rectilignes.

Dans le cas des graphes non planaires, le problème de la minimisation du nombre de croisements (que les arêtes soient rectilignes ou non) reste ouvert.

Sur le plan pratique, le problème a donné lieu à de nombreuses études, vu son importance, en particulier pour la conception des circuits électroniques. Il existe un certain nombre de méthodes heuristiques fondées sur des hypothèses plus ou moins restrictives qui visent à respecter les contraintes pratiques de fabrication. Par exemple, on peut chercher à partitionner l'ensemble des sommets en différents groupes de façon à minimiser le nombre des arcs qui relient ces groupes entre eux (Donath et Hoffmann [7]) Mais les contraintes spécifiques du dessin de circuits électroniques rendent ces heuristiques inintéressantes pour notre problème de dessin de graphes hiérarchisés.

## 1.2. Problèmes spécifiques des graphes hiérarchisés

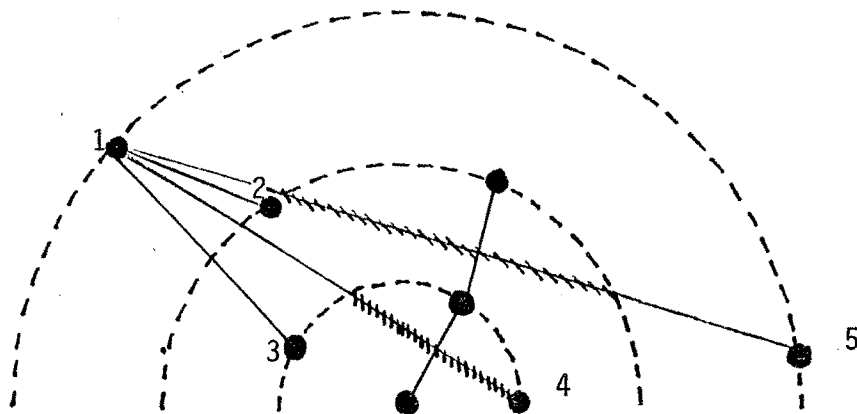
Pour nos graphes hiérarchisés, le problème est à la fois plus simple et plus compliqué.

Plus simple, parce que l'obligation de respecter certaines contraintes de disposition (en particulier, la forme des niveaux hiérarchiques) diminue la part d'arbitraire a priori sur le placement des sommets.

Plus compliqué, parce que ces contraintes de disposition, ne se formalisent pas toutes théoriquement (par exemple, la représentation des c.f.c. dans la forme réduite éclatée, ou bien la nécessité de maintenir une distance minimale de séparation entre les sommets), vu la complexité nouvelle qu'elles introduisent, voire éventuellement leur caractère subjectif puisque variable d'un utilisateur à l'autre.

Nous nous sommes intéressés surtout au problème de la représentation en niveaux rectilignes parallèles, pour deux raisons :

1\* la forme circulaire des niveaux ne donne pas lieu à une formalisation mathématique simple du problème des croisements car un arc rectiligne joignant deux sommets du même niveau ou de deux niveaux voisins coupera, s'il est presque diamétral, d'autres intervalles que celui qui correspond à l'écart des niveaux de ses extrémités. Ceci précisément à cause de la courbure des cercles :



- Figure 13 -

L'arc 1-4 et l'arc 1-5 passent dans des zones (parties hachurées des arcs) qui ne correspondent pas aux niveaux de leurs extrémités.

Aucune étude strictement combinatoire n'est donc possible, puisque le problème des croisements est lié aux tailles relatives des cercles de niveau, et donc aux contraintes matérielles de représentation.

2\* avec des niveaux rectilignes parallèles, une étude combinatoire est possible dans un certain nombre de cas particuliers, d'ailleurs intéressants en eux-mêmes.

En effet, dans ce cas, le problème de la minimisation des croisements est un problème de détermination de "bonnes" permutations des sommets sur les niveaux, indépendamment du tracé des niveaux (plus ou moins serrés, espacés régulièrement ou non etc ...).

Nous allons étudier le problème pour trois classes de graphes :

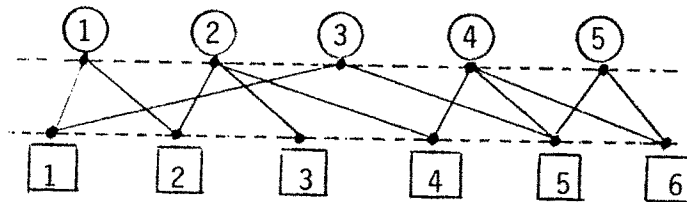
- 1 - les graphes bipartis
- 2 - les graphes multipartis
- 3 - les graphes sous forme réduite éclatée.

La méthode que nous avons utilisée débouche sur une formalisation mathématique intéressante dans le cas des graphes bipartis, et susceptibles de généralisation (au moins sous forme d'une heuristique) dans le cas des graphes multipartis.

Par ailleurs, elle fournit un outil heuristique approprié pour la visualisation en deux phases (graphe réduit et graphe éclaté) des graphes mis sous la forme réduite éclatée.

## 2. LA MINIMISATION DU NOMBRE DE CROISEMENTS D'ARETES POUR DES GRAPHES BIPARTIS CONNEXES

Parmi les graphes hiérarchisés, les graphes bipartis sont les plus simples possibles puisqu'ils n'ont que deux niveaux :



- Figure 14 -

### Remarque :

L'orientation éventuelle des arcs n'entre pas en ligne de compte pour les graphes bipartis qu'on peut définir de manière équivalente comme des graphes 2-coloriables.

En pratique, on rencontre souvent des graphes de ce type, pour lesquels les niveaux sont définis par une référence sémantique extérieure à la représentation : graphes objectifs-moyens, matériel-composants, travaux-tâches élémentaires etc ...

Dans tous ces cas, la connaissance a priori des deux classes de sommets incite à donner une représentation du type de la figure 14.

Cette catégorie de graphes a donc un intérêt spécifique, indépendamment de la simplification qu'elle apporte dans l'étude du problème des croisements d'arcs.

2.1. Dénombrément des croisements d'arêtes

a) Matrice d'incidence d'un graphe biparti

Un graphe biparti étant caractérisé par la répartition de ses sommets en deux classes X et Y où :

$$X = \{X_i, i = 1, 2, \dots, k\}$$

$$Y = \{Y_j, j = 1, 2, \dots, m\}$$

On le note :  $G(X, Y, U)$  et on appelle matrice d'incidence du graphe biparti G la matrice rectangulaire booléenne à k lignes et m colonnes A telle que :

$$a_{ij} = 1 \text{ si } X_i \text{ est lié à } Y_j \text{ par un arc de } U.$$

$$a_{ij} = 0 \text{ sinon}$$

b) Dessin du graphe

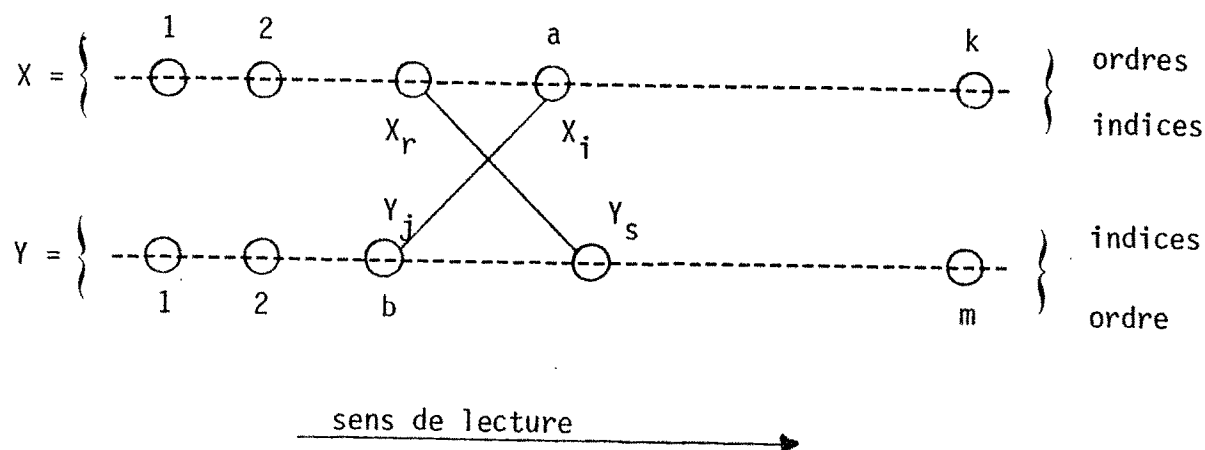
Pour représenter un graphe biparti par un *dessin*, nous le mettons sous la forme de la figure 14 en rangeant les deux classes X et Y sur deux segments de droite parallèles. Par convention, X est l'ensemble des sommets du haut et Y l'ensemble des sommets du bas.

Chaque façon de dessiner le graphe est ainsi définie par un couple  $(p_1, p_2)$  de permutations respectivement sur  $\{1, 2, \dots, k\}$  et  $\{1, 2, \dots, m\}$  en interprétant l'ordre des sommets de gauche à droite sur le dessin à l'aide de permutations sur l'ensemble des indices de ces sommets.

$p_1(a) = i \iff X_i$  est le  $a^{\text{ème}}$  sommet du haut à partir de la gauche sur le dessin ;

$p_2(b) = j \iff Y_j$  est le  $b^{\text{ème}}$  sommet du bas à partir de la gauche sur le dessin.





- Figure 15 -

Compter les croisements revient à chercher pour toute arête  $(X_i, Y_j)$  toutes les arêtes  $(X_r, Y_s)$  telles que :

$$(p_1^{-1}(r) < p_1^{-1}(i)) \quad \text{et} \quad (p_2^{-1}(s) > p_2^{-1}(j))$$

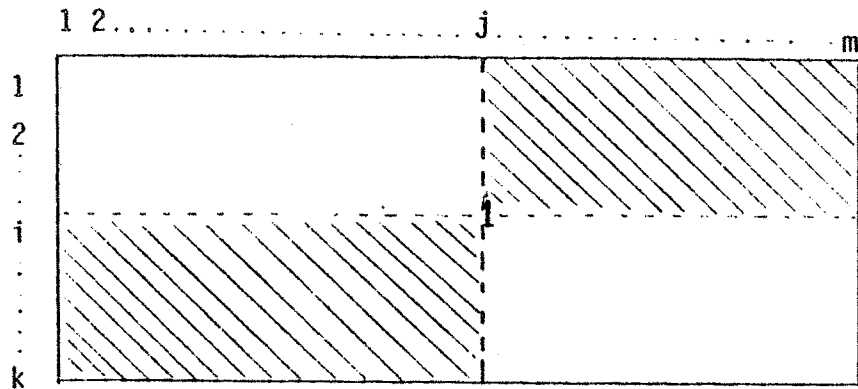
ou :  $(p_1^{-1}(r) > p_1^{-1}(i)) \quad \text{et} \quad (p_2^{-1}(s) < p_2^{-1}(j))$

Soit  $A' = A(p_1, p_2)$  la matrice obtenue en réarrangeant selon  $p_1$  et  $p_2$  les lignes et les colonnes de la matrice  $A$  définie précédemment, le nombre de croisements d'arêtes du dessin défini par  $(p_1, p_2)$  est obtenu en parcourant tous les  $A'_{ij}$  égaux à 1 et en faisant la somme cumulée de tous les 1 contenus au dessus et à droite (ou en dessous et à gauche) de  $A'_{i,j}$ .

Ce qui correspond aux formules :

$$x(p_1, p_2) = \sum_{i=1}^k \sum_{j=1}^m (a'_{ij} \sum_{r < i} \sum_{s > j} a'_{rs})$$

ou :  $x(p_1, p_2) = \sum_{i=1}^k \sum_{j=1}^m (a'_{ij} \sum_{r > i} \sum_{s < j} a'_{rs})$



Toutes les conventions et notations qui viennent d'être définies sont conservées dans la suite, pour les graphes bipartis.

Ainsi, sur le graphe de la figure 14, le nombre de croisements est égal à 7.

La matrice d'incidence associée est :

	1	2	3	4	5	6
①	1	1	0	0	0	0
②	0	1	1	1	0	0
③	1	0	1	0	1	0
④	0	0	0	1	1	1
⑤	0	0	0	0	1	1

La méthode décrite donne : 4 croisements pour l'arc (3,1), et 1 croisement pour chacun des arcs (3,3) (4,4) et (5,5).

## 2.2. La méthode des degrés relatifs : principe heuristique

Pour essayer de ne pas avoir trop de croisements d'arêtes sur le dessin, une idée naturelle est de chercher, une permutation  $p_1$  étant donnée sur les sommets du haut, à rendre les arcs les plus verticaux possibles, c'est-à-dire à construire une permutation  $p_2$  qui place les sommets du bas le plus possible en dessous des sommets du haut qui leur sont liés.

Pour celà, nous attribuons des abscisses  $x_i \in [0,1]$  aux sommets de  $X$  et calculons des abscisses  $y_j$  pour les sommets de  $Y$  en faisant la moyenne des abscisses de leurs précédents.

Ceci revient à calculer le barycentre  $y$  des abscisses  $x_i$  pondérées par les *degrés relatifs* des sommets  $X_i$  liés au sommet  $Y_j$  considéré ;

Par exemple, pour le graphe de la figure 14, on a :

$$y_1 = \frac{1}{2} (x_1 + x_3)$$

et : 
$$y_5 = \frac{1}{3} (x_3 + x_4 + x_5)$$

etc ...

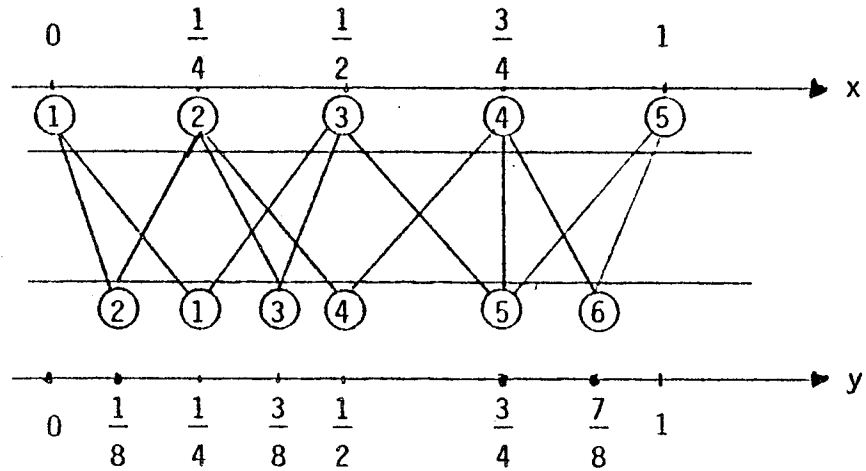
Si on suppose que l'on a :  $x_1 = 0$ ,  $x_2 = \frac{1}{4}$ ,  $x_3 = \frac{1}{2}$ ,  $x_4 = \frac{3}{4}$  et  $x_5 = 1$ ,

les  $y_j$  sont tous compris entre 0 et 1 et valent respectivement :

$$\frac{1}{4}, \frac{1}{8}, \frac{3}{8}, \frac{1}{2}, \frac{3}{4}, \frac{7}{8} .$$

Nous pouvons interpréter les  $x_i$  et les  $y_j$  de façon à obtenir deux permutations  $p_1$  et  $p_2$  respectivement sur les  $X_i$  et les  $Y_j$ , avec

la même convention de lecture que précédemment :



- Figure 16 -

Si nous notons  $\Pi(x)$  l'ordre des composantes du vecteur  $x = (x_i)_{i=1,2,\dots,k}$

et  $\Pi(y)$  l'ordre des composantes du vecteur  $y = (y_j)_{j=1,2,\dots,m}$

on a, pour le dessin de la figure 15 :  $(p_1, p_2) = (\Pi(x), \Pi(y))$

$$p_1 = (1,2,3,4,5) \quad , \quad p_2 = (2,1,3,4,5,6)$$

Remarque 1 :

La convention de lecture choisie est arbitraire ; on pourrait classer les sommets de droite à gauche selon les  $x_i$  et les  $y_j$  croissants le seul point important est d'avoir évidemment la même convention pour les deux niveaux. Suivant l'ordre de rangement on obtient deux couples de permutations symétriques qui sont équivalents du point de vue des croisements d'arêtes, puisque chaque dessin est l'"image-miroir" de l'autre : les dessins sont *identiques à un retournement près*.

Remarque 2 :

Sur l'exemple ci-dessus, on peut noter que le passage de la figure 14 à la figure 16 améliore le dessin , puisque l'on n'a plus que 6 croisements dans le second cas au lieu de 7 dans le premier.

### 2.3. Formalisation itérative

La procédure qui vient d'être décrite pour calculer une permutation  $p_2$  sur les  $Y_j$  à partir d'une permutation  $p_1$  et d'abscisses  $x_i$  attribuées aux  $X_i$  peut être utilisée en sens inverse pour déterminer une permutation sur les  $X_i$  à partir des abscisses  $y_j$ .

D'où l'idée d'envisager une utilisation itérative de cette heuristique des degrés relatifs en reprenant comme point de départ pour les  $y_j$  le résultat du calcul précédent à partir des  $x_i$  :

$$y_1 = \frac{1}{4}, y_2 = \frac{1}{8}, y_3 = \frac{3}{8}, y_4 = \frac{1}{2}, y_5 = \frac{3}{4} \text{ et } y_6 = \frac{7}{8}$$

En calculant les  $x_i$  par la même formule barycentrique, on obtient les nouvelles valeurs :

$$x_1 = \frac{3}{16}, x_2 = \frac{1}{3}, x_3 = \frac{11}{24}, x_4 = \frac{17}{24} \text{ et } x_5 = \frac{13}{16}$$

Contrairement à ce qui s'était produit pour les  $y_j$  lors du premier calcul, le classement des  $X_i$  n'est pas modifié par ces abscisses.

Remarque :

Le fait d'utiliser une formule de calcul "barycentrique" assure que tous les  $x_i$  et tous les  $y_j$  sont toujours dans  $[0,1]$  si les  $x_i$  initiaux sont tous dans  $[0,1]$ .

Nous allons maintenant donner une formalisation matricielle de la méthode des degrés relatifs.

NB.: Dans tout ce qui suit, les indices *supérieurs* des vecteurs  $x$  et  $y$  ne sont pas des exposants, mais des numéros d'itération.

Considérons les deux matrices rectangulaires  $A_Y$  et  $A_X$  suivantes :

-  $A_Y$  est la matrice obtenue à partir de la matrice d'incidence  $A$  en divisant chaque  $a_{ij}$  (qui vaut 0 ou 1) par la somme des éléments de la  $i$ ème ligne.

-  $A_X$  est la matrice obtenue en divisant chaque  $a_{ij}$  par la somme des éléments de la  $j$ ème colonne, puis en transposant la matrice calculée.

$A_Y$  est, comme  $A$ , une matrice de  $\mathcal{M}_{k,m}$  et  $A_X$  est une matrice de  $\mathcal{M}_{m,k}$ .

Calculer les valeurs  $y_j^0$  à partir des valeurs initiales  $x_i^0$  fixées arbitrairement revient à calculer le produit :

$$y^0 = A_X x^0 \text{ où } x^0 = (x_i^0), i = 1, \dots, m$$
$$\text{et } y^0 = (y_j^0), j = 1, \dots, m$$

On calcule ensuite les nouvelles valeurs  $x_i^1$  à partir des  $y_j^0$  en faisant le produit :

$$x^1 = A_Y y^0$$

Puis on peut poursuivre le calcul en faisant un nouvel "aller-retour" entre les niveaux, etc ...

Pour tout  $p \geq 0$ , on a :

$$(E_1) \quad y^p = A_X x^p$$

$$(E_2) \quad x^{p+1} = A_Y y^p$$

Chaque "aller-retour" constitue donc un pas d'itération de la méthode, qui s'écrit :

$$(P_1) \quad x^{p+1} = (A_Y \cdot A_X) x^p = M_1 x^p, \quad p \geq 0$$

$$(P_2) \quad y^{p+1} = (A_X \cdot A_Y) y^p = M_2 y^p, \quad p \geq 0$$

On peut mener le processus itératif sur  $M_1$  (resp.  $M_2$ ) à partir de  $x^0$  (resp.  $y^0$ ) et calculer quand on le souhaite le vecteur itéré  $y^p$  (resp.  $x^{p+1}$ ) en faisant une demi-itération supplémentaire par la formule  $(E_1)$  (resp.  $(E_2)$ ).

Exemple : Pour le graphe de la figure 14 associé à la matrice :

$$A = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{vmatrix}$$

On a les matrices de degrés relatifs :

$$A_Y = \begin{vmatrix} 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.33 & 0.33 & 0.33 & 0 & 0 \\ 0.33 & 0 & 0.33 & 0 & 0.33 & 0 \\ 0 & 0 & 0 & 0.33 & 0.33 & 0.33 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{vmatrix} \quad A_X = \begin{vmatrix} 0.5 & 0 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.33 & 0.33 & 0.33 \\ 0 & 0 & 0 & 0.5 & 0.5 \end{vmatrix}$$

Et les matrices d'itérations :

$$M_1 = \begin{vmatrix} 0.5 & 0.25 & 0.25 & 0 & 0 \\ 0.166 & 0.5 & 0.166 & 0.166 & 0 \\ 0.166 & 0.166 & 0.44 & 0.11 & 0.11 \\ 0 & 0.166 & 0.11 & 0.44 & 0.27 \\ 0 & 0 & 0.166 & 0.416 & 0.416 \end{vmatrix} \quad M_2 = \begin{vmatrix} 0.416 & 0.25 & 0.166 & 0 & 0.166 & 0 \\ 0.25 & 0.416 & 0.166 & 0.166 & 0 & 0 \\ 0.166 & 0.166 & 0.33 & 0.166 & 0.166 & 0 \\ 0 & 0.166 & 0.166 & 0.33 & 0.166 & 0.166 \\ 0.11 & 0 & 0.11 & 0.11 & 0.39 & 0.27 \\ 0 & 0 & 0 & 0.166 & 0.416 & 0.416 \end{vmatrix}$$

$M_1$  et  $M_2$  sont des matrices carrées de tailles respectives  $k$  et  $m$ .

A partir des formules  $(P_1)$  et  $(P_2)$ , on voit que la méthode des degrés relatifs s'écrit itérativement :

$$x^0, x^n = M_1 x^{n-1} = (M_1)^n x^0, \quad n \geq 1.$$

$$y^0, y^n = M_2 y^{n-1} = (M_2)^n y^0, \quad n \geq 1.$$

On reconnaît une itération de type méthode de la puissance sur la matrice  $M_1$  ou  $M_2$  (cf. Varga [21] et Wilkinson [26]).

#### 2.4. Propriétés des matrices d'itérations

Soit  $\alpha_i$  et  $\beta_j$  les coefficients (degrés relatifs) des matrices  $A_Y$  et  $A_X$ ,  $\alpha_i = 1 / \sum_j a_{ij}$  et  $\beta_j = 1 / \sum_i a_{ji}$ . On a :

$$(K_1) \quad M_1(i,j) = \alpha_i \sum_{t=1}^m \beta_t a_{it} a_{jt} \quad i, j = 1, 2, \dots, k$$

$$(K_2) \quad M_2(i,j) = \beta_j \sum_{t=1}^k \alpha_t a_{ti} a_{tj} \quad i, j = 1, 2, \dots, m$$

Nous allons établir un certain nombre de propriétés pour la matrice  $M_1$  ; la symétrie en les  $\alpha_i$  et  $\beta_j$  des formules  $(K_1)$  et  $(K_2)$  permet de vérifier aisément que toutes ces propriétés sont également valables pour la matrice  $M_2$ .

En posant :

$$\theta_{ij} = \sum_{t=1}^m \beta_t a_{it} a_{jt} \quad i, j = 1, 2, \dots, k ; \quad i \neq j$$

et

$$n_i = \sum_{t=1}^m \beta_t a_{it}^2 = \sum_{t=1}^m \beta_t a_{it} \quad i = 1, 2, \dots, k$$

(A étant booléenne,  $a_{ij}^2 = a_{ij}$   $i = 1, 2, \dots, n$ )



On a :  $M_1(i,j) = \alpha_i \theta_{ij}$  et  $M_1(i,i) = \alpha_i n_i$

$M_1(i,j)$  se décompose en deux facteurs  $\alpha_i$  et  $\theta_{ij}$ .

$\theta_{ij}$  est proportionnel au nombre des  $Y_t$  successeurs communs à  $X_i$  et  $X_j$ , et inversement proportionnel au nombre total de prédécesseurs  $1/\beta_t$  de chaque  $Y_t$ .

C'est un coefficient de liaison entre  $X_i$  et  $X_j$  dont la valeur est d'autant plus grande que :

1°)  $X_i$  et  $X_j$  ont davantage de leurs successeurs en commun,

2°) ces successeurs communs  $Y_t$  sont moins reliés à d'autres sommets  $X_r$  que  $X_i$  et  $X_j$  (rôle des  $\beta_t$ ).

L'intensité relative de la liaison entre  $X_i$  et  $X_j$ , appréciée par rapport à  $X_i$  est inversement proportionnelle au nombre total des successeurs de  $X_i$  (rôle de  $\alpha_i$ ).

De façon analogue,  $n_i$  est interprétable comme un coefficient d'"auto-liaison" pour  $X_i$

Propriété 1 :  $M_1$  est à silhouette<sup>(\*)</sup> symétrique

Démonstration : Pour  $i \neq j$ ,  $M_1(i,j) = 0 \iff M_1(j,i) = 0 \iff \theta_{ij} = 0$ .

Comme  $\theta_{ij} = \theta_{ji}$ ,  $\text{silh}(M_1)$  est symétrique.

En général,  $\alpha_i \neq \alpha_j$ , et donc  $M_1$  n'est pas symétrique :

$$M_1(i,j) = \alpha_i \theta_{ij} \neq \alpha_j \theta_{ij} = M_2(j,i)$$

---

(\*) Définition : on appelle silhouette de  $M_1$  et on note  $\text{silh}(M_1)$  la matrice booléenne  $(s_{ij})$  telle que :

$$s_{ij} = 1 \iff M_1(i,j) \neq 0 \quad i,j = 1,2,\dots, k$$
$$s_{ij} = 0 \iff M_1(i,j) = 0$$

Propriété 2 :  $0 \leq M_1(i,j) \leq \frac{1}{2}$  pour  $i \neq j$

Démonstration :

Soit  $k_{ij}$  le nombre de successeurs  $Y_t$  communs à  $X_i$  et à  $X_j$ , on a :  $\alpha_i \leq \frac{1}{k_{ij}}$ . De plus, pour tous ces  $Y_t$ ,  $\beta_t$  vaut au plus  $\frac{1}{2}$ , puisqu'ils sont liés au moins à 2 sommets,  $X_i$  et  $X_j$  ; d'où :  $M_1(i,j) \leq \frac{1}{k_{ij}} \times k_{ij} \times \frac{1}{2} = \frac{1}{2}$ .

Propriété 3 :  $M_1(i,i) \geq \frac{1}{k}$ ,  $i = 1, 2, \dots, k$

Démonstration :

$$\beta_t \geq \frac{1}{k} \implies M_1(i,i) \geq \alpha_i \sum_{t=1}^m \frac{a_{it}}{k} \geq \frac{1}{k_i} \times \frac{k_i}{k} = \frac{1}{k}$$

où  $k_i$  est le degré du sommet  $X_i$ .

Propriété 4 :  $M_1$  est irréductible si et seulement si  $G$  est connexe

Démonstration :

Soit  $S$  le graphe dont  $\text{silh}(M_1)$  est la *matrice d'incidence* (au sens classique, et non pas au sens des matrices rectangulaires associées aux graphes bipartis en 2.1.).

On a :  $M_1$  irréductible  $\iff \text{silh}(M_1)$  irréductible  $\iff S$  fortement connexe.

Comme  $\text{silh}(M_1)$  est symétrique (cf. Propriété 1),

$S$  fortement connexe  $\iff S$  connexe

Or, par définition de  $M_1$ ,  $S$  est le graphe construit sur  $X$  en mettant une arête entre  $X_i$  et  $X_j$  si et seulement si il existe  $Y_t$  lié à la fois à  $X_i$  et  $X_j$ .

A toute chaîne de S entre  $X_i$  et  $X_j$  on peut donc associer une chaîne alternée  $X_i - Y_t - X_{\ell} \dots Y_s - X_j$  dans G, le graphe biparti de départ.

On a donc : S connexe  $\iff$  G connexe, et, par transitivité des équivalences

$M_1$  irréductible  $\iff$  G connexe.

Propriété 5 :  $M_1$  est stochastique en lignes (\*)

Démonstration :

$$\sum_j M_1(i,j) = \sum_j \alpha_i \sum_{t=1}^m \beta_t a_{it} a_{jt}$$

$$= \sum_{j=1}^k \frac{1}{\sum_{j_1=1}^m a_{ij_1}} \sum_{t=1}^m \frac{1}{\sum_s a_{st}} a_{it} a_{jt}$$

$$= \frac{1}{\sum_{j_1=1}^m a_{ij_1}} \sum_{t=1}^m \frac{1}{\sum_{s=1}^k a_{st}} a_{it} \sum_j a_{jt}$$

$$= \frac{1}{\sum_{j_1=1}^m a_{ij_1}} \sum_{t=1}^m \frac{\sum_{j=1}^k a_{jt}}{\sum_{s=1}^k a_{st}} a_{it}$$

$$= \frac{\sum_{t=1}^m a_{it}}{\sum_{j_1=1}^m a_{ij_1}} = 1.$$

(\*) Définition :  $M_1$  est stochastique en ligne ssi  $M_1(i,j) \geq 0$   $i,j = 1,2,\dots,k$   
 et  $\sum_{j=1}^k M_1(i,j) = 1$   $i = 1,2,\dots,k$

Propriété : La plus grande valeur propre d'une matrice stochastique en ligne e simple et vaut +1, le vecteur propre associé  $u^1$  a toutes ses composantes égale entre elles :  $u_1^1 = u_2^1 = \dots u_k^1 = 1$

Considérons les valeurs propres  $\lambda_i$  de  $M_1$  ordonnées par modules décroissants :

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_k|$$

Si  $G$  est connexe, les propriétés 4 et 5 qui précèdent donnent :

$$|\lambda_1| = \lambda_1 = 1 > |\lambda_2| \geq |\lambda_3| \dots \geq |\lambda_k|$$

Supposons que  $M_1$  admette une base de vecteurs propres  $(u^1, u^2 \dots u^k)$  associée aux valeurs propres *ordonnées* ci-dessus.

En écrivant les vecteurs itérés dans cette base, on a :

$$x^0 = \sum_{i=1}^k a_i u^i$$

$$(M_1)^n x^0 = \sum_{i=1}^k \lambda_i^n a_i u^i = a_1 u^1 + \sum_{i=2}^k \lambda_i^n a_i u^i$$

Comme  $\max_{i=2, \dots, k} |\lambda_i| < 1$ ,  $(M_1)^n x^0 \longrightarrow a_1 u^1$  quand  $n \rightarrow +\infty$ , si  $a_1 \neq 0$

(Sinon  $(M_1)^n x^0 \rightarrow 0$ )  $a_1$  est la première composante de  $x^0$  dans la base  $(u^1, u^2 \dots u^k)$  des vecteurs propres.

En pratique, avec un  $x^0$  choisi au hasard, la condition  $a_1 \neq 0$  (et assez grand) est vérifiée.

Ceci établit la classique convergence *numérique* des itérations  $(P_1)$  et  $(P_2)$ .

Remarque :

on a  $\lambda_1 = 1$ , ce qui rend inutile ici toute normalisation des vecteurs itérés (cf. méthode de la puissance [26]). De plus, si  $x^0$  est dans  $\mathcal{C} = [0,1]^k$ , tous les itérés  $x^n$  resteront dans  $\mathcal{C}$ .

## 2.5. Convergence en ordre de la méthode des degrés relatifs

Comme au paragraphe précédent, l'étude est menée sur la matrice  $M_1$ , mais les résultats obtenus s'établissent de la même façon avec  $M_2$ , du fait qu'on passe de  $(x_i)$  à  $(y_j)$  par la demi-itération  $(E_1)$  (voir § 2.3.).

### 2.5.1. Définition de la convergence en ordre

Dans le cas où le graphe biparti étudié est connexe, la méthode décrite par la formule d'itération  $(P_1)$  converge vers le vecteur  $u^1$ , pour presque tout point de départ  $x^0$ .

$u^1$  ayant toutes ses composantes égales, il ne permet pas de définir une permutation  $\Pi(u^1)$  à partir de laquelle on pourrait classer les sommets sur le dessin (cf. § 2.2.).

Supposons maintenant que les hypothèses de convergence de  $(P_1)$  sont vérifiées, nous pouvons alors étudier si  $\Pi(x^r)$  converge également, d'où la :

Définition : la méthode des degrés relatifs *converge en ordre* vers la permutation  $p_1^*$  si :

1) il existe un rang  $r_0$ , pour la suite  $(x^r)$   $r \geq 0$  tel que, pour tout  $r \geq r_0$ ,  $\Pi(x^r)$  reste constante et égale à  $p_1^*$ .

2)  $p_1^*$  ne dépend pas du vecteur de départ  $x^0$ . (à un retournement près du dessin).

Nous allons caractériser cette convergence en ordre et essayer ensuite d'apprécier les résultats obtenus du point de vue du problème des croisements d'arêtes.

### 2.5.2. Caractérisation

Sous certaines hypothèses restrictives, nous allons établir la convergence en ordre de la méthode et caractériser la permutation-limite  $p_1^*$  ; nous discuterons ensuite ces hypothèses, qui sont des conditions suffisantes de la convergence en ordre.

Théorème :

La méthode des degrés relatifs converge en ordre vers  $p_1^*$  égal (à un retournement près) à  $\Pi(u^2)$ , où  $u$  est le vecteur propre associé à la seconde valeur propre  $\lambda_2$ .

Cet énoncé suppose remplies les quatre conditions suivantes :

- (H<sub>1</sub>)  $M_1$  admet une base de vecteurs propres ( $u^1, u^2 \dots u^k$ )
- (H<sub>2</sub>) le vecteur de départ  $x^0$  n'est pas colinéaire à  $u^1$  (sinon la suite des itérés ( $x^n$ ) reste stationnaire en  $x^0$ ) et les composantes  $a_1$  et  $a_2$  de  $x^0$  (sur  $u^1$  et  $u^2$  respectivement) sont non nulles.
- (H<sub>3</sub>)  $\lambda_1$  et  $\lambda_2$  sont des valeurs propres simples
- (H<sub>4</sub>) l'ordre défini par les composantes de  $u^2$  est *strict* .

a) Démonstration :

Ecrivons les vecteurs de la suite ( $x^r$ ) dans une base de vecteurs propres ( $u^1, u^2 \dots u^k$ ) (hypothèse (H<sub>1</sub>)).

On a , à la constante multiplicative  $a_1 \neq 0$  (hypothèse (H<sub>2</sub>)) près :

$$\frac{1}{a_1} x^r = u^1 + \lambda_2^r \frac{a_2}{a_1} u^2 + \dots + \lambda_k^r \frac{a_k}{a_1} u^k$$

En posant :  $\epsilon_i^r = \frac{1}{a_1} (1 - x_i^r)$  pour  $i = 1, 2, \dots, k$ , il vient :

$$\begin{aligned} \epsilon_i^r &= \lambda_2^r [a_2 u_i^2 + (\frac{\lambda_3}{\lambda_2})^r a_3 u_i^3 + \dots + (\frac{\lambda_k}{\lambda_2})^r a_k u_i^k] \\ &= \lambda_2^r [a_2 u_i^2 + \epsilon_i^r] \text{ avec } \epsilon_i^r \xrightarrow{r \rightarrow +\infty} 0 \end{aligned}$$

En effet, l'hypothèse (H<sub>3</sub>) nous garantit simultanément  $\lambda_2 \neq 0$  (pour  $k > 2$ )

et  $|\frac{\lambda_i}{\lambda_2}| < 1$  pour  $i = 3, 4, \dots, k$  et (H<sub>2</sub>) assure  $a_2 \neq 0$ .

Etudier l'ordre des composantes de  $x^r$  revient à comparer entre eux les  $\epsilon_i^r$ .

$u^2$  vecteur propre associé à  $\lambda_2$  a toutes ses composantes réelles, d'après (H<sub>3</sub>), et, à une normalisation près, comprises dans  $[-1, +1]$ ; l'hypothèse (H<sub>4</sub>) permet de définir  $\Pi(u^2)$  sans ambiguïté ; à  $\Pi(u^2)$  près,  $u_1^2 < \dots < u_k^2$ .

On a :  $\epsilon_i^r = \lambda_2^r [a_2 u_i^2 + \epsilon_i^r]$  pour  $i = 1, 2, \dots, k$ .

La différence de deux  $\epsilon_i^r$  consécutifs s'écrit donc :

$$\epsilon_{i+1}^r - \epsilon_i^r = \lambda_2^r [a_2 (u_{i+1}^2 - u_i^2) + (\epsilon_{i+1}^r - \epsilon_i^r)]$$

Quand  $r \rightarrow +\infty$ , le terme prépondérant dans cette différence est  $a_2$

$(u_{i+1}^2 - u_i^2)$  :

$$\text{soit } \epsilon^* = \min(|u_{i+1}^2 - u_i^2|, i = 1, 2, \dots, k-1)$$

D'après (H<sub>4</sub>),  $\epsilon^* > 0$ .

En prenant  $r_0$  tel que :  $|\epsilon_i^{r_0}| < \frac{\epsilon^*}{4}$  pour tout  $i$ , le facteur  $(\epsilon_{i+1}^{r_0} - \epsilon_i^{r_0})$  est inférieur en valeur absolue à  $\frac{\epsilon^*}{2}$  et comme  $\frac{\epsilon_i^r}{\lambda_2^r}$  reste alors compris entre  $(u_1^2 - \frac{\epsilon^*}{2})$  et  $(u_1^2 + \frac{\epsilon^*}{2})$ , l'ordre des  $\epsilon_i^r$  ne change plus, vu le choix effectué pour  $\epsilon^*$ .

D'où le résultat :  $r_0 : r \geq r_0 \quad \Pi(\epsilon^r) = \Pi(\text{sign}(a_1 \ a_2) \ u^2) = p_1^*$ .

Remarque :

Dans  $p_1^*$ ,  $\text{sign}(a_1 \ a_2)$  dépend de  $x_0$ . Mais, d'après la remarque 1 § 2.2., les permutations  $p_1^*$  et  $\Pi(u^2)$  donnent des dessins identiques, à un retournement près.

C'est donc l'ordre  $\Pi(u^2)$  qui définit l'ordre  $\Pi_r(\epsilon^r)$  et qui guide ainsi le processus de convergence des  $x^r$ .

$\Pi(u^2)$  définit le nombre limite  $\chi^*$  de croisements d'arêtes obtenu par la méthode des degrés relatifs.

#### b) Rôle des hypothèses

- l'hypothèse  $(H_1)$  n'est sans doute pas nécessaire, mais elle facilite la démonstration : Supprimer  $(H_1)$  oblige à se contenter d'une forme de Jordan pour  $M_1$ , au lieu d'une forme diagonale.
- l'hypothèse  $(H_2)$  est sans grand inconvénient pratique : un vecteur  $x^0$  pris au hasard sera en général satisfaisant.
- nous ne disposons pas d'une interprétation satisfaisante pour  $(H_3)$ ; mais simplement de la conjecture suivante :

Conjecture 1 : si  $(H_3)$  est vérifiée,  $(H_4)$  est inutile.



Nous allons donner deux exemples pour lesquels  $\lambda_2$  n'est pas valeur propre simple, et qui permettent de mieux situer la portée de l'hypothèse (H<sub>3</sub>) et son lien avec les ex-aequos dans  $\Pi(u^2)$ , c'est-à-dire les composantes égales de  $u^2$ .

1) Graphe complet  $K_{k,m}$  : (cf. Roy [ 20 ] ) : Si G est complet, A est universelle et  $M_1$  a tous ses coefficients égaux à  $1/k$ . Dans ce cas très particulier,  $M_1$  étant symétrique, on peut calculer facilement la matrice de déflation :

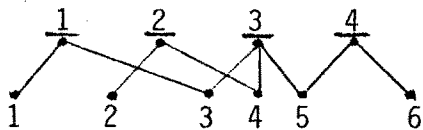
$$M_1' = M_1 - \frac{t_{u^1} \times u^1}{u^1 \times t_{u^1}} \text{ et comme } u_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, M_1' \text{ est la matrice}$$

nulle ; on a donc :

$$\lambda_2 = \lambda_3 = \dots = \lambda_k = 0$$

$\lambda_2$  n'est donc pas valeur propre simple.

2) Triple chaîne : considérons le graphe suivant :



Il est formé de trois chaînes de même longueur articulées sur le sommet 3 du niveau du haut : 3-3-1-1, 3-4-2-2 et 3-5-4-6.

Pour ce graphe on a :

$$M_1 = \begin{vmatrix} \frac{3}{4} & 0 & \frac{1}{4} & 0 \\ 0 & \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} \end{vmatrix}$$

D'après la structure du graphe, on voit facilement que  $\lambda_2 = \frac{3}{4}$  est valeur propre triple, associée au sous-espace propre défini par le trièdre :

$$u^2 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ +1 \end{pmatrix} \quad u^3 = \begin{pmatrix} -1 \\ -1 \\ 0 \\ +2 \end{pmatrix} \quad u^4 = \begin{pmatrix} -2 \\ 0 \\ +1 \\ +1 \end{pmatrix}$$

Dans ce cas particulier, suivant le point de départ  $x^0$ , on constate qu'on obtient un classement variable des composantes, il n'y a donc pas convergence en ordre au sens de la définition.

Dans ces deux exemples, non seulement  $\lambda_2$  n'est pas valeur propre simple, mais son ordre de multiplicité est  $k-1$ . Ceci n'est pas surprenant pour les graphes bipartis *complets* puisque toutes les permutations sont équivalentes, la convergence en ordre n'a pas grande signification.

L'autre exemple est plus complexe, bien que toujours lié à des propriétés de symétrie du graphe qui se traduisent sur le vecteur propre  $u^2$  de la manière suivante : on peut avoir des composantes  $u_i^2$  égales sans que les lignes et les colonnes correspondantes de  $M_1$  soient identiques.

Ceci nous amène à distinguer dans  $\Pi(u^2)$  les "vrais" et "faux" ex-aequos :

Si les lignes  $i_1, i_2, \dots, i_p$  de  $M_1$  sont identiques, on a

$u_{i_1}^2 = u_{i_2}^2 = \dots = u_{i_p}^2$ . Les sommets  $X_{i_1}, X_{i_2}, \dots, X_{i_p}$  sont de "vrais" ex-aequos

car ils sont aussi associés à des composantes égales dans tous les autres vecteurs propres  $u^3, \dots, u^k$ , et il n'est évidemment pas question de chercher un rang  $r_0$  à partir duquel ils soient classés entre eux de manière stable dans  $x^r$ .



Soit  $\mathcal{D}$  le compact :

$$\mathcal{D} = \{x = (x_i) : -1 \leq x_1 \leq x_2 \dots \leq x_k \leq 1\}.$$

Montrons que :

$$x^r \in \mathcal{D} \implies x^{r+1} \in \mathcal{D} \quad r \geq 1 ;$$

$$x_2^{r+1} - x_1^{r+1} = \frac{1}{4} [(x_1^r + 2x_2^r + x_3^r) - (3x_1^r + x_2^r)]$$

$$= \frac{1}{4} [x_2^r + x_3^r - 2x_1^r]$$

$$= \frac{1}{4} [(x_2^r - x_1^r) + (x_3^r - x_1^r)] \geq 0$$

et de même :  $x_3^{r+1} - x_2^{r+1} \geq 0$  etc ...

Donc si  $x^0$  est dans  $\mathcal{D}$  la suite  $(x^r)$  reste dans  $\mathcal{D}$ , et l'ordre  $\Pi(x^r) = \Pi(-u^2)$  ne change pas au cours des itérations.

Vu la structure particulière de  $M_1$ , on vérifie facilement que  $u^2$  ne contient pas d'ex-aequos :

$$M_1 u^2 = \lambda_2 u^2 \implies \frac{3}{4} u_1^2 + \frac{1}{4} u_2^2 = \lambda_2 u_1^2 < u_1^2$$

D'où  $u_2^2 \neq u_1^2$  et par un raisonnement analogue de proche en proche :

$$u_1^2 > u_2^2 > \dots > u_k^2$$

$\Pi(u^2)$  est donc bien défini sans ambiguïté et fournit le meilleur classement possible.

## 2.6. Etude expérimentale

Pour le problème de la minimisation du nombre de croisements d'arêtes, nous avons étudié la dynamique de la convergence en ordre de la méthode des degrés relatifs et les éléments propres de la matrice  $M_1$ .

### 2.6.1. Notations

Dans ce qui suit, nous aurons à distinguer :

\* *l'optimum*  $\omega = \chi(\hat{p}_1, \hat{p}_2)$ .  $\hat{p}_1$  et  $\hat{p}_2$  étant le (ou les) couple(s) de permutations qui permettent d'avoir un dessin du graphe biparti comportant le minimum absolu de croisements d'arêtes, pour tous les dessins possibles.

$$* \text{ le minimum } \mu(x^0) = \min_{r \geq 0} [\chi(\Pi_1^r, \Pi_2^r)]$$

où  $\Pi_1^r = \Pi(x^r)$  et  $\Pi_2^r = \Pi(y^r)$ .  $\mu(x^0)$  est le nombre de croisements minimal fourni par la méthode des degrés relatifs quand on part de  $x^0$ .

\* *la limite*  $\chi^* = \chi(p_1^*, p_2^*)$  dans les cas où l'on a convergence en ordre de la méthode (cf. théorème 2.5.).

A défaut de connaître  $\omega$ , il est intéressant d'en obtenir une approximation grâce à  $\mu$  et  $\chi^*$ , si cela est possible. Il est clair que :

$$\omega \leq \mu(x^0)$$

En général on aura :  $\mu(x^0) \leq \chi$  pour presque tout  $x^0$ .

### 2.6.2. Méthode de travail

Pour effectuer cette étude, nous avons utilisé :

- un programme standard de détermination des valeurs propres et vecteurs propres (algorithme QR).

- un programme d'itération de la méthode des degrés relatifs, qui détermine à chaque itération  $r$  :  $\pi_1^r$ ,  $\pi_2^r$ , et  $\chi(\pi_1^r, \pi_2^r)$ .

Les exemples traités sont de deux types :

a) Des graphes simples, de taille assez petite, pour lesquels  $\omega$  est connu et qui sont construits "sur mesure" pour tester la méthode dans les conditions de la convergence en ordre.

b) Des graphes très touffus, de grande taille (jusqu'à 150 sommets et plus) qui sont des cas réels d'application.

Ces exemples proviennent du CTGREF<sup>(\*)</sup> de Grenoble et représentent des graphes de décision pour des problèmes de politique agricole régionale.

$\omega$  n'étant pas connu, leur traitement reste heuristique. Ce sont ces graphes qui ont suscité l'étude spécifique des graphes bipartis.

### 2.6.3. Résultats

#### a) Éléments propres

Le calcul des éléments propres de  $M_1$  nous a permis de constater sur tous les exemples traités que tous les éléments propres sont *réels*, même lorsque les hypothèses de simplicité de  $\lambda_1$  et  $\lambda_2$  n'étaient pas vérifiées.

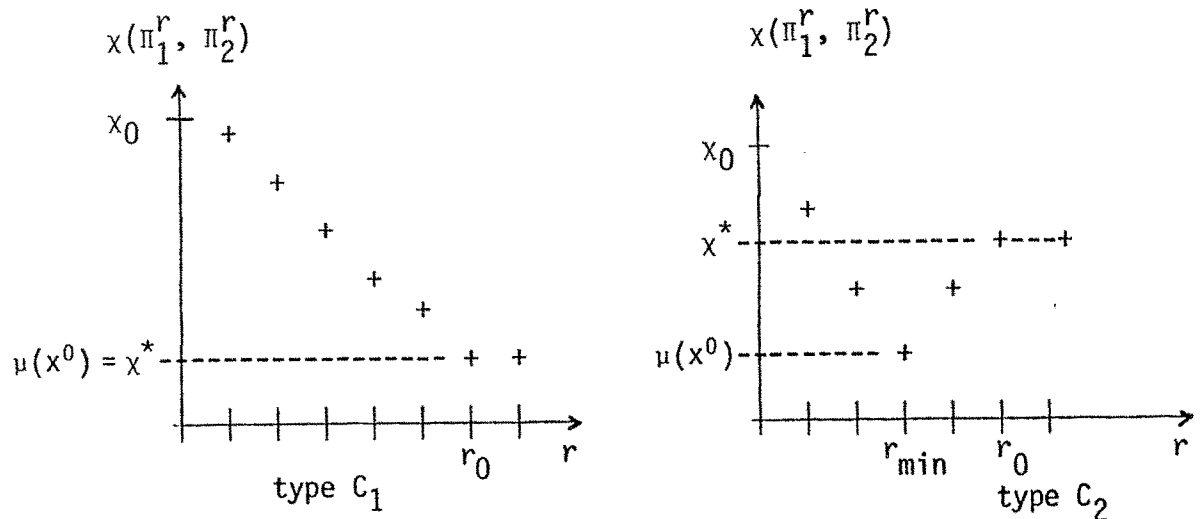
L'hypothèse faite précédemment sur ces éléments propres ( $M_1$  diagonalisable) ne semble donc pas restrictive en pratique.

---

(\*) Centre Technique du Génie Rural des Eaux et Forêts - Ministère de l'Agriculture.

b) dynamique de convergence

L'expérimentation a fait apparaître deux catégories de graphes, d'après l'allure du processus de convergence :



- Figure 17 -

Pour  $C_1$  on a :  $\mu(x^0) = x^* \leq x_0$  pour presque tout  $x^0$  et pour

$C_2$  :  $\exists x^0 / \mu(x^0) < x^* \leq x_0$ ,  $\mu$  étant atteint pour  $r = r_{\min} < r_0$  valeur de  $r$  à partir de laquelle  $\pi(x^r)$  reste constante.

Pour les graphes "sur mesure" on a obtenu :

- la convergence de type  $C_1$
- la solution optimale :  $\omega = \mu = x^*$

Pour les graphes du CTGREF, on a :

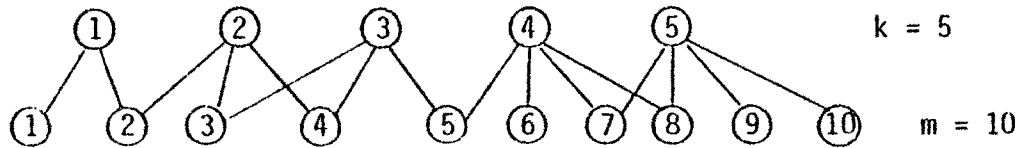
- une convergence de type  $C_2$
- $\omega$  inconnu, mais  $\mu(x^0)$  semble proche de  $\omega$ .

Lorsque  $x^0$  est un vecteur quelconque (non colinéaire à un des vecteurs propres) on observe que :

- la dépendance de  $\mu$ ,  $r_{\min}$  et  $r_0$  par rapport à  $x^0$  est faible
- $r_{\min}$  et  $r_0$  sont des fonctions croissantes de  $k$ .

Exemples :

a) Le graphe suivant donne une convergence de type  $C_1$  :



On a :  $\omega = \mu = \chi^* = 2$

b) Pour deux graphes du CTGREF nous avons obtenu :

k	m	$r_{\min}$	$\mu$	$\chi^*$	$\chi_0$	$r_0$	$\mu/\chi_0$	$\chi^*/\chi_0$
10	28	4	302	385	485	18	0.63	0.79
24	98	5	4419	5018	8844	25	0.49	0.57

Dans les exemples étudiés,  $\lambda_2$  est voisine de 0,90 et  $\lambda_3$  légèrement inférieure (de quelques centièmes) ce qui explique les valeurs assez élevées de  $r_0$ , puisque la vitesse de la convergence en ordre dépend de  $\frac{\lambda_3}{\lambda_2}$  ;



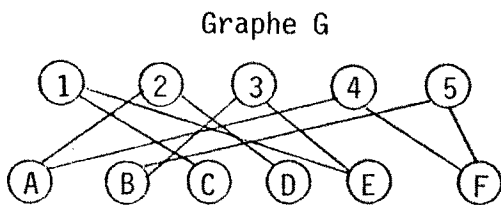
Les autres vecteurs propres :  $u^3, u^4 \dots u^k$  donnent des permutations  $\Pi(u^i)$  avec des nombres de croisements de plus en plus élevés.

Dans certains cas particuliers,  $u^3, u^4$  etc ... apparaissent liés aux symétries du graphe.

2.7. Un problème voisin : la bande diagonalisation des matrices

D'après 2.1., toute matrice booléenne A rectangulaire  $k \times m$  (carrée si  $k = m$ ) peut être considérée comme matrice d'incidence d'un graphe biparti  $G = (X, Y, U)$ , X et Y étant respectivement associés aux lignes et aux colonnes de la matrice A.

Les dessins qui suivent montrent la parenté qui existe entre le problème de la minimisation du nombre de croisements d'arêtes et celui de la bande diagonalisation des matrices :

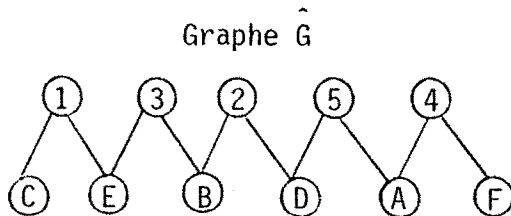


$x(p_1, p_2) = 19$

Matrice A

	A	B	C	D	E	F
1			x		x	
2	x			x		
3		x			x	
4	x					x
5		x		x		

après classement des sommets selon les permutations optimales  $(\hat{p}_1, \hat{p}_2)$  :



$x(\hat{p}_1, \hat{p}_2) = \omega = 0$

Matrice  $\hat{A}$

	C	E	B	D	A	F
1	x	x				
3		x	x			
2			x	x		
5				x	x	
4					x	x



degrés relatifs pour les matrices tridiagonalisables :

$$\chi^* = \mu = \omega$$

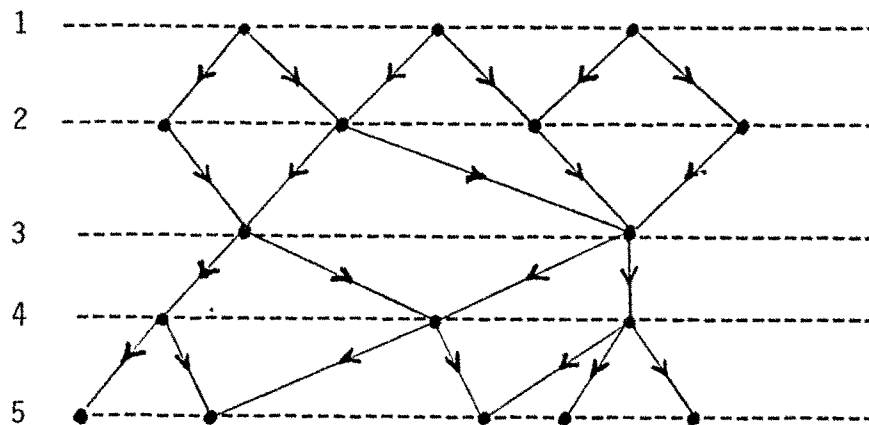
Ce résultat peut se généraliser avec une largeur de bande  $\geq 2$ . pour A.

### 3. LES GRAPHES MULTIPARTIS

Parmi les graphes hiérarchisés comportant plus de deux niveaux, les plus simples sont les graphes multipartis. Un graphe est de ce type si et seulement si :

\* tout sommet fait partie d'un chemin maximum (hiérarchie des rangs, hiérarchie inverse des rangs et hiérarchie serrée sont équivalentes),

\* il ne contient pas de shunts (arcs entre des niveaux non consécutifs) :



- Figure 18 -

### 3.1. Le problème des croisements

Dans un tel graphe, le problème de la minimisation du nombre de croisements se pose toujours en termes de permutations des sommets à l'intérieur de chaque niveau. Mais il est impossible d'utiliser les méthodes valables pour les graphes bipartis en cherchant les permutations les meilleures sur toutes les paires de niveaux consécutifs.

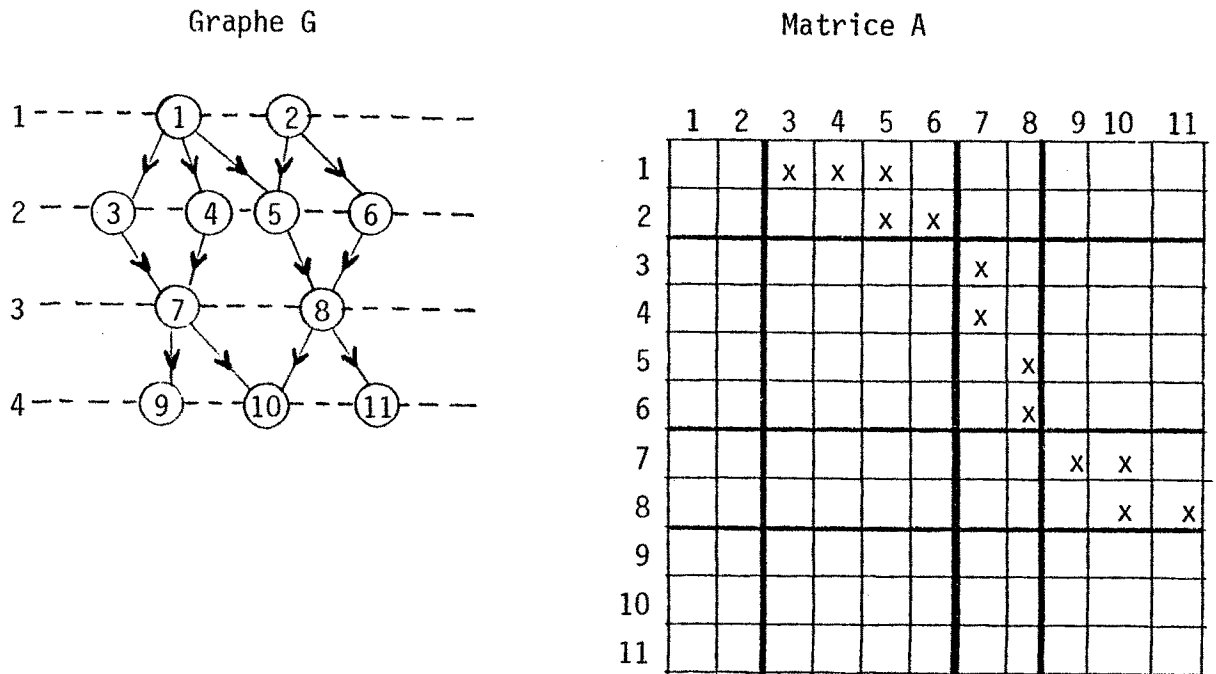
En effet, il est peu probable que la suite de ces solutions locales donne une bonne solution globale, à cause des incompatibilités de classement sur les niveaux intermédiaires : si l'on retient une certaine permutation  $p$  des sommets du niveau 2 par rapport au niveau 1 et une permutation  $p'$  par rapport au niveau 3,  $p$  et  $p'$  ne coïncideront généralement pas et la meilleure solution sera une permutation  $p''$  qui donnera un nombre de croisements supérieur à la somme des deux minima mais inférieur au nombre total de croisements résultant du choix de  $p'$  ou de  $p$ .

Le problème devient donc beaucoup plus complexe, puisqu'il s'agit de trouver pour chaque niveau la permutation qui soit la meilleure, non plus par rapport aux niveaux voisins, mais par rapport aux répercussions de ce choix sur l'ensemble de la représentation.

Les méthodes heuristiques utilisées (voir Laugier et Martinez [12] et Warfield [25] ) dans le cas des graphes multipartis consistent en général à se donner une liste de "configurations indésirables" et de règles élémentaires de transformation puis à chercher sur la matrice du graphe toutes les améliorations locales possibles. Leur inconvénient est d'imposer pour chaque cas un parcours systématique de la matrice où l'on cherche à appliquer la règle choisie, sans détruire les améliorations préalablement obtenues, ce qui conduit à des temps de calcul assez importants si l'on souhaite effectuer un travail vraiment exhaustif.

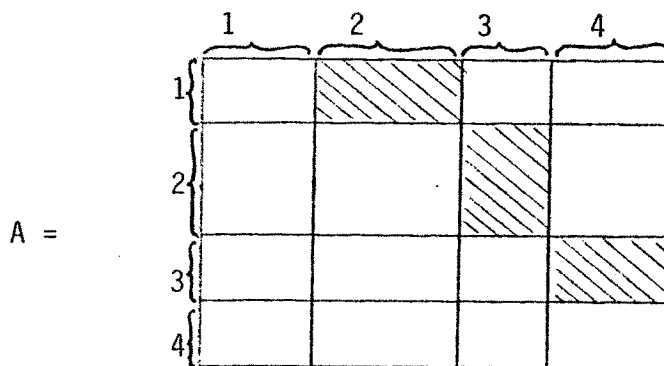
### 3.2. Extension de la méthode des degrés relatifs

La matrice d'incidence d'un graphe multiparti peut être mise sous forme bloc-triangulaire, si les lignes et les colonnes respectent l'ordre des niveaux :

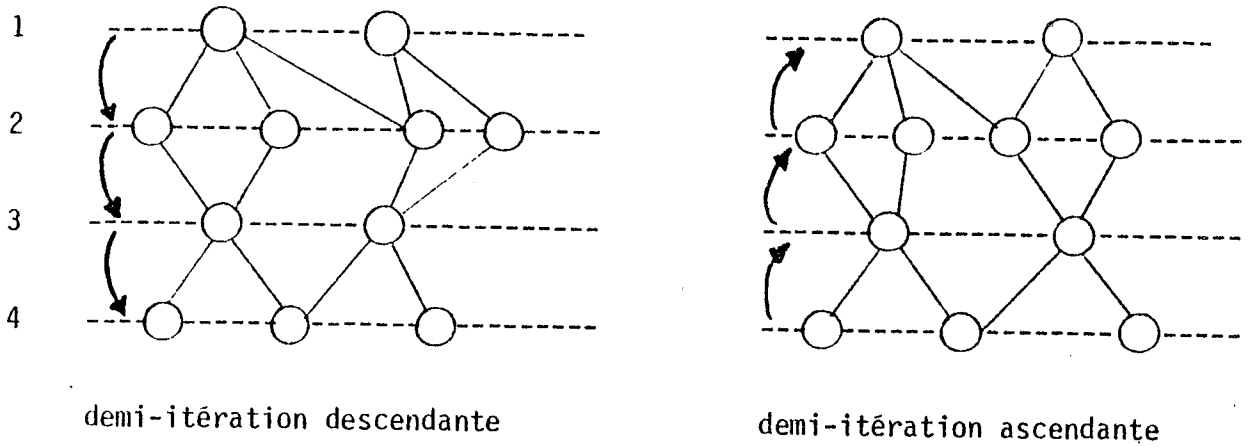


- Figure 19 -

Dans la matrice A, on a trois blocs rectangulaires qui décrivent les arcs existant entre les quatre niveaux du graphe :

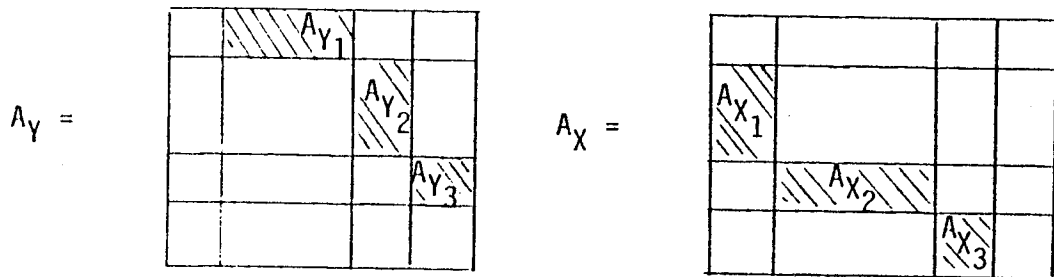


On peut définir la méthode itérative à partir des deux niveaux extrêmes :



- Figure 20 -

On introduit comme au § 2.3. les matrices de degrés relatifs :



D'où la suite des calculs par niveaux :

$$\left\{ \begin{array}{l} Y^0 = A_{Y_1} X^0 \\ Z^0 = A_{X_2} Y^0 \\ W^0 = A_{X_3} Z^0 \end{array} \right. \quad \text{puis} \quad \left\{ \begin{array}{l} W'^0 = W^0 \\ Z'^0 = A_{Y_3} W'^0 \\ Y'^0 = A_{Y_2} Z'^0 \end{array} \right.$$

On tire :  $X'^0 = A_{Y_1} Y'^0 = X^1 = (A_{Y_1} A_{Y_2} A_{Y_3} A_{X_3} A_{X_2} A_{X_1}) X^0$

$$\text{On pose : } M_1 = A_{Y_1} A_{Y_2} A_{Y_3} A_{X_3} A_{X_2} A_{X_1} = A'_{Y} A'_{X}$$

$$M_2 = A_{X_3} A_{X_2} A_{X_1} A_{Y_1} A_{Y_2} A_{Y_3} = A'_{X} A'_{Y}$$

$M_1$  et  $M_2$  sont comme précédemment les produits en sens inverse l'un de l'autre de deux matrices rectangulaires  $A'_{X}$  et  $A'_{Y}$ .

Nous n'avons pas expérimenté numériquement cette extension de la méthode, car les graphes que nous avons à traiter sont plus généraux que les graphes multipartis.

#### 4. FORME REDUITE ECLATEE : PROBLEMES SPECIFIQUES

Par rapport aux graphes multipartis, un graphe quelconque que l'on veut dessiner sous forme réduite éclatée pose trois problèmes nouveaux :

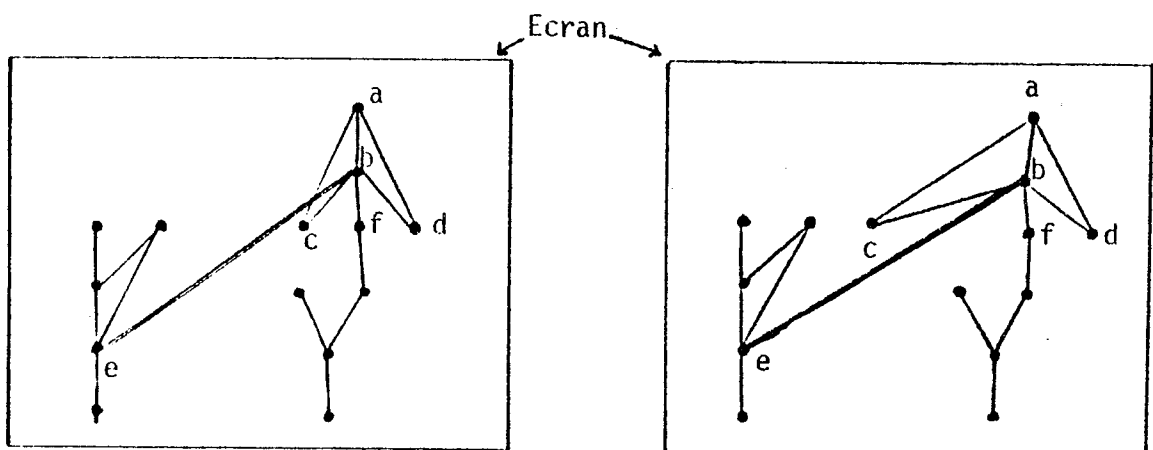
- \* les shunts
- \* la hiérarchie serrée
- \* les composantes fortement connexes.

##### 4.1. Les shunts

Bien que la hiérarchie serrée ait pour effet de raccourcir certains arcs inutilement longs, un certain nombre de shunts peuvent subsister, qui risquent de créer des croisements d'arêtes. On peut essayer de se ramener approximativement au cas d'un graphe multiparti en introduisant des sommets fictifs le long des shunts. Cependant, on aura du mal à construire des solutions dans lesquelles les shunts resteront rectilignes. En fait, une différence majeure apparaît avec le cas des graphes bipartis et multipartis : il ne suffit plus de rechercher pour chaque niveau une permutation optimale sur les sommets, il faut aussi tenir compte des *distances* entre les sommets pour ce qui concerne les croisements dûs aux shunts. En effet, pour qu'un shunt ne coupe pas une autre partie du dessin, il faut pouvoir

dégager de part et d'autre de son trajet un "couloir" qui ne soit traversé par aucun autre arc. Ceci suppose qu'on puisse écarter suffisamment les autres sommets de part et d'autre de cette zone et donc qu'on donne éventuellement aux arcs gênants une orientation assez éloignée de la verticale, alors qu'a priori, on aurait plutôt tendance à les placer en dessous de leurs prédécesseurs (cf. heuristique des degrés relatifs).

Ceci peut s'obtenir concrètement de deux manières : soit l'espace physique dont on dispose est assez large, soit on resserre les niveaux pour avoir un bon rapport dimension horizontale / dimension verticale, ce qui nuit évidemment à une utilisation rationnelle de l'espace disponible. Or un des critères que nous avons posé au début de ce chapitre était d'avoir une répartition homogène des sommets sur le dessin, dans la mesure du possible.



- Figure 21 -

- Figure 22 -

Le passage de la figure 21 à la figure 22 nécessite d'écarter nettement c vers la gauche, ce qui ne sera pas toujours possible concrètement, ni même réellement souhaitable : c et d sont reliés à a et b de la même façon, et il peut être plus intéressant de les laisser l'un à côté de l'autre, ce qui favorise le repérage visuel de leur similitude de situation. (par exemple, mettre c entre d et f).



Remarque :

Dans le cas concret d'où cet exemple est tiré, c et d représentent le Brésil et l'Argentine.

En résumé, le problème des shunts est aussi un problème d'espace disponible concrètement, et non plus seulement un problème de détermination d'une permutation optimale pour chaque niveau comme c'était le cas pour les graphes multipartis. Même si la solution théorique était connue, on ne pourrait pas forcément l'appliquer complètement à tout coup, à cause des contraintes matérielles.

4.2. La hiérarchie serrée

L'utilisation de la hiérarchie serrée implique l'existence de chemins de portée totale non maximale, c'est-à-dire que l'on a des sommets sans antécédent qui peuvent apparaître à des niveaux intermédiaires (sommets "suspendus"). Ces sommets introduisent un arbitraire supplémentaire, puisque l'on ne peut les placer que d'après les positions de leurs successeurs.

4.3. Les composantes fortement connexes

On doit respecter la contrainte de voisinage entre sommets d'une même c.f.c., ce qui complique la construction du dessin en interdisant certaines permutations sur les niveaux et en introduisant un risque supplémentaire de croisements : les croisements entre des shunts et les traits horizontaux symbolisant les c.f.c.

4.4. Heuristique des degrés relatifs

L'utilisation qui a été faite ici de la méthode des degrés relatifs tient compte des problèmes particuliers à la forme réduite éclatée.

Le calcul heuristique est effectué en deux phases :

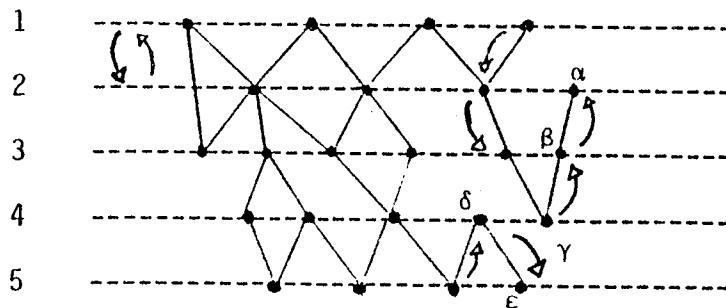
1) On commence par calculer un dessin du graphe réduit par la méthode dérivée de celle du paragraphe sur les graphes multipartis : on initialise les positions des sommets du haut par quelques itérations sur

le sous graphe biparti des deux premiers niveaux, puis on effectue une demi-itération descendante, avec deux modifications :

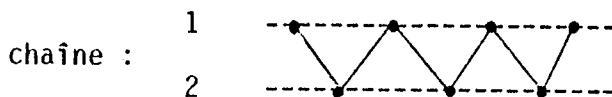
- \* les coefficients de pondération par les degrés relatifs tiennent compte du nombre réel d'arcs existant entre deux c.f.c.

- \* les sommets sans antécédent qui apparaissent au niveau  $i$  sont empilés dans une zone auxiliaire et, lorsque le placement au niveau  $i+1$  est achevé, on démarre une demi-itération ascendante sur les sommets qui sont dans la pile pour voir s'ils ont des successeurs déjà placés. Si oui, on place ces sommets et on relance une demi-itération descendante sur la pile etc ... jusqu'à ce qu'il n'y ait plus de nouveau sommet placé (le principe de parcours est le même que celui qui a été décrit au chapitre 2 dans l'algorithme de raccourcissement des arcs). Lorsque le traitement de la pile est terminé, on y range tous les sommets du niveau  $(i+1)$  qui sont apparus sans précédent puis on reprend la demi-itération descendante normale au niveau  $i+2$ .

Voyons un exemple :



L'itération sur les niveaux 1 et 2 donne le classement correspondant à la



Au niveau 2, le sommet  $\alpha$  va dans la pile, puis le sommet  $\beta$  au niveau 3 ;

au niveau 4, le sommet  $\gamma$  est placé et permet à  $\beta$  et  $\alpha$  de quitter la pile. Ensuite le sommet  $\delta$  entre dans la pile, et le placement des sommets du dernier niveau permet de calculer une position pour  $\delta$  puis pour  $\epsilon$ .

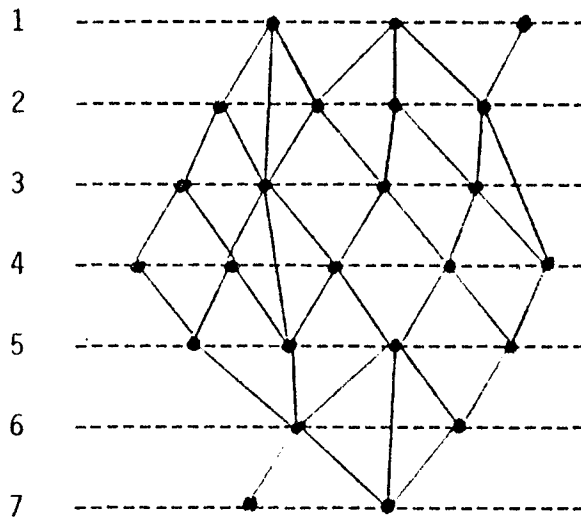
2) On éclate niveau par niveau les c.f.c. comportant plus d'un sommet en employant la même méthode : le positionnement relatif des sommets d'une même c.f.c. se fait en calculant par l'heuristique, une série de valeurs par rapport aux précédents et une série de valeurs par rapport aux suivants, puis en effectuant un classement qui tient compte des deux séries :

Ceci permet de respecter la contrainte de connexité pour les c.f.c., puisque les sommets d'une même c.f.c. seront répartis après ce classement dans l'intervalle séparant la c.f.c. de sa voisine de droite et de sa voisine de gauche. Outre son efficacité spécifique pour le problème traité, cette heuristique a l'intérêt d'être rapide. Elle a aussi certains inconvénients qui découlent de son caractère très local :

- pas d'optimisation globale,
- pas de solution efficace pour le problème des shunts,
- manque de finesse (ex-aequos).

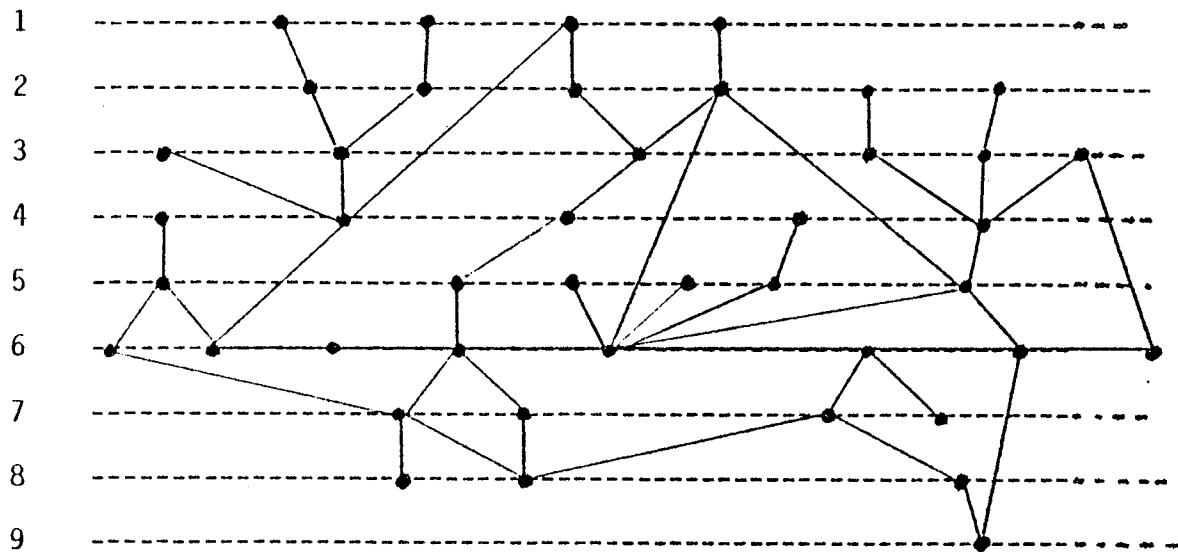
Cependant, moyennant quelques aménagements ajoutés dans les programmes de calcul : séparation des ex-aequos, redilatation éventuelle de l'intervalle de répartition, maintien d'un écart suffisant entre les sommets, ..., elle donne des dessins utilisables que l'on peut ensuite corriger à l'écran, si l'on souhaite améliorer la représentation (cf. Chapitre 5).

En fait, les résultats obtenus sont d'autant meilleurs que le graphe est plus proche d'un graphe multiparti, avec une distribution régulière des sommets sur les niveaux, le "graphe idéal" ayant la forme suivante (cf. paragraphe sur les graphes multipartis).



En effet, l'existence d'un bon maillage et le fait que le nombre de sommets par niveau varie régulièrement, garantissent qu'il y aura peu d'ex-aequos. De plus, les problèmes causés par la présence des shunts sont limités du fait de leur faible longueur.

Par contre, le graphe suivant est un exemple typique de graphe difficile à traiter par l'heuristique considérée :



En effet, toutes les sources de difficultés y sont réunies :

- shunts longs et nombreux,
- beaucoup de sommets "suspendus",

- absence de maillage,
- existence d'une grosse c.f.c. (problème de compatibilité entre les deux classements heuristiques à l'éclatement),
- graphe biparti des deux premiers niveaux non connexe (initialisation inefficace).

Pour un tel graphe, qui peut par ailleurs être dessiné sans aucun croisement d'arêtes, l'heuristique peut donner de très mauvais résultats.

Le cas est malgré tout relativement rare. En effet, dans les cas concrets que nous avons traités, nous disposons d'une aide supplémentaire non négligeable : la nomenclature. Or, quand il y a beaucoup d'arbitraire dans les choix, l'ordre de la numérotation des sommets devient un facteur important et le fait que cet ordre soit celui d'une nomenclature concrète (qui n'est en général pas bâtie au hasard) aide à faire les "bons choix".

En fait, si la structure du graphe est assez fortement maillée, l'heuristique suit le maillage et il n'y a pas trop de problèmes (hormis les shunts). Si, en revanche, le graphe comporte surtout de grandes filières isolées les unes des autres, l'ordonnancement devient plus aléatoire, mais dans ce cas, il y a moins d'arcs, donc moins de chances qu'ils se coupent !

Compte tenu des réserves que nous venons de faire, les trois objectifs fixés au départ sont assez bien remplis par la méthode des degrés relatifs :

adaptation aux formes standard : calculs portant sur les abscisses - dessin obtenu en deux temps : graphe réduit puis éclatement.

lisibilité : placement assez naturel au barycentre des précédents ou des suivants.

rapidité : pas de retour en arrière : chaque emplacement calculé est définitif

De nombreux dessins obtenus par cette méthode sont donnés au chapitre 6.

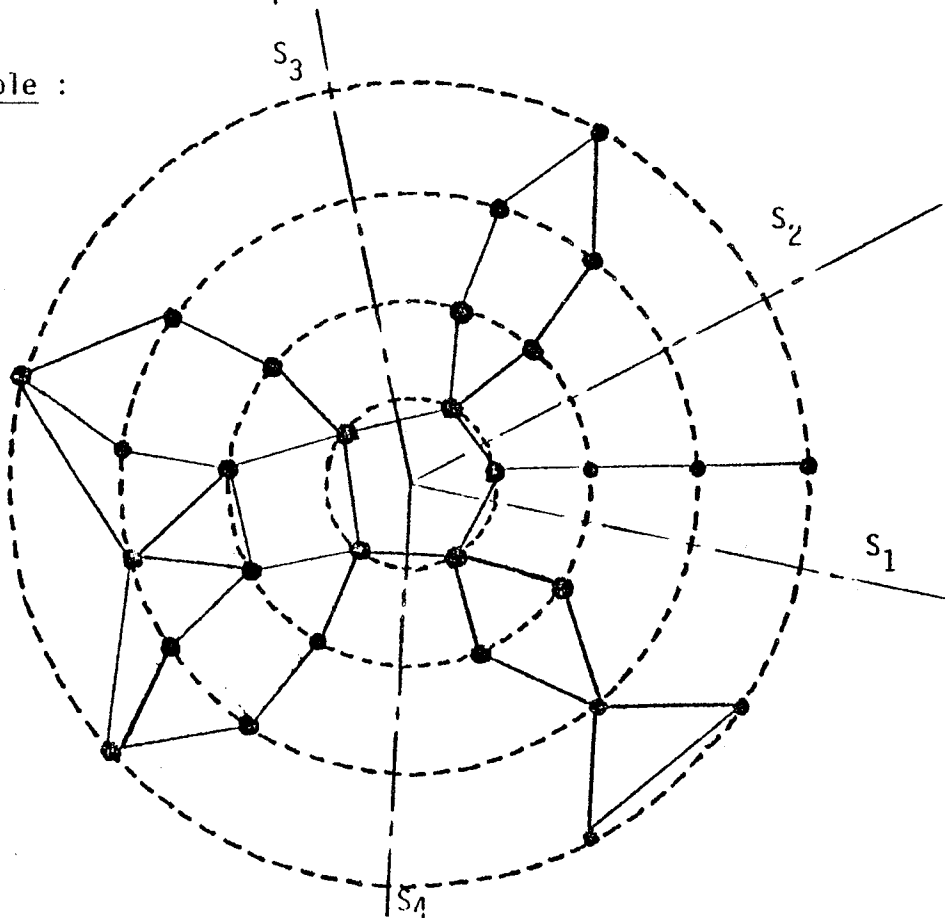
### 5. FORME CENTRIFUGE : L'HEURISTIQUE DES ANGLES POLAIRES

Pour utiliser au mieux la forme du dessin, on va ici, placer les sommets en utilisant un système de coordonnées polaires  $(\rho, \theta)$  où  $\rho$  sera calculé en fonction des niveaux, et où l'heuristique se réduira donc à la détermination des angles polaires  $\theta$  des sommets du graphe.

On commence par chercher si les centres constituent un ensemble d'articulation. Si oui, on positionne indépendamment les sommets des sous graphes ainsi trouvés dans des secteurs circulaires puisque ces sous graphes ne sont reliés que par l'intermédiaire des centres.

Les secteurs circulaires ont une taille angulaire proportionnelle au nombre de sommets qu'ils contiennent.

Exemple :



- Figure 23 -

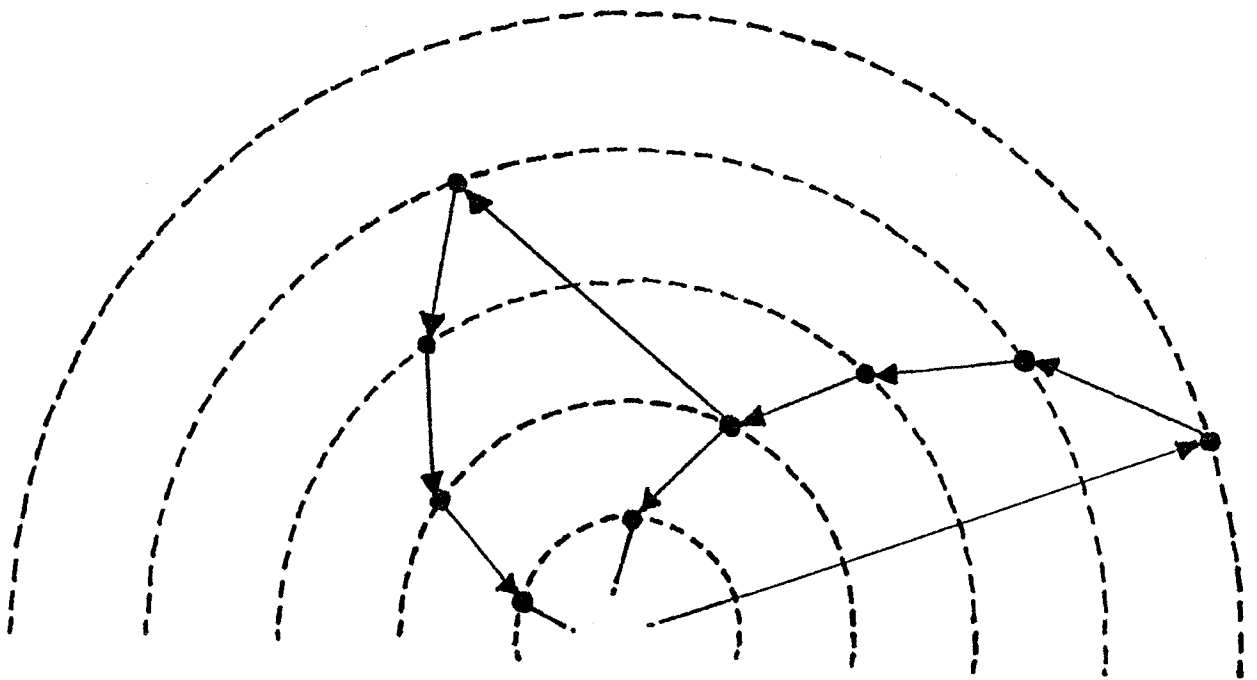
Ici, le plan est découpé en quatre secteurs :

$S_1$  contient 3 sommets,  $S_2$  en contient 5 et  $S_3$  et  $S_4$  respectivement 11 et 5.

Dans chaque secteur, on commence par positionner un sommet de faible degré, éloigné du centre, puis on positionne ses successeurs et prédécesseurs à intervalle angulaire régulier, en tournant toujours dans le même sens (sauf pour les sommets liés à plusieurs sommets déjà placés, pour qui on fera la moyenne des angles de leurs voisins) et en donnant la priorité au sommet le plus proche du centre, lorsque l'on a plusieurs candidats au positionnement (les autres étant mis sur une liste d'attente). Chaque fois qu'on arrive sur un centre, on prend le premier sommet dans la liste d'attente. S'il y a plusieurs sommets centres, on les positionne sur le plus petit cercle d'après le nombre d'arcs qui les lient à chaque secteur.

Quand les sommets sont tous placés, on repasse en coordonnées cartésiennes pour l'affichage du dessin.

Dans les cas où les arcs ne sont pas trop nombreux, on obtient un dessin de ce type :



- Figure 24 -

(Les circuits articulés sur les centres sont représentés par l'heuristique sous une forme rappelant des pales d'hélices). Ici aussi, chaque position calculée est définitive, ce qui donne des résultats assez rapides.

La méthode d'empilement des successeurs non placés revient à simuler un algorithme récursif de placement des sommets par la gestion explicite d'une pile. Le fait que l'on parte des sommets périphériques (qui sont toujours de degré plus faible que les autres) d'une part diminue la dose d'arbitraire inhérente à ce type d'algorithme, et, d'autre part, garantit que l'on n'empilera pas trop de sommets en attente de placement puisque les sommets de fort degré auront une partie de leurs sommets adjacents déjà placés.

Pour les raisons invoquées au § 1.2., on n'a pas étudié le problème de la minimisation du nombre de croisements. Le déplacement interactif des sommets à l'écran est bien sûr une possibilité offerte à l'utilisateur. Le lecteur se reportera au chapitre 6 où sont donnés de nombreux exemples de dessins.





## CHAPITRE 4

TRAITEMENT PREPARATOIRE DES TABLEAUX DE DONNEES :

LE SEUILLAGE

Si le graphe à hiérarchiser est défini par la liste de ses arcs ou par sa matrice d'incidence, il peut être traité directement.

Mais, en général, les données concrètes (portant, par exemple, sur des échanges entre pays, ou entre branches industrielles) sont présentées sous la forme d'un tableau carré à coefficients entiers ou réels.

La manière la plus simple de transformer ces tableaux en des matrices booléennes qu'on puisse interpréter comme des matrices d'incidence de graphes, consiste à en prendre la silhouette :

Exemple :

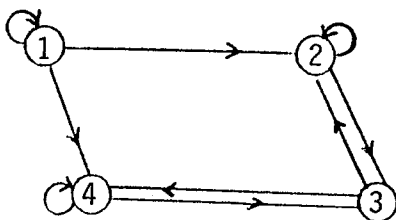
	1	2	3	4
1	150	100	0	10
2	0	120	5	15
3	0	100	0	50
4	0	20	75	10

Matrice de données M

1	1	0	1
0	1	1	1
0	1	0	1
0	1	1	1

Matrice A = silhouette de M.

Graphe associé :



En procédant ainsi, on ne conserve qu'une vision très schématique de la structure du tableau initial, puisque les différences quantitatives dans M disparaissent dans A, et donc aussi du graphe associé.

Cette trop grande simplification entraîne que l'on n'a guère de chances d'obtenir une structure hiérarchique utile pour la compréhension du système. Pour extraire des données numériques des informations intéressantes et faire ressortir les lignes de force du réseau de relations étudié, il semble logique de privilégier les données qui pèsent le plus et de négliger les valeurs *quantitativement* trop faibles pour être *qualitativement* significatives ; cette idée nous amène à effectuer diverses opérations de *seuillage* sur les tableaux de données initiales.

## 1. LE SEUILLAGE

### Définitions :

Soit M une matrice carrée  $n \times n$  à coefficients entiers ou réels, on appelle seuillage de M par la condition C(s) l'application qui à M associe une matrice booléenne A de même taille de la façon suivante :

si C(s) est vérifiée pour  $m_{ij}$ , alors  $a_{ij} = 1$ , sinon  $a_{ij} = 0$   
pour  $i, j = 1, 2, \dots, n$

s est un paramètre appelé *niveau de seuil* (il peut être vectoriel)

A est appelée *matrice seuillée* au niveau s.

Cette opération permet d'associer la valeur 1 aux  $a_{ij}$  seulement lorsque les  $m_{ij}$  correspondants dépassent une certaine valeur, qui dépend du choix effectué pour la condition  $C(s)$  et du niveau de seuil  $s$ .

## 2. DIVERSES CONDITIONS DE SEUIL

### 2.1. Seuillage absolu

On construit la matrice seuillée  $A$  en appliquant à chaque  $m_{ij}$  une condition  $C_a(s)$  de la forme :

$$m_{ij} \geq s$$

### 2.2. Seuillages relatifs

On construit  $A$  en tenant compte pour chaque  $m_{ij}$  de son poids relatif dans sa ligne et/ou sa colonne, ce qui donne trois expressions possibles pour la condition de seuil,

\* seuillage relatif en ligne  $C_\ell(s)$  : 
$$\frac{m_{ij}}{\sum_j m_{ij}} \geq s$$

\* seuillage relatif en colonne  $C_c(s)$  : 
$$\frac{m_{ij}}{\sum_i m_{ij}} \geq s$$

\* seuillage relatif mixte  $C_m(s)$  : 
$$\left( \frac{m_{ij}}{\sum_j m_{ij}} \geq s' \right) \text{ ou } \left( \frac{m_{ij}}{\sum_i m_{ij}} \geq s'' \right)$$

Dans ce dernier cas, le niveau de seuil est un vecteur  $s = (s', s'')$ .

### 2.3. Remarques

On peut évidemment construire autant de seuillages *hybrides* que l'on veut, par combinaison de seuillages absolus et relatifs, par exemple :

$$C_h(s) : [(m_{ij} \geq s_1) \text{ et } (\frac{m_{ij}}{\sum_i m_{ij}} \geq s_2)] \text{ ou } (\frac{m_{ij}}{\sum_j m_{ij}} \geq s_3)$$

L'inconvénient de conditions de ce type est la nécessité de fixer simultanément les trois composantes ( $s_1, s_2, s_3$ ) du paramètre vectoriel de niveau.

L'utilisation des seuillages absolus est classique en analyse d'image : on travaille sur des images "digitalisées", découpées en petites cases auxquelles sont associés des niveaux d'intensité lumineuse qui jouent le même rôle que les coefficients numériques de nos tableaux de données initiales.

En choisissant bien les niveaux de seuil, on peut par exemple isoler des cellules sur un fond plus clair, ou des noyaux de cellules, plus sombres que les cytoplasmes.

Il est clair que, dans ce contexte, des seuillages relatifs n'auraient aucun sens, puisque le découpage de l'image est totalement arbitraire.

Au contraire, dans le cas de matrices de données concernant les interactions dans un système, les lignes et les colonnes correspondent aux éléments du système, et l'utilisation de seuillages relatifs peut être fort utile.

### 2.4. Utilisation des seuillages

Le choix de l'une ou de l'autre des deux méthodes dépend de l'application envisagée, chaque méthode ayant ses inconvénients : le seuillage absolu privilégie la valeur numérique brute, sans tenir compte

de la différence éventuelle des poids globaux représentés par les deux extrémités d'un arc donné, d'où une élimination trop brutale d'informations utiles à la compréhension de la structure des flux ; par exemple, si dans un tableau d'import-export, le coefficient  $m_{ij}$  est faible en valeur, mais s'il représente 80% des exportations totales du pays j, sa suppression par seuillage peut alors être ennuyeuse.

Au contraire, les seuillages relatifs, et particulièrement le seuillage mixte, respectent bien ce genre de données, mais négligent les valeurs réelles des flux, ce qui peut aboutir à des résultats paradoxaux, surtout dans les cas où l'on emploie des niveaux de seuils élevés : par exemple, un sommet du graphe qui représenterait un gros volume total pourrait se trouver complètement isolé si tous les flux entrants (et/ou sortants) ont des valeurs voisines (et par conséquent des poids relatifs équivalents qui peuvent être faibles si les arcs sont nombreux). Néanmoins l'emploi de seuils relatifs nous paraît préférable à celui de seuils absolus, pour 2 raisons :

- \* le seuillage absolu interdit par nature toute péréquation liée aux sommets et par là se montre souvent peu "structurant".

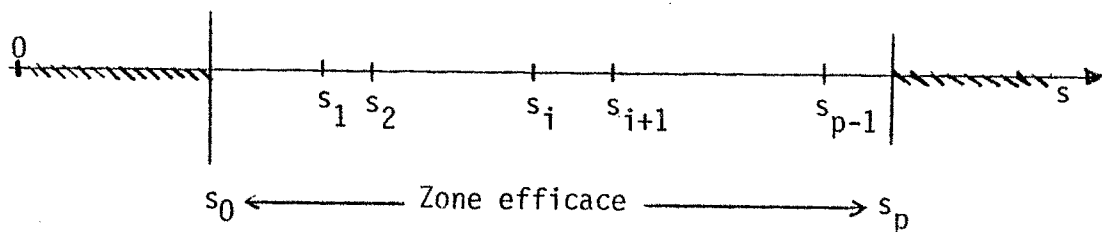
- \* il est plus facile et plus parlant de tenir compte des valeurs brutes par des moyens visuels (épaisseurs variables attribuées aux arcs) ou de rechercher les valeurs initiales exactes des coefficients du tableau après un seuillage relatif que de construire des données complémentaires sur les poids relatifs après un seuillage absolu.

### 3. LES PROBLEMES D'INTERPRETATION

Une fois adoptée la méthode de seuillage, il faut sélectionner le (ou les) niveaux de seuil qui sont les plus "significatifs", d'après la structure hiérarchique à laquelle ils aboutissent.

Les critères de sélection sont à la fois objectifs et subjectifs ; en général le problème se pose de la manière suivante (quel que soit le type de seuillage utilisé) : pour un niveau de seuil faible ou nul, le graphe est très touffu, l'analyse des structures hiérarchiques ne donne pas grand chose (souvent un seul niveau) ; les différents sommets étant très liés les uns aux autres, ils jouent des rôles peu différenciés dans la structure du système. Pour un niveau de seuil très élevé, les arcs ont tous été supprimés, les sommets sont tous isolés et on n'a pas non plus de structure intéressante (hiérarchie des rangs réduite à un niveau, hiérarchie des écartements impossible à construire). Entre ces deux cas extrêmes, il y a une "zone efficace" plus ou moins grande à l'intérieur de laquelle apparaissent différentes structures, qui correspondent à des plages de seuil de largeurs variables  $[s_i, s_{i+1}[$

En général on obtient schématiquement :



- Figure 25 -

On a  $p$  plages de seuillage  $[s_i, s_{i+1}[$  correspondant chacune à une certaine structure hiérarchique pour le graphe associé à la matrice seuillée par  $s$ ,  $s_i \leq s < s_{i+1}$ . Le nombre de plages  $p$  est borné par le nombre d'arcs contenus dans le graphe (au maximum,  $p = n^2$  si  $n$  est le nombre de sommets du graphe), puisque chaque changement de structure hiérarchique ne se produit qu'avec la disparition d'au moins un arc. L'intervalle  $[s_0, s_p]$  définit la zone efficace de seuillage, dans laquelle on va chercher à mettre en évidence, sinon les  $p$  structures hiérarchiques existantes, du moins les plus caractéristiques d'entre elles. En pratique, on effectue divers seuillages à intervalle régulier dans  $[s_0, s_p]$ . Par exemple, si on a :

$s_0 = 1\%$  et  $s_p = 35\%$  (en seuillage relatif) on étudie les structures hiérarchiques obtenues pour des seuils 5, 10, 15, 20, 25 et 30%, puis, suivant les résultats obtenus, on refait ou non des seuillages intermédiaires.

Pour effectuer une sélection dans la zone efficace, sur la base des dessins obtenus, plusieurs critères interviennent :

1 \* la lisibilité du graphe, d'autant meilleure qu'il y a moins d'arcs.

2 \* la différenciation hiérarchique, ce qui revient à choisir les images comportant le plus grand nombre de niveaux hiérarchiques.

3 \* l'adéquation de la structure observée à l'idée qu'on se faisait a priori de la structure du système.

Ces trois critères sont plutôt subjectifs, aussi a-t-on intérêt à leur adjoindre trois critères objectifs de sélection, qui sont liés aux plages de seuillage.

4 \* le niveau de seuil : une structure hiérarchique reflètera d'autant mieux l'agencement interne du système correspondant qu'elle sera obtenue pour un seuil plus bas.

5 \* la largeur de la plage de seuillage : plus la plage de seuillage  $[s_i, s_{i+1}[$  associée à une structure hiérarchique donnée est large, plus cette structure apparaît comme stable dans le système considéré, par opposition aux structures transitoires qui sont liées à des seuils précis, et donc à des plages de seuillage étroites, et qui risquent ainsi d'être moins pertinentes pour la compréhension du système.

6 \* la ressemblance avec les structures voisines : ce critère est voisin du précédent en ce sens qu'il permet aussi d'apprécier la stabilité d'une structure hiérarchique ; en effet si on compare la structure obtenue par la plage de seuillage  $[s_i, s_{i+1}[$  aux structures obtenues pour



les plages de seuillages voisines  $[s_{i-1}, s_i[$  et  $[s_{i+1}, s_{i+2}[$  on peut espérer évaluer qualitativement si une des bornes  $s_i$  ou  $s_{i+1}$  marque un changement de structure important.

Le premier critère et, dans une moindre mesure, le second, poussent à utiliser des niveaux de seuil assez élevés, pour obtenir des structures à la fois claires et assez finement hiérarchisées. Le quatrième critère est donc contradictoire avec les deux premiers, mais on aura rarement des structures intéressantes à des niveaux de seuil très bas.

Le troisième critère est nocif, puisqu'il pousse à privilégier a priori, telle ou telle structure observée.

Or, la seule objectivité ici possible est l'exhaustivité : parcourir la zone de seuillage avec un pas de variation du niveau de seuil suffisamment fin pour ne laisser échapper aucun des schémas structurels possibles. Ensuite, tout ce que peut apporter la méthode est d'ordre critique et purement négatif : indiquer à l'utilisateur que son idée a priori était fautive si la structure qu'il s'attendait à voir apparaître n'existe pour aucun niveau de seuil, mais en aucun cas l'existence d'une image attendue ne peut certifier une plus grande portée explicative du schéma préconçu.

Les cinquième et sixième critères posent le problème de la stabilité des structures observées, sur lequel nous reviendrons au chapitre 6. En pratique, on construira des matrices *multiseuillées* en "empilant" les matrices obtenues pour des niveaux de seuils croissants :

0	1	0	1	0
0	0	1	1	0
0	1	0	0	1
0	1	1	0	0
0	0	1	1	0

seuil  $s_1$

0	1	0	0	0
0	0	0	1	0
0	1	0	0	1
0	1	0	0	0
0	0	1	0	0

seuil  $s_2 > s_1$

0	2	0	1	0
0	0	1	2	0
0	2	0	0	2
0	2	1	0	0
0	0	2	1	0

matrice multiseuillée  
( $s_1, s_2$ )

Prendre dans la matrice multiseuillée tous les coefficients  $m_{ij} \geq k$  restitue la matrice booléenne du même seuil.

Cette méthode permet de tester rapidement toute une série de seuils, et de mieux approximer les différentes plages  $[s_i, s_{i+1}]$  dans la zone efficace.



## CHAPITRE 5

DESCRIPTION DES LOGICIELS
---------------------------

### 1. TRAITEMENT DES GRAPHEs BIPARTIS

En dehors des programmes standard utilisés pour l'étude expérimentale de la méthode itérative des degrés relatifs (méthode QR pour le calcul des valeurs propres et méthode de la puissance), nous avons écrit un programme utilisant la méthode des degrés relatifs de manière heuristique.

Ce programme est écrit en Algol W et contient 370 instructions. Son principe est le suivant :

1. A partir du graphe initial (stocké sous forme de listes d'arcs) on construit un graphe de taille plus petite en regroupant arbitrairement en classes les sommets sur leurs niveaux respectifs.

2. On applique au graphe obtenu un pas d'itération de la méthode des degrés relatifs, ce qui permet de permuter éventuellement certains groupes de sommets entre eux. (Pour calculer la pondération par les degrés relatifs on dénombre tous les arcs entre sommets des groupes).

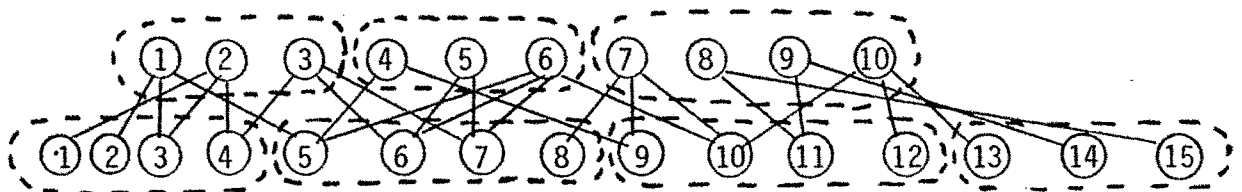
3. On "éclate" les classes du premier niveau en classant les sommets à l'intérieur de chaque classe par le critère des degrés relatifs les liant aux classes de l'autre niveau, et on réordonne éventuellement les sommets de chaque classe par rapport à ceux des classes voisines.

4. On effectue la même démarche pour éclater les groupes de l'autre niveau, avec cette différence que le critère des degrés relatifs est appliqué par rapport aux *sommets* du premier niveau, puisque les classes y ont déjà été éclatées.

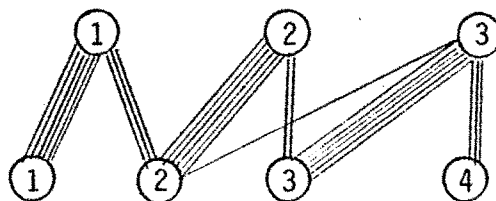
Ce programme a été utilisé sur les mêmes exemples que ceux qui sont cités au chapitre 3 (graphes du CTGREF) et il s'est révélé efficace et surtout très rapide : Sur l'exemple (10,28) qui donnait un minimum de 302 croisements il a fourni une configuration à 306 croisements (à partir du même point de départ) et sur l'exemple (24,98) il a calculé 4900 croisements en 5 s, alors que la méthode de la puissance donnait un minimum de 4400 croisements en 5 itérations, soit 30 s de calcul !

Soit un graphe biparti  $(k,m)$  on s'arrange pour construire des classes ayant respectivement sur chaque niveau environ  $\sqrt{k}$  et  $\sqrt{m}$  sommets, ce qui donne  $\sqrt{k}$  et  $\sqrt{m}$  classes.

Exemple : graphe initial (10,15) avec 31 croisements :



On a 3 classes en haut, et 4 en bas :



- Figure 26 -



## 2. LES LOGICIELS LØVIGREC et CØPERNIC

LØVIGREC (Logiciel d'Ordonnancement et de Visualisation Interactive de Graphes Réduits Eclatés et des Composantes) et CØPERNIC permettent de hiérarchiser et de visualiser des graphes, respectivement sous la forme réduite éclatée et sous la forme centrifuge (cf. chapitre 2)

### 2.1. Infrastructure

Ces deux logiciels fonctionnent au CICG (Centre Interuniversitaire de Calcul de Grenoble) respectivement depuis Juin 1978 et Janvier 1979, sur une machine virtuelle de 512K de l'ordinateur IBM 360/67 sous CP-CMS.

Ils fournissent, par l'intermédiaire du logiciel GRIGRI (Leduc-Leballeur [13], Lucas [15]) des dessins sur un terminal graphique Tektronix 4010. Des dessins sur papier peuvent être obtenus grâce au logiciel SOFTBEN sur le traceur BENSON. Ils ont été écrits en Fortran et représentent un ensemble de programmes assez volumineux : plus de 3000 instructions Fortran et un encombrement de l'ordre de 75K au chargement. Nous allons décrire le fonctionnement de ces logiciels du point de vue de l'utilisateur en faisant ressortir les différentes options offertes et les éléments de dialogue.

Ensuite, nous esquisserons une description interne des programmes, en nous limitant essentiellement à la structure de données, les principes des algorithmes de calcul ayant été donnés aux chapitres précédents.

Nous essaierons surtout de faire ressortir les principes adoptés pour le découpage des sous-programmes et d'opérer une classification fonctionnelle de ceux-ci.

### 2.2. Description externe

Ces deux logiciels peuvent être utilisés de façon indépendante. Néanmoins LØVIGREC "contient" CØPERNIC, du fait que la méthode de hiérarchisation des écartements est utilisée pour représenter la structure interne des composantes fortement connexes d'un graphe "réduit éclaté".

Nous allons décrire le fonctionnement de LØVIGREC, puis nous préciserons les quelques points sur lesquels CØPERNIC diffère de LØVIGREC.

Nous traiterons successivement :

- les différents types d'entrées possibles pour les données initiales.
- la structure du dialogue avec l'utilisateur
- les sorties graphiques.

### 2.2.1. Les entrées

Les formes d'entrées actuellement disponibles sont de quatre types :

- T1 - liste des coefficients non nuls
- T2 - matrice carrée à coefficients entiers ou réels, positifs ou nuls.
- T3 - matrice multiseuillée (cf. chapitre 4)
- T4 - liste des arcs du graphe.

Le programme d'initialisation générale demande à l'utilisateur d'indiquer le type d'entrée dont il se sert et prend en charge l'altération des noms de fichiers de données, le chargement et l'exécution des programmes, à partir des indications fournies par l'utilisateur. Pour lancer l'exécution de ce programme initial il suffit de taper la commande IMAG au clavier.

Etant donnée la manière dont le graphe est représenté en mémoire (cf. Description interne), seul le type T4 est directement interprétable.

Le type T3 requiert l'indication du niveau de seuil adopté par l'utilisateur.

Les types T1 et T2 exigent une opération préalable de seuillage, opération qui permet de se ramener au type T3, selon une des méthodes décrites au chapitre 4.

Le cas où le graphe est donné sous la forme d'une matrice d'incidence booléenne est considéré comme faisant partie du type T3 (avec un seul niveau de seuil).

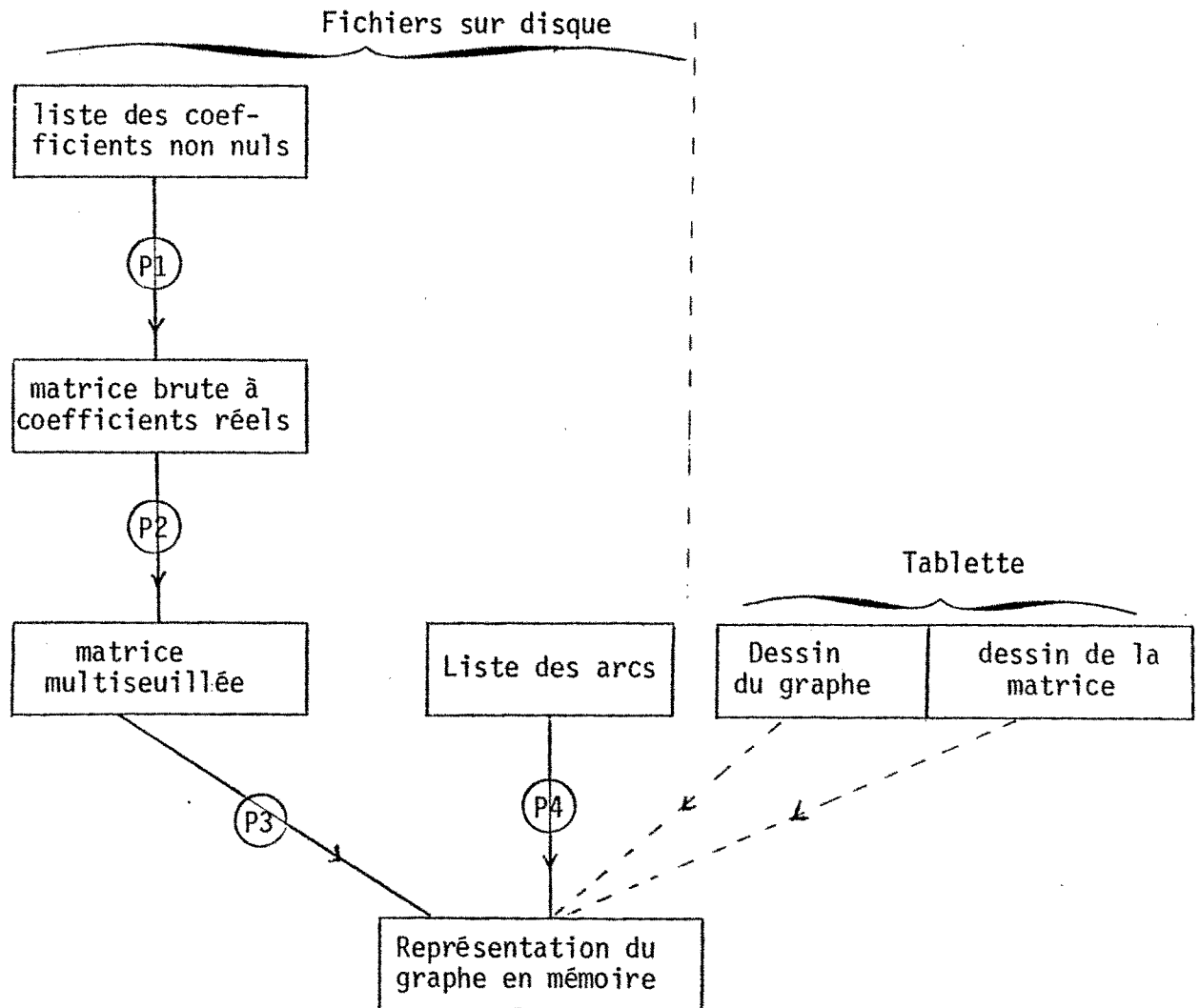


Deux autres formes d'entrées sont envisagées, pour l'introduction de données à partir d'une tablette graphique :

T5 - dessin du graphe sur la tablette

T6 - pointage des coefficients non nuls d'une matrice sur la tablette

Soit le schéma :



- Figure 27 -

Les différents programmes de passage d'un type de données à un autre type sont P1, P2, P3 et P4.

P1 est un programme de reformatage et P4 un programme de lecture qui ne demandent aucune intervention à l'utilisateur.

P2 est le programme de seuillage dont l'activation entame une phase de dialogue avec l'utilisateur en lui demandant successivement d'indiquer :

- la méthode de seuillage voulue
- le nombre de seuils souhaités (maximum : 9)
- les niveaux de seuils (par ordre croissant)

Quatre méthodes de seuillage sont disponibles :

- seuillage absolu
- seuillage relatif en lignes
- seuillage relatif en colonnes
- seuillage relatif mixte.

Dans un souci de simplification, les seuillages hybrides ont été exclus, et le seuillage mixte est effectué avec le même niveau en ligne et en colonne. Ainsi, un seul paramètre de niveau  $s$  est demandé à l'utilisateur.

Dans les notations du chapitre 4 les quatre méthodes retenues ici s'écrivent donc :  $C_a(s)$ ,  $C_l(s)$ ,  $C_c(s)$  et  $C_m(s,s)$ .

P3 est le programme d'interprétation d'une matrice multiseuillée. Il demande à l'utilisateur de choisir un seuil, ce qui s'effectue en tapant le numéro d'ordre du seuil pour lequel il veut voir le graphe (1 si la matrice est simplement booléenne, d'après la remarque faite plus haut).

La structure du graphe étant stockée en mémoire, les sous programmes de réduction et de hiérarchisation sont activés sauf si l'utilisateur indique que le graphe est sans circuit : alors la hiérarchisation est effectuée directement. Le résultat en est fourni à l'utilisateur sous la forme suivante :

NOMBRE DE NIVEAUX : nivmax  
NOMBRE DE CFC : ncf

Ensuite sont affichés les tableaux des niveaux et des c.f.c., suivis de la question : VOULEZ-VOUS DES IMAGES ?

Jusqu'à ce point de l'exécution le dialogue s'est effectué par des ordres d'écriture et de lecture à la console, et aucune commande graphique n'a encore eu lieu.

Ceci permet d'utiliser le programme de réduction et hiérarchisation sans être obligatoirement connecté à un terminal graphique. (cas d'un terminal indisponible, essais préalables).

Maintenant, si la réponse est OUI, le programme entre dans sa phase proprement graphique, et le dialogue avec l'utilisateur se fait par des "MENU" et des "MESSAGE" (primitives de GRIGRI) ; nous allons décrire le dialogue à l'écran.

#### 2.2.2. Structure du dialogue graphique

L'instrument essentiel du dialogue est la primitive "MENU" de GRIGRI qui permet à l'utilisateur de faire son choix parmi une série de possibilités en désignant sur l'écran un nom dans une liste : MENU affiche dans le coin de l'écran la liste des noms désignant les différentes options possibles. L'utilisateur effectue alors son choix en pointant un nom avec le réticule. La primitive "MESSAGE" est utilisée pour les parties du dialogue qui requièrent de la part du programme ou de l'utilisateur la fourniture de données alphanumériques (numéros, valeurs etc ...).

Le principe retenu pour la mise en place des "MENU" a été une structure hiérarchisée, correspondant sur le plan graphique au passage d'un type d'affichage à un autre ; la structure comprend quatre niveaux :

- \* un niveau préliminaire qui permet à l'utilisateur d'effectuer un *affichage global*, ou de changer l'affichage global présent sur l'écran.

- \* un niveau de *modification* de l'image globale et de demande *d'informations nouvelles*.

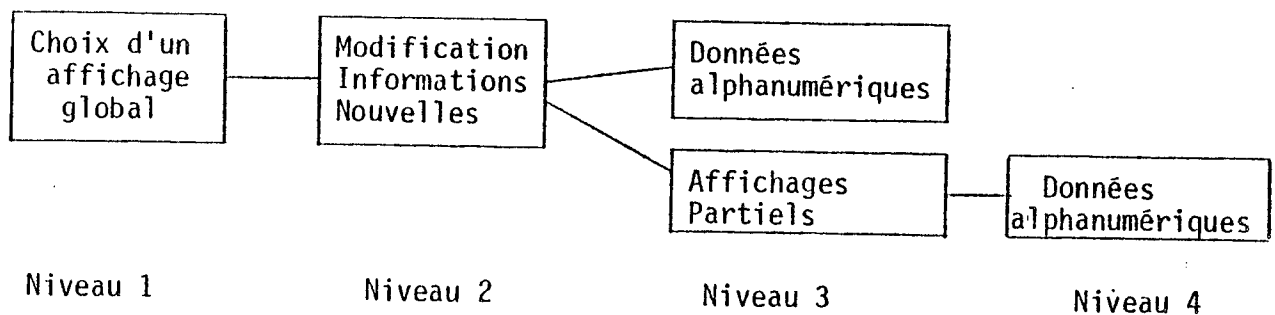
\* un niveau décrivant les deux types d'informations possibles :

1 \* *type alphanumérique* : affichage de données avec l'image globale

2 \* *type graphique* : *images partielles*

\* un niveau redonnant la possibilité d'afficher des données avec les *images partielles*.

D'où le schéma :



- Figure 28 -

Dans chaque MENU, la commande FIN permet la remontée au niveau précédent, sauf au niveau 1, où la commande ARRET termine la phase d'affichage.

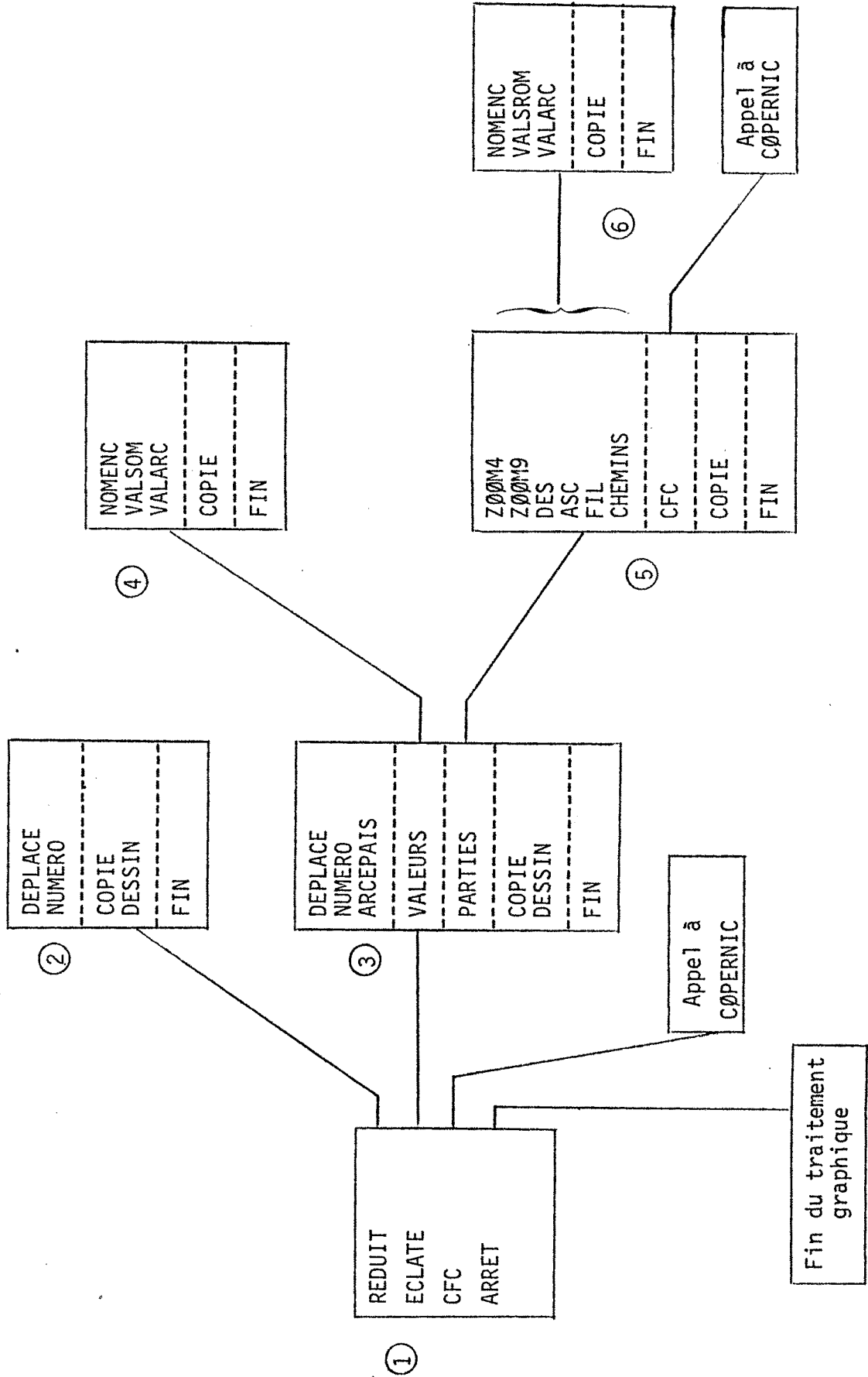
### 2.3. Les sorties

Les sorties prévues sont de trois types :

- A - dessins à l'écran
- B - dessins sur papier
- C - stockage sur fichier

Les sorties à l'écran sont générées par le programme, en utilisant les primitives de GRIGRI. Les images ainsi obtenues sont modifiables par interactivité.

- Figure 29 -



Ces dessins sur papier permettent de conserver une trace de l'image existant sur l'écran à un moment donné, sous une forme plus ou moins élaborée.

Le stockage sur fichier (qui n'existe pas actuellement) a pour but de conserver une image sous forme numérique (liste de coordonnées) pour une réutilisation ultérieure (par exemple, lors d'une phase de simulation sur le graphe), avec l'avantage de ne pas recommencer à ce moment toutes les phases de calcul précédant le dessin conservé (hiérarchisation, calcul des coordonnées, modifications éventuelles du dessin).

Dans les MENU, les dessins sur papier correspondent aux commandes COPIE et DESSIN.

Le stockage sur fichier prendra place dans les MENU aux mêmes endroits que la commande DESSIN, sous la forme d'une commande FICHER.

Toutes les autres commandes servent à l'affichage de dessins ou de données alphanumériques et à la modification interactive des dessins, elles font donc partie de la rubrique "sorties écrans".

1. Les sorties écrans se répartissent en trois classes :

- les dessins globaux (et leurs modifications)
- les informations alphanumériques auxiliaires.
- les dessins partiels.

Les dessins globaux sont affichés sur appel des commandes :

REDUIT, ECLATE ou CFC

Leurs modifications sont effectuées par les commandes DEPLACE, NUMERO et ARCEPAIS.

Les affichages de données auxiliaires (regroupés dans la commande VALEURS du menu ③ ) sont obtenus par les commandes NOMENC, VALSOM et VALARC (menus ④ et ⑥ ).

Les affichages de dessins partiels (regroupés dans la commande PARTIES du menu ③ ) correspondent aux commandes :

ZØØM4, ZØØM9, DES, ASC, FIL, CHEMINS et CFC

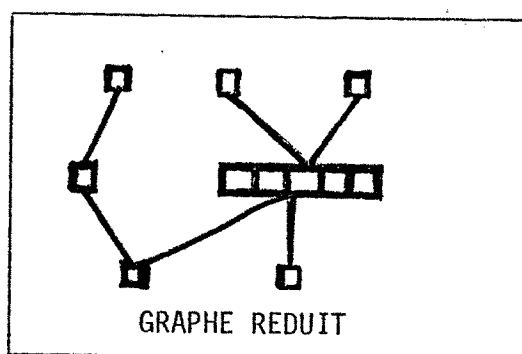
Remarque :

Vu son importance particulière, la commande CFC est considérée à la fois comme un dessin global et comme un dessin partiel.

Nous allons maintenant décrire ces dix neuf commandes du point de vue de l'utilisateur, c'est-à-dire en donnant ce qui est vu sur l'écran, ainsi que les paramètres éventuellement liés à certaines d'entre elles .

MENU ① ;

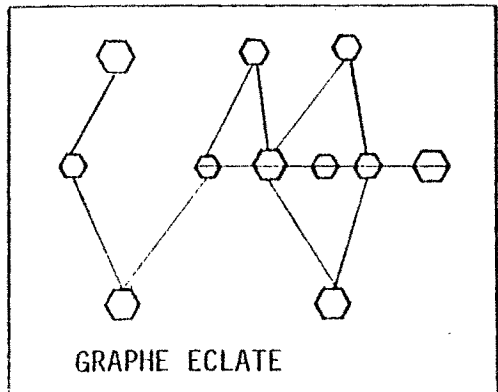
\* commande REDUIT : provoque l'affichage du graphe réduit (cf. chapitre 2) sous la forme suivante :



Chaque cfc est symbolisée par un ● "domino" contenant autant de petits carrés qu'il y a de sommets du graphe dans cette cfc.  
Le message "GRAPHE REDUIT" s'affiche en bas de l'écran.

Les commandes de MENU ② sont ensuite proposées au choix de l'utilisateur.

\* commande ECLATE : provoque l'affichage du graphe éclaté :



Chaque sommet est symbolisé par un hexagone.

Les tailles des sommets peuvent être variables (cf. données auxiliaires). Le message "GRAPHE ECLATE" s'affiche en bas de l'écran.

L'utilisateur est sollicité par le MENU ③ .

\* commande CFC : cette commande envoie à l'utilisateur un message : "SOMMET CONCERNE ? Z", qui s'inscrit en haut de l'écran.

Celui-ci doit répondre en tapant au clavier le numéro d'un des sommets de la cfc qui l'intéresse dans la zone Z du message. Ensuite, la commande provoque un appel à CØPERNIC, qui donnera une représentation centrifuge de la cfc considérée.

MENU ② ;

\* commande DEPLACE : cette commande permet de modifier l'image du graphe réduit de la manière suivante :

elle envoie un message : "COMBIEN DE SOMMETS DEPLACEZ-VOUS ? Z" en haut de l'écran. L'utilisateur répond un nombre k (au plus égal à 10) dans la zone Z.

Ensuite l'utilisateur désigne alternativement une des cfc qu'il veut déplacer puis le nouvel endroit de l'écran où il veut la mettre. Quand les 2k points attendus ont été fournis, une image tenant compte des modifications demandées est affiché.

En cas d'erreur de pointage d'une cfc, la cfc concernée n'est pas déplacée.



Remarque :

Les changements de niveaux ne sont pas admis, les modifications ne se font donc qu'horizontalement.

\* commande NUMERO : cette commande permet d'afficher à côté de chaque carré le numéro du sommet correspondant (qui est évidemment arbitraire pour les cfc contenant plusieurs sommets). Cet affichage ne fait pas partie des données alphanumériques auxiliaires, puisque le numéro donné à chaque sommet est un numéro d'ordre arbitrairement fixé par le programme (par exemple dans l'ordre de lecture des lignes de la matrice d'incidence).

Dans tous les cas où une commande exige que l'utilisateur donne un numéro de sommets (par exemple la commande CFC), la commande NUMERO est exécutée préalablement, pour éviter toute erreur de désignation, l'utilisateur ne connaissant pas a priori les numéros affectés aux sommets sur l'écran.

(Les commandes COPIE et DESSIN sont décrites ultérieurement).

MENU ③ ;

Les commandes DEPLACE et NUMERO fonctionnent sur le graphe éclaté exactement de la même façon que sur le graphe réduit, avec la seule différence qu'elles portent sur les sommets (hexagones) au lieu des cfc ("dominos").

\* commande ARCEPAIS : cette commande permet de modifier le dessin du graphe éclaté en attribuant aux arcs une épaisseur liée à la valeur réelle du coefficient correspondant. Si les données permettant d'effectuer cette opération n'existent pas, la commande est sans effet.

ARCEPAIS n'aurait pas de sens dans le MENU ② , car chaque arc entre deux cfc peut représenter plusieurs arcs du graphe éclaté.

Les commandes VALEURS et PARTIES provoquent respectivement l'affichage du MENU ④ et du MENU ⑤ , sans changement du dessin présenté à l'écran (graphe éclaté).

MENU (4) :

\* commande NOMENC : cette commande affiche à côté de chaque hexagone sa nomenclature abrégée (3 lettres). Si les sommets du graphe n'ont pas de nom, cette commande est sans effet.

\* commande VALSOM : cette commande affiche le message suivant :

SOMMET CONCERNE ?  $\underline{Z_1}$  VALEUR BRUTE :  $\underline{Z_2}$

L'utilisateur tape dans  $Z_1$  le numéro du sommet. En réponse, la valeur numérique réelle associée à ce sommet (lors du seuillage) est affichée dans  $Z_2$ .

\* commande VALARC : cette commande affiche le message :

SOMMET INITIAL :  $\underline{Z_1}$  FINAL :  $\underline{Z_2}$  VALEUR BRUTE :  $\underline{Z_3}$

L'utilisateur tape dans  $Z_1$  et  $Z_2$  les numéros des extrémités de l'arc qui l'intéresse. La réponse en  $Z_3$  est la valeur du coefficient correspondant dans la matrice non seuillée.

En cas d'erreur, la réponse est : "ARC INEXISTANT".

Si les données brutes n'existent pas (graphe non issu d'un seuillage) les commandes VALSOM et VALARC sont sans effet.

MENU (5)

Ce menu récapitule toutes les possibilités de dessins partiels extraits du graphe éclaté, à savoir :

1 \* des effets de loupe : ZØØM4 et ZØØM9.

2 \* des chemins arrivant et/ou partant d'un sommet donné (FIL, ASC, et DES) ou entre deux sommets donnés (CHEMINS).

3 \* des composantes fortement connexes en hiérarchie centrifuge. (commande CFC : voir MENU (1) ).

\* commande ZØØM4 : cette commande donne une image agrandie d'une portion du graphe, en affichant sur tout l'écran un quart de la surface du dessin, centré sur un point fourni par l'utilisateur.

Elle envoie donc le message "POINTEZ LE CENTRE DE LA LOUPE" puis attend que l'utilisateur ait désigné un point avec le réticule, dans la région qui l'intéresse.

En réponse, la commande substitue au graphe éclaté l'image locale agrandie, accompagnée du message : AGRANDISSEMENT DEMANDE.

La commande ZØØM9 fonctionne comme ZØØM4, mais pour 1/9ème de la surface. L'effet de loupe est donc plus fort.

\* commande ASC : cette commande affiche sur l'écran tous les chemins (élémentaires ou non) qui ont un sommet donné x pour extrémité finale à un ordre donné k.

Le dessin a la forme d'une arborescence de racine x, et contient tous les sommets qui sont des prédécesseurs de x à une distance au plus égale à k.

Le message suivant est envoyé pour demander les paramètres :

SOMMET CONCERNE ?  $\underline{Z_1}$     ORDRE VOULU ?  $\underline{Z_2}$

L'utilisateur tape le numéro de x dans  $Z_1$  et l'ordre k dans  $Z_2$ .

Les commandes DES et FIL fonctionnent comme ASC, mais DES fournit l'arborescence des successeurs et FIL la réunion des deux arborescences précédentes (cf. Figure 30) dessinée chacune sur un demi-écran.

\* commande CHEMINS : cette commande fonctionne à partir de trois paramètres, deux numéros de sommets x et y et un ordre de calcul k, qui sont fournis par l'utilisateur dans les trois zones du message :

SOMMET INITIAL ?  $\underline{Z_1}$     SOMMET FINAL ?  $\underline{Z_2}$     ORDRE VOULU ?  $\underline{Z_3}$

Alors, la commande détermine tous les chemins allant de x à y et qui sont de longueur au plus égale à k, puis elle les affiche sous la même forme arborescente que les commandes ASC, DES et FIL (cf. figure 30 ).

A toutes ces visualisations partielles est associé le MENU ⑥, qui reprend les commandes du MENU ④, lui-même associé au graphe éclaté.

Les sorties papiers sont de deux types :

\* commande COPIE : cette commande effectue une copie fidèle du dessin présent sur l'écran au moment où elle est appelée (commande TEKBen de GRIGRI) et la range dans le fichier destiné au traceur BENSON.

\* commande DESSIN : cette commande permet de tracer un dessin sur papier avec une présentation plus soignée. L'utilisateur peut choisir successivement :

- le format du dessin, soit grand format (90 cm x 65 cm) soit format A3 (42 cm x 29,7 cm).

- un titre et un sous titre, qui viendront s'inscrire dans un cartouche surmontant le graphe.

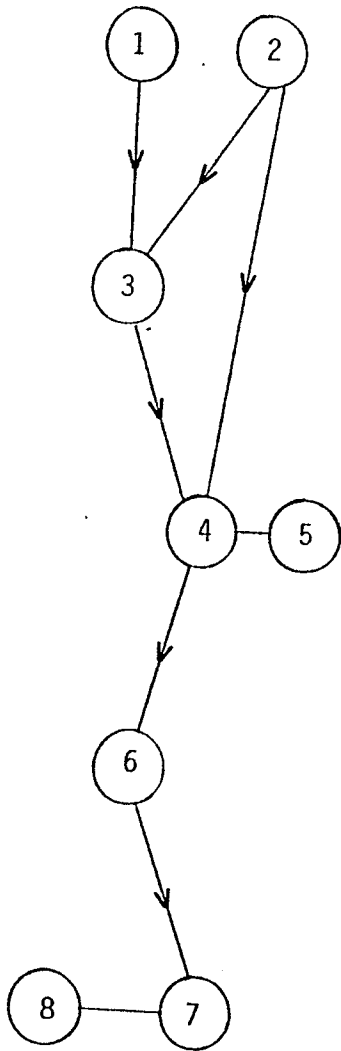
- le sens des flèches sur le graphe (de haut en bas ou le contraire)

Si les sommets possèdent des dénominations, elles seront écrites dans un cartouche "LEGENDE".

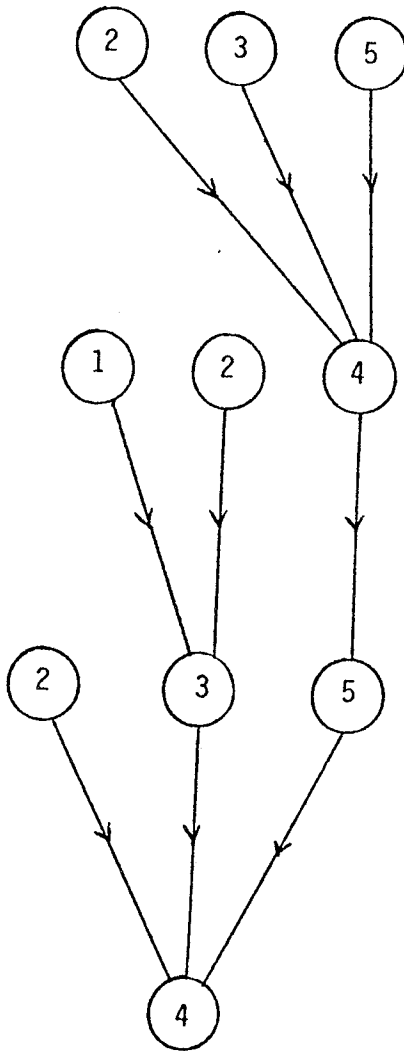
La figure 31 montre une photoréduction (1 : 2) d'un tel dessin au format A3.

La commande DESSIN n'est associée qu'aux dessins globaux (REDUIT, ECLATE, CFC).

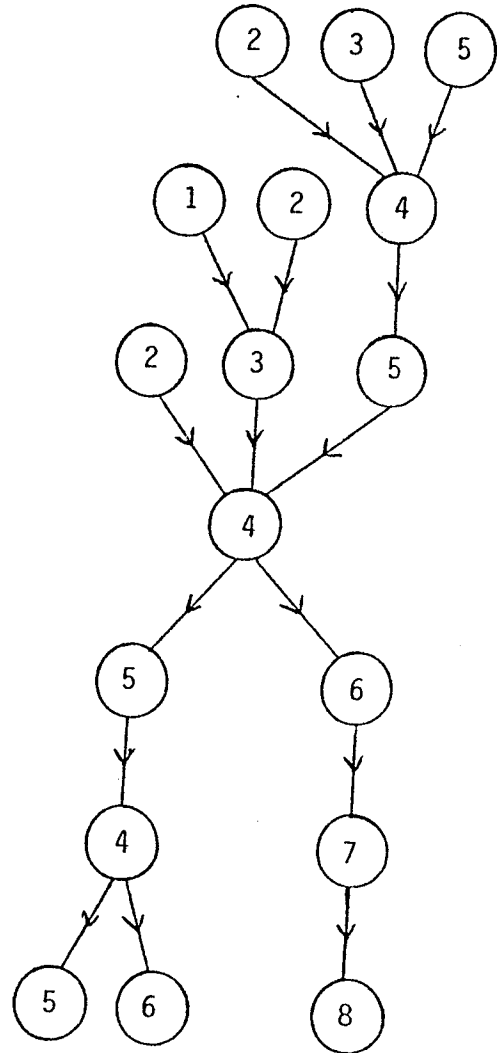
EXEMPLES



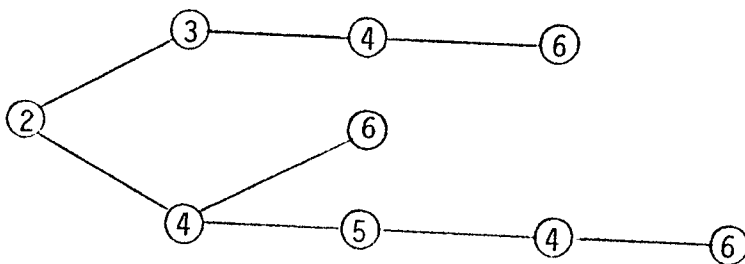
GRAPHE ECLATE



ASC (4,3)



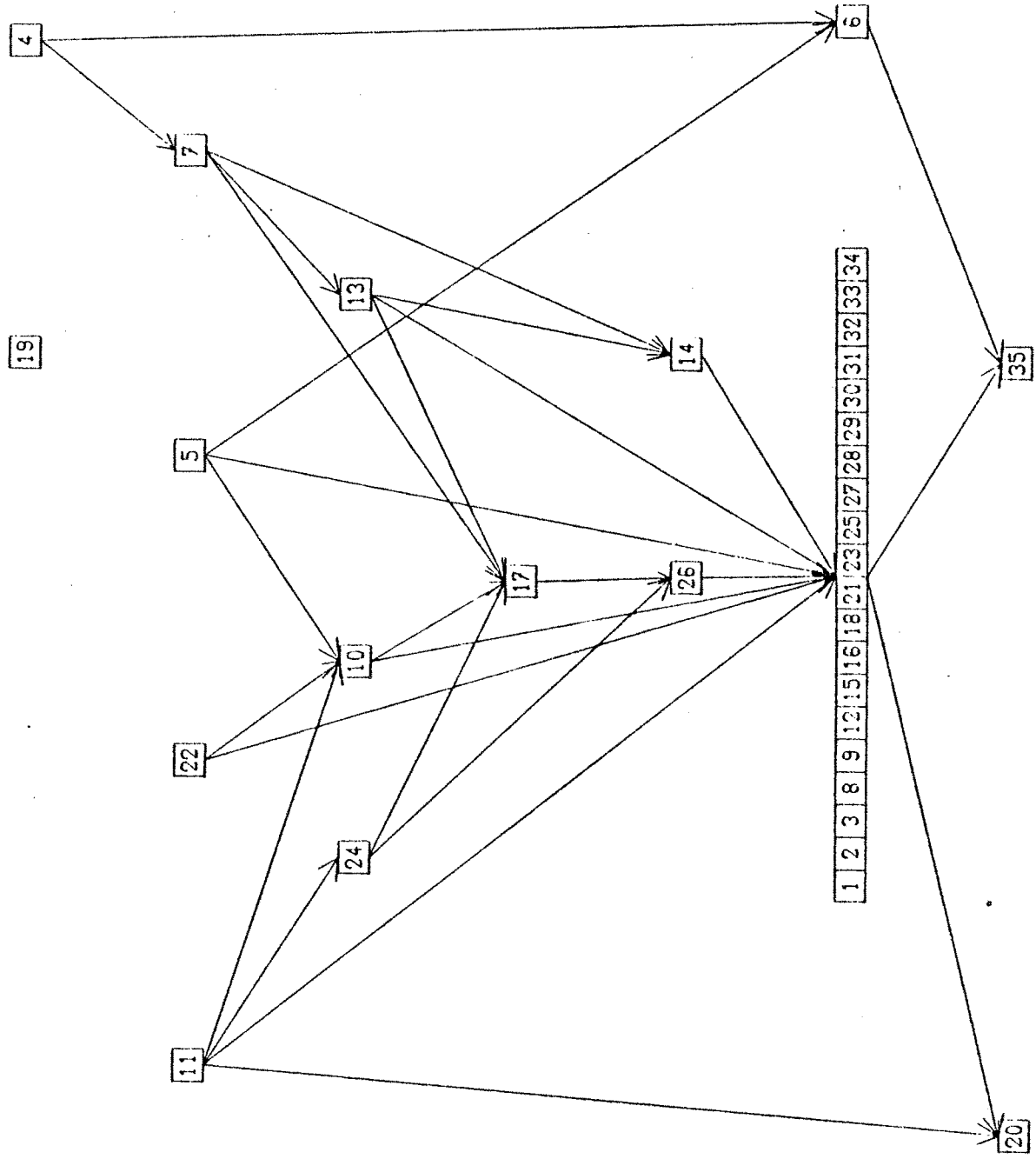
FIL (4,3)



CHEMINS (2,6,4)

# TES FRANCAIS EN 35 BRANCHES 1975

## GRAPHE REDUIT SEUIL RELATIF MIXTE 11x



- Figure 31 -

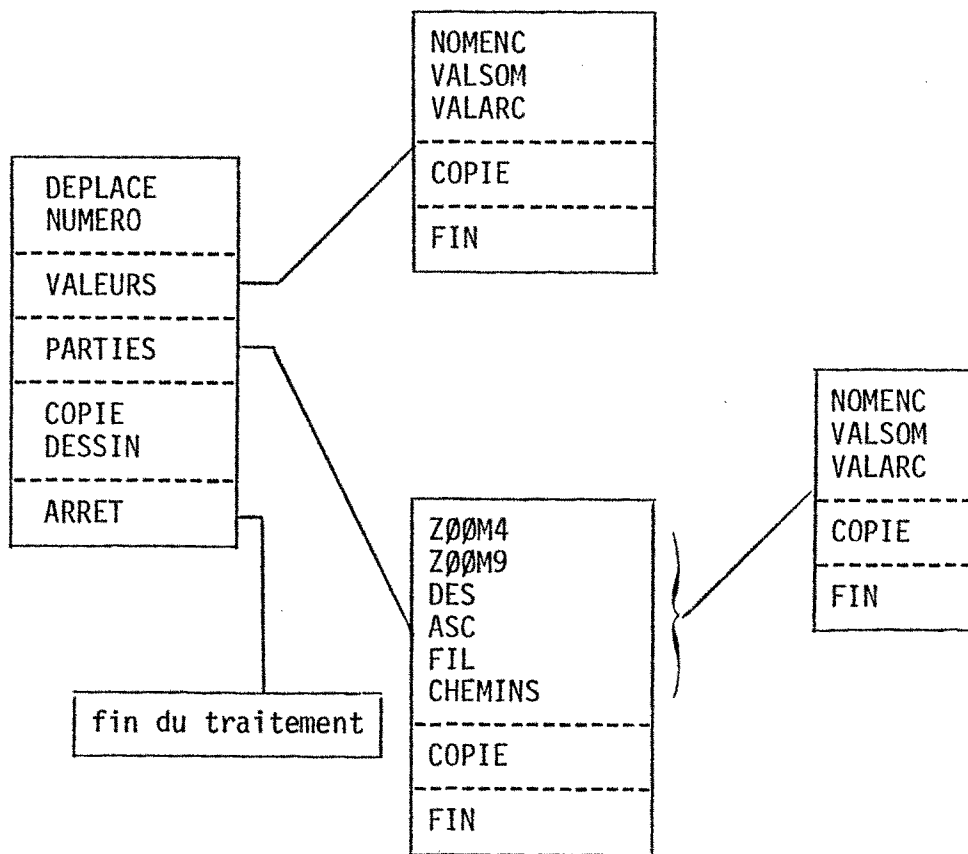
# LEGENDE

- 1 AGRI. SYLV. PECHE
- 2 VIANDE ET LAIT
- 3 ALIMENTAIRE
- 4 CHARBON ET COKE
- 5 PETROLE GAZ NAT.
- 6 EAU GAZ ELEC
- 7 SIDERURGIE
- 8 METX. NON FERREUX
- 9 MAT. CONSTRUCTION
- 10 VERRERIE
- 11 CHIMIE ET FIBRES
- 12 PHARMACIE
- 13 FONDERIE
- 14 MECANIQUE
- 15 MAT. ELEC. PROF.
- 16 EQUIP. MENAGER
- 17 CONSTR. AUTO
- 18 NAVAL, AERC. ARMES
- 19 TEXTILE HABILLEMENT
- 20 CUIR & CHAUSSURE
- 21 BOIS ET MEUBLES
- 22 PAPIER ET CARTON
- 23 EDITION
- 24 MAT. PLASTIQUES
- 25 BAT. GENIE CIVIL
- 26 PEPA. & COMM. AUTO
- 27 HOTELLERIE
- 28 TRANSPORTS
- 29 POSTE & TELECOMM
- 30 SERV. MAR. ENTREP.
- 31 SERV. MAR. PARTIC.
- 32 IMMOBILIER
- 33 ASSURANCES
- 34 SERV. FINANCIERS
- 35 SERV. NON MAR.

Le logiciel CØPERNIC est construit de la même façon : mêmes types d'entrées (plus l'appel à partir d'une c.f.c. dans LØVIGREC), même hiérarchie de menus.

Il comprend un élément de dialogue particulier : l'utilisateur doit préciser s'il veut calculer une hiérarchie centrifuge directe, inverse, ou non orientée. (cf. chapitre 2).

Son utilisation indépendamment de LØVIGREC est guidée par un programme d'initialisation générale, activé par la commande initiale au clavier : "KEPLER" (cette commande lance l'exécution de la même façon que la commande "IMAG" pour LØVIGREC). La structure du dialogue est la suivante :



-- Figure 32 --

On n'a ici que trois niveaux, le premier niveau de sélection d'une image globale étant inutile.

Les seules différences avec le schéma de dialogue de LOVIGREC sont les suppressions de ARCEPAIS et CFC.

### 2.3. Description interne

#### 2.3.1. La structure de données

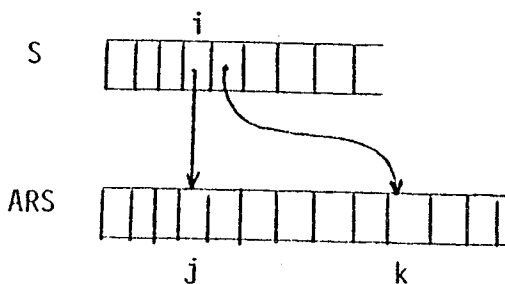
L'ensemble des variables et des tableaux utilisés pour décrire les différents types de données manipulées par les sous-programmes est réparti en 27 zones de mémoire qui sont des "commons à étiquettes" en Fortran.

Ces 27 zones sont regroupées en cinq classes, d'après les rôles joués dans les programmes par les données qu'elles reçoivent.

Parmi ces cinq classes, deux décrivent les propriétés structurelles des graphes étudiés ; la classe ST contient les données et résultats relatifs à la structure globale des graphes dans LØVIGREC. La classe C joue le même rôle dans CØPERNIC. La classe G décrit les différents types de données auxiliaires. Les deux autres classes regroupent les tableaux utilisés par GRIGRI, la classe A pour les affichages de dessin et la classe T pour les affichages de textes.

#### \* La classe ST :

La structure du graphe est décrite par quatre groupes de tableaux : ST1 : quatre tableaux d'entiers S, P, ARS, et ARP. Ces tableaux décrivent le graphe initial sous la forme suivante :





Le tableau ARS contient la liste des successeurs du sommet  $i$  de  $j = S(i)$  à  $k - 1 = S(i+1) - 1$ .

Si le sommet  $i$  n'a pas de successeur,  $S(i) = S(i+1)$ . P et ARP contiennent les mêmes données dans l'ordre des prédécesseurs. (S, ARS) permet de descendre dans le graphe et (P, ARP) de monter. Suivant le sens de parcours, on est amené à utiliser l'une ou l'autre des deux structures.

Pour un graphe ayant  $n$  sommets et  $p$  arcs, l'encombrement total est :  $2(n + p + 1)$ .

ST2 : six tableaux d'entiers RS, RP, GRS, GRP, DMS et DMP.

Les tableaux (RS, GRS) et (RP, GRP) donnent les listes d'arcs du graphe réduit avec les mêmes conventions que pour ST1. Les tableaux DMS et DMP indiquent pour chaque arc entre deux c.f.c., le degré de multiplicité qu'il a, en terme de nombre d'arcs du graphe initial.

Si le graphe initial est sans circuit, DMS et DMP valent 1 partout, et les autres tableaux sont identiques à ceux de ST1.

ST3 : trois tableaux entiers CFC, TA et NSØM.

Le tableau CFC indique à quelle c.f.c. appartient chaque sommet ;  $CFC(i) = j$  : le sommet  $i$  est dans la c.f.c.  $j$ . Les tableaux (TA, NSØM) décrivent l'ensemble des c.f.c. ; La taille de la c.f.c.  $k$  est donnée par  $TA(k+1) - TA(k)$ . Les numéros des sommets faisant partie de cette c.f.c. sont donnés dans NSØM, de NSØM(TA(k)) à NSØM(TA(k+1)-1). Cette structure permet de parcourir les sommets c.f.c. par c.f.c.

ST4 : trois tableaux d'entiers NIV, INIV et LNIV.

Ces trois tableaux décrivent la structure hiérarchique du graphe de la même manière que ST3 celle des c.f.c. :  $NIV(i) = j$  : la c.f.c.  $i$  est au niveau  $j$ . Les tableaux (INIV, LNIV) donnent la répartition des c.f.c. par niveaux, la liste des c.f.c. du niveau  $k$  étant comprise entre LNIV(INIV(k)) et LNIV(INIV(k+1)-1). Ceci facilite les parcours par niveaux du graphe.

Les niveaux sont donnés pour le graphe réduit, puisque les sommets ont les mêmes niveaux que leurs CFC. Ainsi, le niveau du sommet  $i$  est donné par :  $NIV(CFC(i))$ .

Enfin le common STO contient les 3 variables  $N$ ,  $NCF$  et  $NIVMAX$  qui sont respectivement le nombre des sommets du graphe, celui des c.f.c., et celui des niveaux.

STO contient aussi un tableau booléen à 5 cases  $ABSENT(5)$  qui permet de savoir si les données auxiliaires (épaisseurs d'arcs, valeurs brutes etc ...) existent ou non.

La classe C :

L'ensemble des informations utilisées pour le calcul de la structure hiérarchique centrifuge et la détermination des positions des sommets se répartit en 7 zones :

- C1 : le tableau d'entiers  $MC$ , à deux dimensions est la matrice des plus courts chemins.
- C2 : trois tableaux d'entiers  $IC\emptyset$ ,  $IZ\emptyset$ ,  $NSZ$  et un tableau de réels  $THI$ .  
 $IC\emptyset$  donne la structure hiérarchique ;  $IC\emptyset(i)$  contient la différence entre l'écartement du sommet  $i$  et le rayon du graphe (cf. chapitre 2) ( $IC\emptyset(i) = 0$  si  $i$  est un centre).  
 $IZ\emptyset$  donne la répartition des sommets dans les secteurs articulés sur le (ou les) centres ;  $IZ\emptyset(i) = j$  : le sommet  $i$  est dans le secteur  $j$ .  
 $NSZ$  donne le nombre de sommets par secteurs ;  $NSZ(i) = j$  : le secteur  $i$  contient  $j$  sommets.  $THI(k)$  : angle polaire où commence le secteur  $k$ .
- C3 : Deux tableaux d'entiers  $NDEG$  et  $ISHU$   
 $NDEG$  donne le degré de chaque sommet et  $ISHU$  la portée maximale des arcs qui lui sont adjacents.

(Ces deux données interviennent dans les calculs de priorités pour le placement des sommets).

- C4 : 3 tableaux d'entiers NPR, LATT et NUSP  
NPR contient l'indice de priorité des sommets  
LATT sert à enregistrer les sommets en attente de placement  
NSUP à classer les voisins du sommet en court de placement.
- C5 : 2 tableaux de réels RLEV et THETA  
RLEV donne le rayon des cercles concentriques de niveau ;  
RLEV(i) = r : tous les sommets du niveau i ont un rayon polaire égal à r.  
THETA donne les angles polaires des sommets.
- C6 : 2 tableaux de réels CFX et CFY  
Ces tableaux contiennent les coordonnées cartésiennes des sommets.

Les variables simples NT, NC, NCØ, NSC et NZ sont dans le common STO ;  
elles donnent respectivement le nombre de sommets, le numéro de c.f.c.  
(pour les appels à partir de LØVIGREC), le nombre de niveaux hiérarchiques,  
le nombre de centres et le nombre de secteurs, STO contient aussi un tableau  
d'entiers NVU qui donne les numéros des sommets dans LØVIGREC (éventuellement)

\* La classe G :

Les données auxiliaires d'affichage de LØVIGREC sont rangées  
dans 5 zones ;

- G1 : trois tableaux réels ØRDX, ØRDY et TABX, une variable RST.  
ØRDX et ØRDY donnent les coordonnées des centres des c.f.c. et  
TABX les abscisses des centres des sommets.  
RST contient le rayon standard attribué aux sommets  
RST est utilisé d'une part dans le calcul de la séparation minimale  
entre deux sommets (placement et éclatement) et, d'autre part, dans le  
calcul effectif des points d'affichage à partir des coordonnées des  
centres.

G2 : quatre tableaux d'entiers NCS, IDA, IFA, NCA.

Ces tableaux sont destinés à recevoir les informations concernant les tailles des sommets et les épaisseurs des arcs. Si ces informations existent pour le graphe traité, elles sont stockées dans un fichier auxiliaire créé au moment du seuillage (cf. seuillage).

NCS contient des entiers (compris entre 1 et 9) qui définissent le coefficient multiplicatif appliqué au rayon standard RST ; pour chaque sommet  $i$  :  $\text{rayon}(i) = \text{RST} * \text{NCS}(i)$ . Les tableaux (IDA, IFA) donnent la liste des arcs qui subsistent après le seuillage, avec les mêmes conventions que (S, ARS). Le tableau NCA, associé à IFA, indique pour chaque arc un facteur d'épaisseur (entier compris entre 1 et 5).

Dans le cas où ces données n'existent pas, les sommets du graphe ont tous la même taille ( $\text{NCS}(i)$  vaut 4 pour tout  $i$ ) et les arcs sont sans épaisseur.

G3 : deux tableaux LISNØM et ABR

Ces tableaux contiennent la nomenclature du graphe, qui peut être définie par l'utilisateur dans un deuxième fichier auxiliaire.

LISNØM contient le libellé intégral en 16 caractères, et ABR le nom abrégé (3 caractères).

G4 : deux tableaux de réels SØVAL et ARVAL

Ces tableaux contiennent les valeurs brutes associées aux sommets et aux arcs subsistant après seuillage.

Comme les données de G2, elles sont contenues dans le fichier auxiliaire créé au moment du seuillage.

G5 : deux tableaux de réels : DELAI et AMPLI

Dans la perspective d'une extension du logiciel visant à permettre de représenter sur écran graphique des phénomènes de transmission de perturbations dans un réseau, ces deux tableaux sont destinés à contenir deux paramètres associés à chaque arc :

- un délai de transmission de l'extrémité initiale à l'extrémité finale
- un facteur d'amplification des perturbations.

### 2.3.2. La structure graphique

La manipulation des données nécessaires à l'affichage proprement dit des dessins et des textes utilise les primitives de GRIGRI. Nous avons vu précédemment que le dialogue se faisait avec des MENU et des MESSAGE.

Les données à afficher sont décrites dans cinq zones de types A et quatre zones de type T.

Chaque zone contient tous les éléments de description d'une figure de points (type A) ou de textes (type T) : numéro de figure, tableau de sections etc ...

L'ensemble de la structure graphique est composé de :

- \* 3 fenêtres : deux fenêtres fixes (0.,0.,1.,1.) et (-1.,-1.,+1.,+1.) qui servent respectivement aux dessins de LØVIGREC (globaux et partiels sauf CFC et ZØØM) et au dessin de la hiérarchie centrifuge, et une fenêtre mobile (XB, YB, XH, YH) utilisée pour les effets de loupe (commande ZØØM).
- \* 5 figures de points : deux pour LØVIGREC, dessins globaux et dessins partiels ; deux pour CØPERNIC, dessin centrifuge et dessin des cercles de niveaux ; enfin, une figure auxiliaire permet la collecte des coordonnées dans les phases de modification interactive (commande DEPLACE).
- \* 4 figure de textes : deux pour les dessins globaux et deux pour les dessins partiels ; dans chaque cas, l'une sert à l'affichage des numéros et l'autre à l'affichage des noms abrégés (commande NØMENC).

Au total, l'encombrement mémoire de la structure de données est de l'ordre de 56K octets, dont la moitié pour la structure graphique stricto sensu (les 9 figures de points et de textes).

### 2.3.3. Typologie des sous-programmes

Les logiciels LØVIGREC et CØPERNIC totalisent une soixantaine de sous-programmes de tailles très diverses, de quelques instructions à une centaine.

Par rapport à leur rôle dans le programme d'ensemble, ces sous-programmes peuvent être répartis en trois catégories :

- 1 \* modules de bases
- 2 \* sous-programmes spécifiques
- 3 \* phases de traitement

Les modules de bases sont des sous-programmes d'intérêt général qui reviennent souvent dans le traitement des données ou bien qui peuvent être utilisés indépendamment des logiciels dont ils font partie.

Ils se répartissent en deux groupes :

- \* traitement mathématique
- \* traitement graphique

Dans le premier groupe on trouve tous les sous-programmes qui présentent un intérêt en eux-mêmes, indépendamment des logiciels LØVIGREC et CØPERNIC sous-programme de tri des éléments d'un tableau (qui revient très souvent), sous-programme de réduction d'un graphe, de calcul de la matrice des plus courts chemins etc ...

Le second groupe constitue une sorte de catalogue graphique dont les éléments sont des objets graphiques élémentaires manipulés fréquemment pour faire des affichages à l'écran ou bien des sorties sur papier.

Par exemple le programme de traçage d'un polygone régulier utilisé pour construire à l'écran la représentation des niveaux centrifuges, les sous-programmes de traçage d'une flèche, d'un hexagone, ou d'un "domino" (cfc) etc ...

La deuxième catégorie, celle des sous-programmes spécifiques, contient deux types de programmes :

- d'une part les sous-programmes qui sont spécifiques parce qu'ils effectuent une partie de traitement très liée aux logiciels, par exemple le sous-programme qui calcule les pondérations de priorités pour le rangement des sommets dans CØPERNIC, ou le programme d'initialisation itérative sur les deux premiers niveaux, ou le sous-programme qui construit la structure du multigraphe réduit (cf. structure de données).

- d'autre part des sous-programmes qui sont une combinaison spécifique de modules de bases, par exemple, les sous-programmes d'affichages partiels, ou les sous-programmes de calcul des coordonnées (placement des sommets par l'heuristique des degrés relatifs ou bien sous-programme d'éclatement des c.f.c.).

Enfin, la troisième catégorie correspond au niveau le plus global, et donc le plus spécifiquement lié aux logiciels LØVIGREC et CØPERNIC.

Cette catégorie comprend les programmes principaux, ainsi que les sous-programmes qui gèrent une phase entière de traitement : sous-programme de réduction-hiérarchisation-raccourcissement des arcs, sous-programme prenant en charge tout un niveau de dialogue (menu - test- affichage - retour au niveau antérieur) etc ...

Cette décomposition obéit à deux principes fonctionnels, une progression du général au particulier qui va des modules de bases (suffisamment élémentaires pour être utilisables ailleurs) jusqu'au programme principal, et une séparation systématique à chaque niveau entre calcul mathématique et traitement graphique : par exemple, les phases de calcul heuristique de coordonnées font partie du secteur "calcul mathématique", en revanche la génération des tableaux de coordonnées destinés à GRIGRI font partie du traitement graphique. Le cloisonnement systématique permet d'envisager le fonctionnement des phases de calcul en l'absence d'un terminal graphique, et augmente le degré de modularité des logiciels LØVIGREC et CØPERNIC.

## CHAPITRE 6

### EXPERIMENTATION ET CONCLUSIONS CRITIQUES

Nous allons présenter deux exemples d'utilisation de LØVIGREC et de CØPERNIC, illustrés par des dessins de graphes hiérarchisés. Les commentaires qui accompagnent ces deux exemples ont pour but de mettre en évidence les informations qu'il est possible de retirer directement d'une étude attentive de quelques dessins, sans préjuger de leur validité, qui pourrait être remise en cause dans le cadre d'une étude détaillée effectuée par un spécialiste.

Ensuite, nous essaierons de faire un bilan critique d'ensemble et de dégager des axes de recherche pouvant prolonger ce travail d'un point de vue pratique ou théorique.

#### 1. UTILISATION DES LOGICIELS

##### 1.1. Exemples d'utilisation de LØVIGREC

Matrice d'imports-exports de véhicules à moteur pour l'année 1973, entre 26 pays (données fournies par le GAPSET).



USA	CAN	RFA	BLX	FRA	G.B	N.L	ITA	SUI	SUE	AUT	AUS	DAN	MEX	FIN	NOR	N-Z	NGA	VEN	YOU	SIN	BRE	ESP	JAP	ARG	AFS
4.654.940	2.216.553	298.565	298.565	86.828	364.915	4.205	178.657	-	235.681	4.573	1.601	1.431	68.541	70	139	-	-	-	191	0	6.590	20.000	2.304.652	1.742	-
4.130.771	-	135.533	323	15.295	58.145	453	18.036	-	38.747	1.422	172	286	560	724	-	-	-	-	-	-	152	600	221.455	-	-
75.279	7.676	-	620.016	783.438	132.250	95.277	379.307	-	21.884	18.373	1.190	2.003	10.790	1.035	1.493	-	-	-	11.056	-	2.490	20.550	75.039	35	-
109.384	1.467	1.276.910	-	487.523	232.746	97.684	85.300	-	74.920	150	300	910	36	257	1.420	-	-	-	450	-	-	10.334	87.670	13	-
59.904	2.800	877.600	360.860	-	93.555	71.000	322.743	-	11.690	570	355	580	368	360	-	-	-	-	10.525	-	18	74.600	53.563	246	-
71.784	3.207	373.925	119.741	311.644	-	72.201	133.471	-	185.290	6.875	4.375	4.817	51	350	229	-	-	-	127	-	167	7.797	176.624	414	1.214
19.410	1.034	442.535	387.940	246.480	71.345	-	70.894	-	43.750	5.310	40	1.190	-	200	12	-	-	-	1.396	-	47	26.215	74.490	18	-
12.245	819	498.340	161.790	490.000	71.490	7.290	-	-	2.105	5.523	185	190	17	-	-	-	-	-	2.467	-	17	3.083	18.826	122	-
19.914	613	369.884	50.017	118.681	41.342	4.480	99.036	-	67.885	35.795	77	766	10	640	-	-	-	-	228	-	-	6.351	39.681	-	-
21.650	7.935	336.550	36.203	50.213	135.658	7.363	24.047	-	-	4.053	40	20.430	-	39.262	28.253	-	-	-	76	-	46	1.143	15.767	-	-
3.053	580	367.015	44.926	79.270	35.085	1.151	44.142	-	33.605	-	-	942	56	70	28	-	-	-	307	-	-	8.420	20.018	-	-
87.777	13.355	72.075	241	17.992	101.876	307	17.027	-	28.193	165	-	63	18	-	-	-	-	-	-	-	36	1.051	266.651	-	-
9.884	335	191.632	37.201	39.219	64.560	3.770	17.223	-	105.118	8.032	137	-	-	209	976	-	-	-	915	-	13	4.465	39.795	-	-
293.244	27.252	86.389	1.142	13.121	6.400	-	492	-	34	-	22	-	-	-	-	-	-	-	-	-	1.790	3.663	21.764	103	-
7.134	627	78.516	22.580	13.815	30.075	3.159	8.246	-	104.415	360	47	1.186	-	-	430	-	-	-	770	-	-	10.497	38.740	-	-
5.566	3.878	132.231	26.311	24.507	19.497	3.883	3.453	-	74.780	675	-	2.908	-	667	-	-	-	-	-	-	-	1.560	31.785	-	-
6.028	2.166	7.560	-	2.986	87.780	39	5.876	-	985	-	121.940	14	-	-	-	-	-	-	-	-	-	44	49.090	-	-
2.995	290	59.357	3.524	28.972	48.460	1.194	8.926	-	1.890	5.766	161	378	42	-	10	-	-	-	-	-	61	945	35.533	61	-
130.695	44.040	34.915	24	4.640	11.857	12	14.154	-	94	23	68	-	2.355	-	-	-	-	-	-	-	4.854	531	17.655	2.785	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
49.531	11.290	248.590	316	50.278	148.721	3.099	23.975	-	8.795	16	63.919	33	668	40	-	-	-	-	-	-	684	333	132.360	117	-

Matrice initiale des imports-exports de véhicules à moteurs (en milliers de dollars U.S.)  
 Source: Yearbook International Trade Statistics (U.N. 1974)

Dans le tableau initial (voir Figure 33), les coefficients  $M_{ij}$  représentent les dépenses des pays  $i$  pour acheter des véhicules aux pays  $j$ . Les arcs  $(i,j)$  représentent donc des flux monétaires, cumulés sur un an, et évidemment orientés en sens inverse des flux de biens matériels correspondants.

Pour faire ressortir les relations privilégiées dans le graphe, nous avons construit des matrices multiseuillées par seuillage relatif mixte (cf. chapitre 4) avec différents niveaux de seuils, et nous avons fait fonctionner LØVIGREC avec ces différentes matrices :

Seuil	0%	0.1%	0.5%	1%	2.5%	5%	7.5%	10%	15%	17.5%	20%	22.5%	25%	30%	35%
Nombre de cfc	6	10	11	11	11	11	11	11	13	15	17	18	19	22	24
Nbre de niveaux	2	5	3	3	3	3	3	3	5	5	6	6	5	5	4

Le critère retenu pour apprécier l'intérêt de chaque seuil est essentiellement le nombre de niveaux de la hiérarchie des rangs appliquée au graphe correspondant.

Ici, les différents seuillages effectués montrent deux plages bien distinctes :

- une zone de forte stabilité jusqu'à 10% : 11 c.f.c. sur 3 niveaux.
- une zone correspondant à un nouveau découpage hiérarchique : 5 à 6 niveaux, avec une croissance régulière du nombre des c.f.c.

D'où une sélection d'images représentatives de ce premier résultat : celles qui correspondent aux seuils 10%, 20% et 35%.

\* 10% pour voir la structure du graphe dans la zone de stabilité sous sa forme la plus dépouillée (celle qui comporte le moins d'arcs).

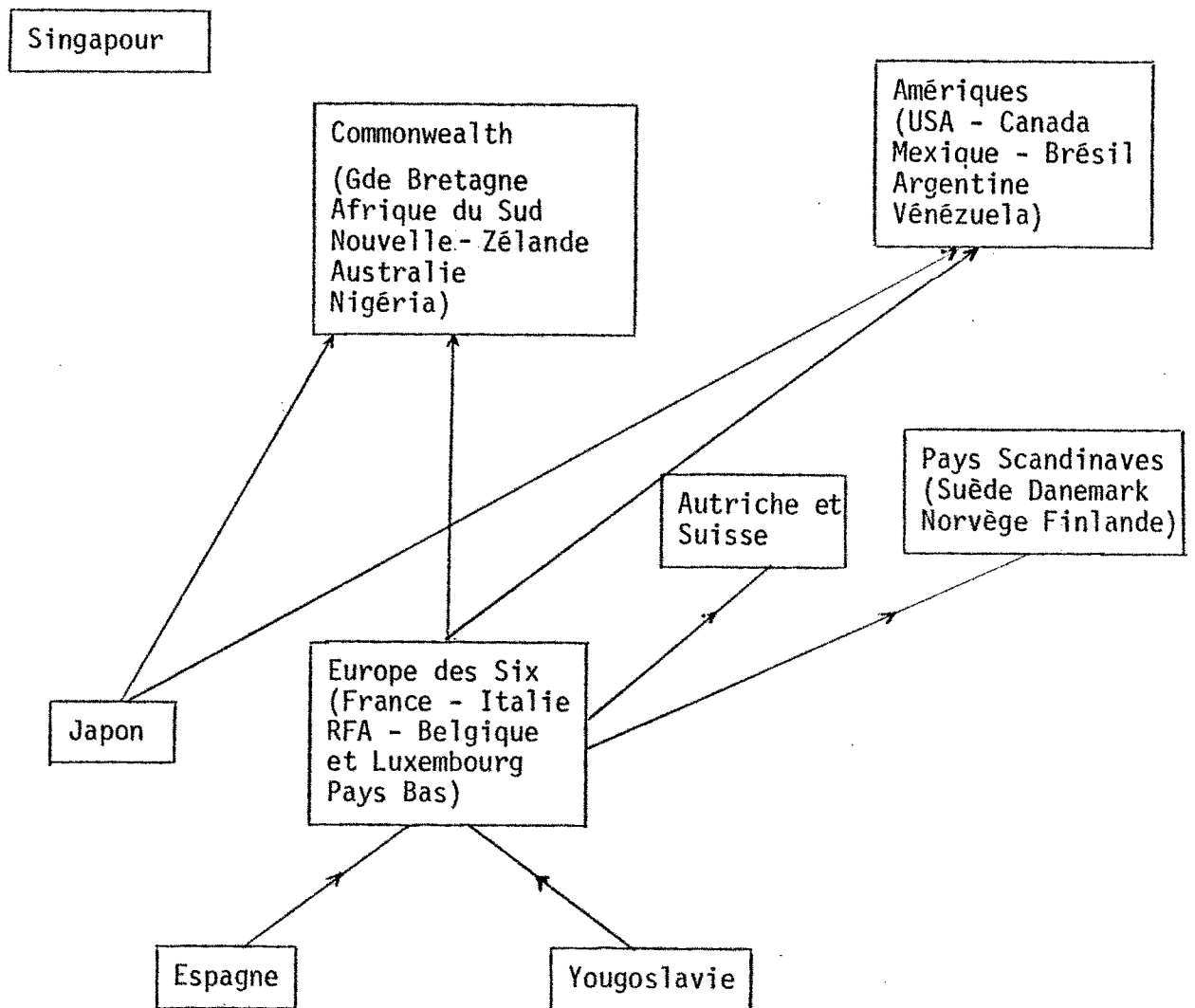
\* 20% pour avoir une image avec le maximum de niveaux

\* 35% qui correspond à la *structure limite* du graphe (presque toutes les c.f.c. sont réduites à un seul sommet) avant sa désagrégation complète.

Au seuil 10%, la structure est très pauvre :

le sommet 22 est isolé, et l'on a une grosse c.f.c. qui contient tous les sommets de 1 à 8 et de 10 à 17, soit tous les pays européens de la liste sauf la Suisse, l'Espagne et la Yougoslavie et avec les U.S.A., le Canada et l'Afrique du Sud.

La structure qui nous semble la plus intéressante est obtenue au seuil 20% ; on peut la schématiser ainsi :



- Figure 34 -

Le découpage des c.f.c. et des filières hiérarchiques suit donc remarquablement les grandes zones géographiques ou économiques, ce qui traduit bien le fait que c'est avec ses voisins ou ses dépendances que chaque pays a le plus d'échanges, les trois principaux groupes de pays repérables sur ce dessin admettant chacun un "pôle principal" USA pour les Amériques, RFA pour l'Europe des Six et, à un degré moindre, GB pour le Commonwealth.

La relation hiérarchique dans le graphe n'est pas interprétable comme une relation de domination pure et simple entre les pays des différents niveaux, d'une part parce que le seuillage est relatif et ne fait donc pas intervenir le volume réel des échanges, et d'autre part, parce que le seuillage est mixte et qu'il préserve donc toutes les relations de forte dépendance entre les extrémités des arcs.

Le même arc peut donc être interprété comme une dépendance à l'importation, à l'exportation, ou encore une dépendance mutuelle. Seule l'analyse concrète de la *forme* des échanges étudiés permettra de trancher entre les différentes hypothèses possible lorsque le niveau absolu des relations ne sera pas décisif.

Au seuil 35%, presque toutes les c.f.c. ont éclaté, seul reste le groupe U.S.A. - Canada - Mexique.

Les clivages entre filières du graphe 20% sont transformés en un découpage en six composantes connexes :

Le Nigéria, Singapour et les Pays Bas sont isolés.

Les autres composantes sont :

le tandem Afrique du Sud - Grande Bretagne

un groupe Japon + Australie + Nouvelle Zélande et Amériques ;

le reste des pays d'Europe forme un ensemble qui garde la trace des clivages précédents (Pays Scandinaves - Suisse et Autriche) et où le rôle prépondérant de la R.F.A. apparaît nettement : il lui reste 9 arcs à ce seuil très élevé.

Le rôle déterminant de certains pays dans le sous-groupe dont ils font partie peut également être mis en évidence par la représentation centrifuge des c.f.c. ; nous avons donc joint aux dessins obtenus par LOVIGREC trois dessins représentant pour les seuils 10%, 15% et 20% la c.f.c. de taille maximum au niveau de seuil considéré. Ces c.f.c. ont été hiérarchisées en hiérarchie inverse (cf. chapitre 2) de façon à ce que les sommets centraux soient les pays les plus exportateurs.

L'observation de ces dessins recoupe bien les remarques qui viennent d'être faites :

\* Au seuil 10%, la grosse c.f.c. du niveau médian comporte 16 sommets, et constitue un sous-graphe assez riche en arcs. Cependant, un certain nombre de caractéristiques se dégagent de l'ensemble. D'abord le rôle important joué par la RFA, la Grande Bretagne et les USA - Ensuite le rejet à la périphérie (pays plus importateurs qu'exportateurs) des pays Scandinaves et de l'Australie et de l'Afrique du Sud.

La présence au centre de pays comme le Mexique ou la Belgique, s'explique par le fait qu'ils sont reliés simultanément à au moins deux des trois pôles cités (RFA - GB - USA).

\* Au seuil 15%, la c.f.c. principale contient 7 sommets qui sont trois pays du Commonwealth (Grande-Bretagne, Afrique du Sud et Australie) et les quatre pays Scandinaves (Suède, Danemark, Finlande et Norvège). La Grande Bretagne au centre, avec d'un côté ses partenaires du Commonwealth et de l'autre le groupe Scandinave où la Suède joue un rôle d'articulation important.

\* Au seuil 20%, la c.f.c. principale regroupe 5 sommets, qui sont les pays de l'Europe des Six.

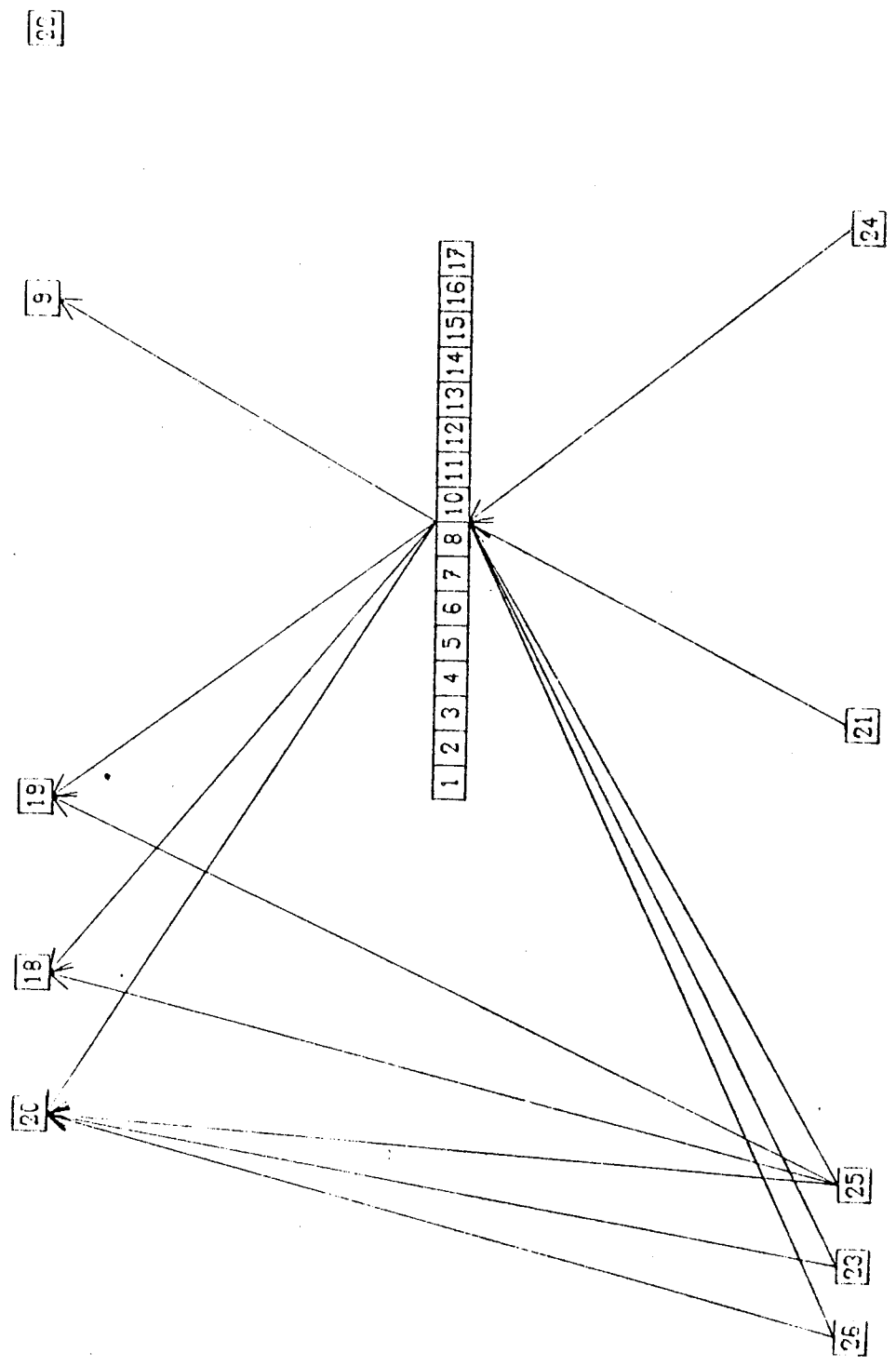
Le rôle majeur de la RFA apparaît nettement : elle sert de trait d'union entre la France et l'Italie d'une part, et le Bénélux d'autre part.

Sur les dessins qui suivent, les sommets ont une taille proportionnelle au montant total des exportations des pays correspondants, et les arcs une épaisseur proportionnelle aux flux réels entre pays.

IMPORT EXPORT VEHICULES A MOTEUR 1973  
 GRAPHE REDUIT SEUIL RELATIF MIXTE 5\*

LEGENDE

- 1 ETATS-UNIS
- 2 CANADA
- 3 ALLEMAGNE OUEST
- 4 BELGIQUE ET LUX
- 5 FRANCE
- 6 GRANDE-BRETAGNE
- 7 PAYS-BAS
- 8 ITALIE
- 9 SUISSE
- 10 AFRIQUE DU SUD
- 11 SUEDE
- 12 AUTRICHE
- 13 AUSTRALIE
- 14 DANEMARK
- 15 MEXIQUE
- 16 FINLANDE
- 17 NORVEGE
- 18 NOUVELLE-ZELANDE
- 19 NIGERIA
- 20 VENEZUELA
- 21 YOUGOSLAVIE
- 22 SINGAPOUR
- 23 BRESIL
- 24 ESPAGNE
- 25 JAPON
- 26 ARGENTINE

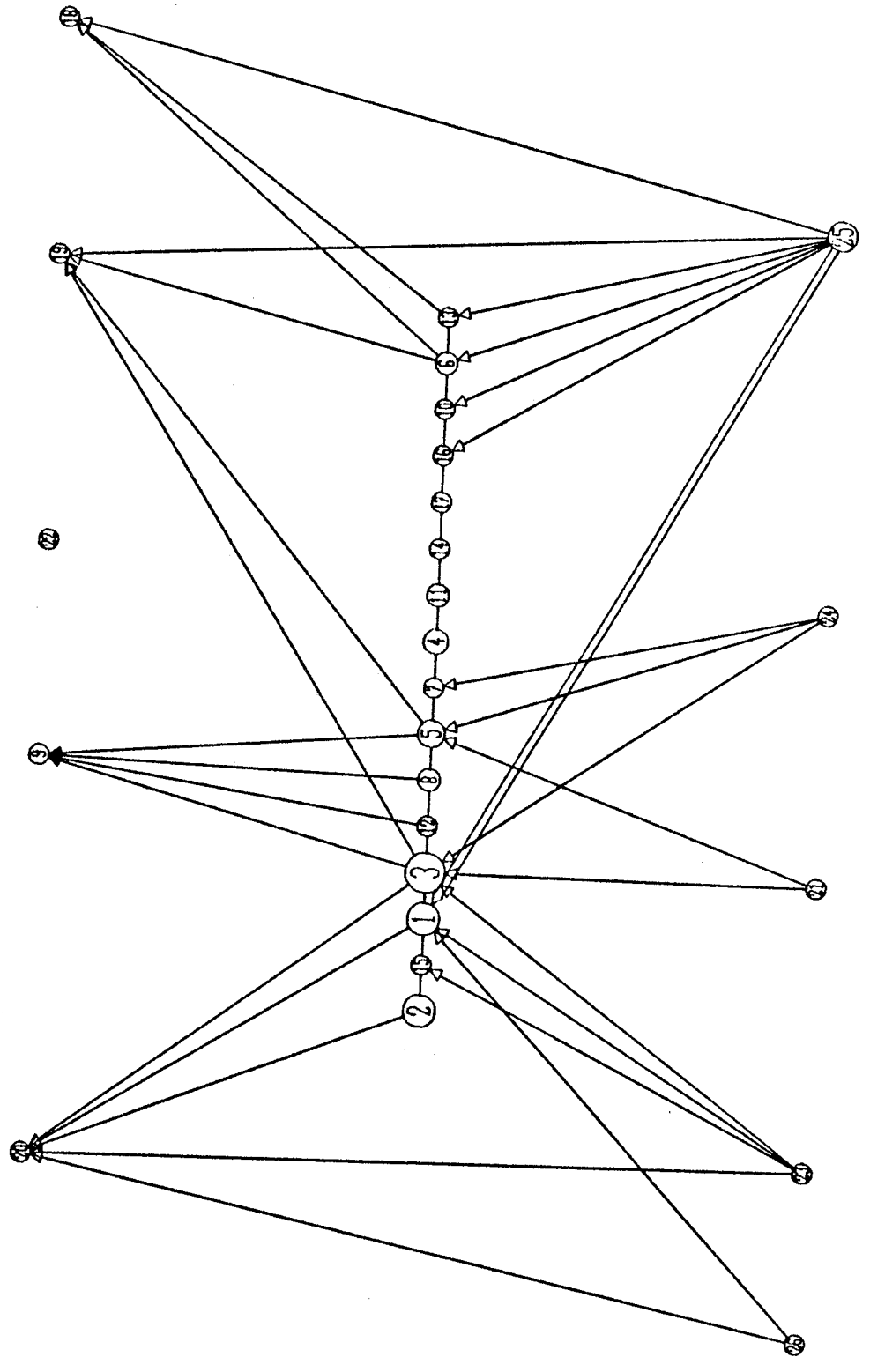


IMPORT EXPORT VEHICULES A MOTEUR 1973  
 GRAPHE ECLATE SEUIL RELATIF MIXTE 10%

LEGENDE

- 1 ETATS-UNIS
- 2 CANADA
- 3 ALLEMAGNE OUEST
- 4 BELGIQUE ET LUX
- 5 FRANCE
- 6 GRANDE-BRETAGNE
- 7 PAYS-BAS
- 8 ITALIE
- 9 SUISSE
- 10 AFRIQUE DU SUD
- 11 SUEDE
- 12 AUTRICHE
- 13 AUSTRALIE
- 14 DANEMARK
- 15 MEXIQUE
- 16 FINLANDE
- 17 NORVEGE
- 18 NOUVELLE-ZELANDE
- 19 NIGERIA
- 20 VENEZUELA
- 21 YOUGOSLAVIE
- 22 SINGAPOUR
- 23 BRESIL
- 24 ESPAGNE
- 25 JAPON
- 26 ARGENTINE

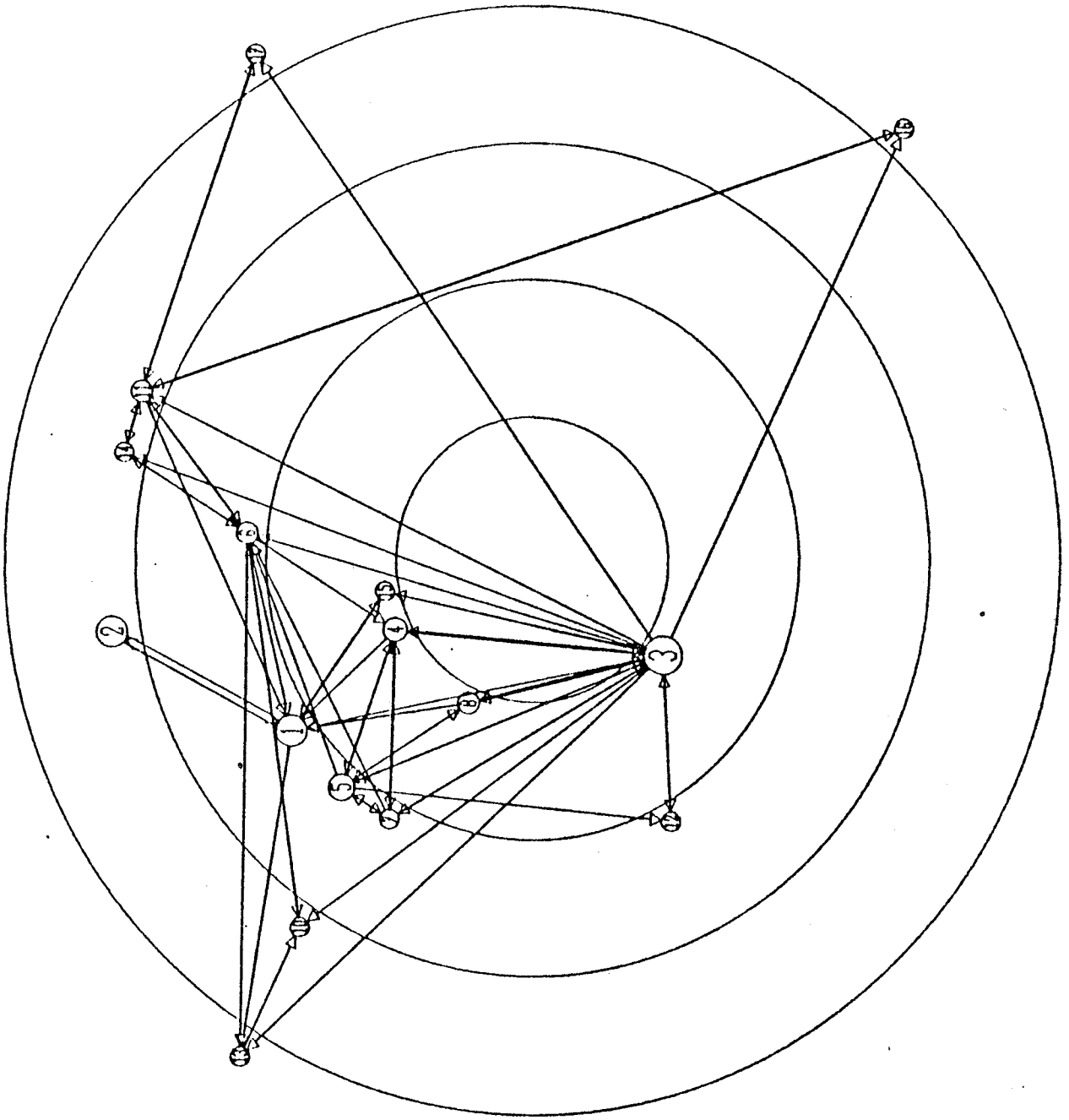
138



EXPORT IMPORT VEHICULES A MOTEUR 1973  
 CFC PRINCIPALE SEUIL 10%

LEGENDE

- 1 CANADA
- 2 ALLEMAGNE OCCIDENTALE
- 3 BELGIQUE ET LUX
- 4 FRANCE
- 5 GRANDE-BRETAGNE
- 6 PAYS-BAS
- 7 ITALIE
- 8 SUISSE
- 9 AFRIQUE DU SUD
- 10 SUÈDE
- 11 AUTRICHE
- 12 AUSTRALIE
- 13 DANEMARK
- 14 MEXIQUE
- 15 FINLANDE
- 16 NORVEGE
- 17 NOUVELLE-ZELANDE
- 18 NIGERIA
- 19 VENEZUELA
- 20 YOUGOSLAVIE
- 21 SINGAPOUR
- 22 BRÉSIL
- 23 ESPAGNE
- 24 JAPON
- 25 ARGENTINE

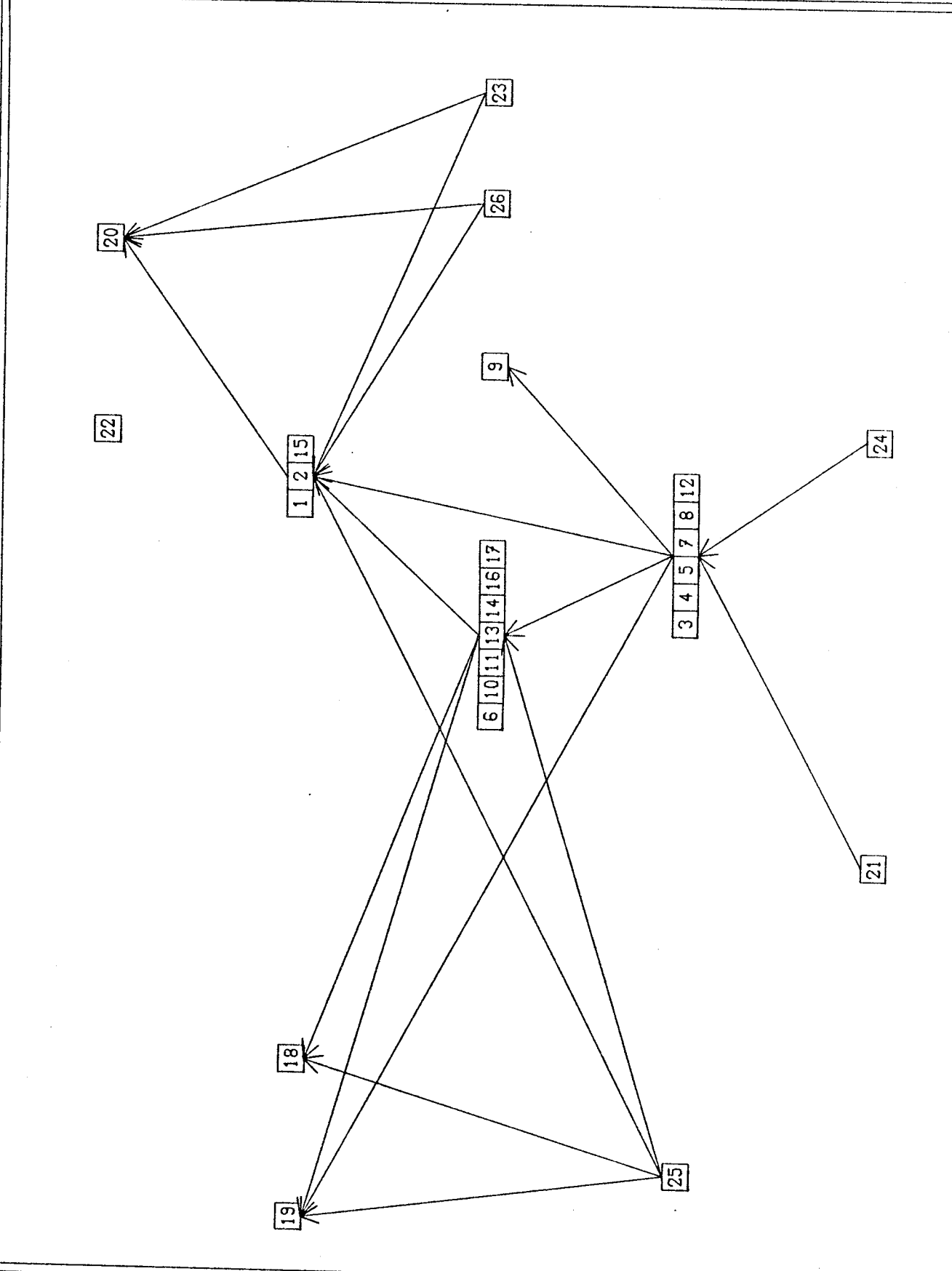




IMPORT EXPORT VEHICULES A MOTEUR 1973  
 GRAPHE REDUIT SEUIL RELATIF MIXTE 15\*

LEGENDE

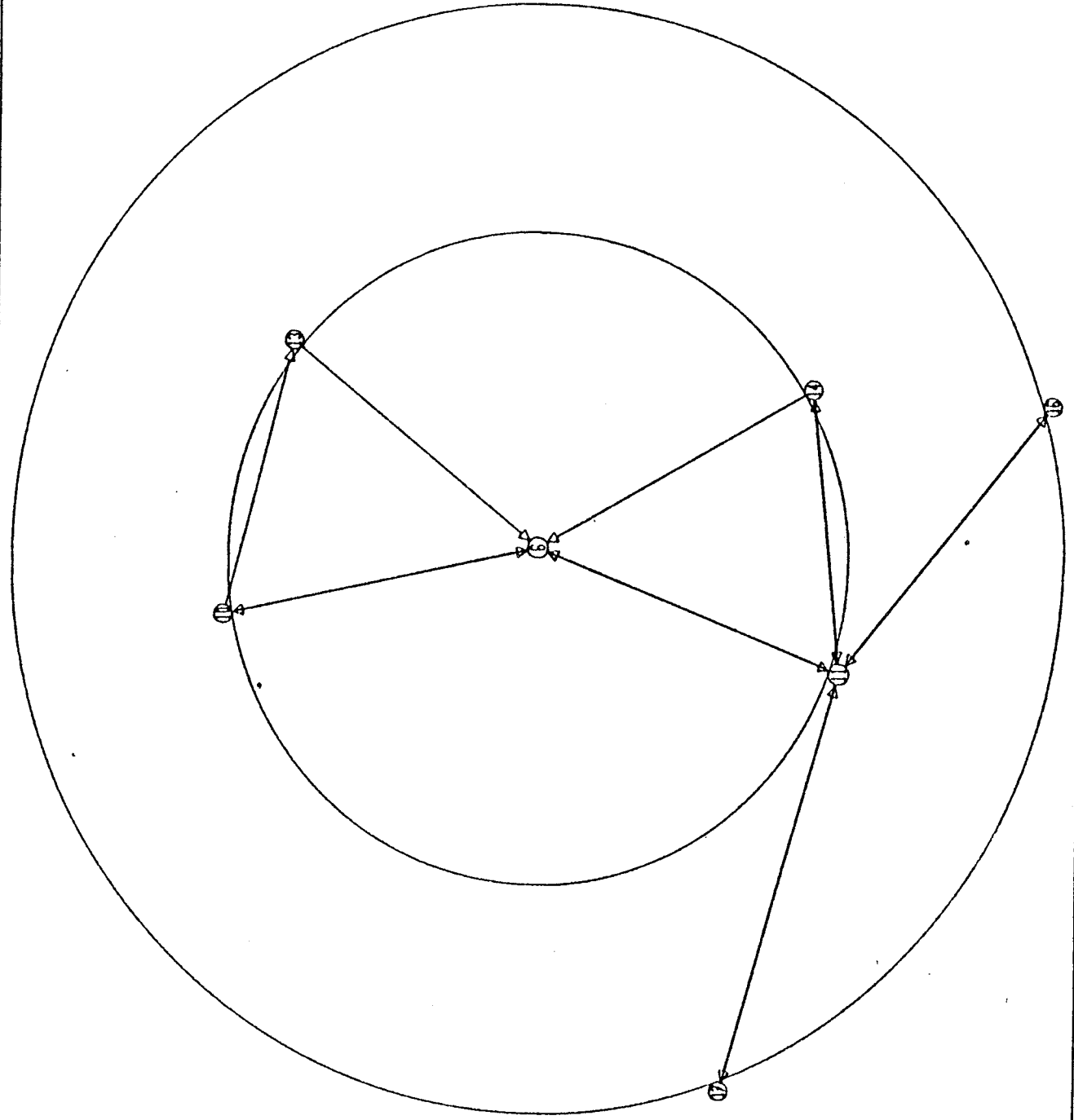
- 1 ETATS-UNIS
- 2 CANADA
- 3 ALLEMAGNE OUEST
- 4 BELGIQUE ET LUX
- 5 FRANCE
- 6 GRANDE-BRETAGNE
- 7 PAYS-BAS
- 8 ITALIE
- 9 SUISSE
- 10 AFRIQUE DU SUD
- 11 SUEDE
- 12 AUTRICHE
- 13 AUSTRALIE
- 14 DANEMARK
- 15 MEXIQUE
- 16 FINLANDE
- 17 NORVEGE
- 18 NOUVELLE-ZELANDE
- 19 NIGERIA
- 20 VENEZUELA
- 21 YUGOSLAVIE
- 22 SINGAPOUR
- 23 BRESIL
- 24 ESPAGNE
- 25 JAPON
- 26 ARGENTINE



IMPORT EXPORT VEHICULES A MOTEUR 1975  
CFC PRINCIPALE SEUIL 15x

LEGENDE

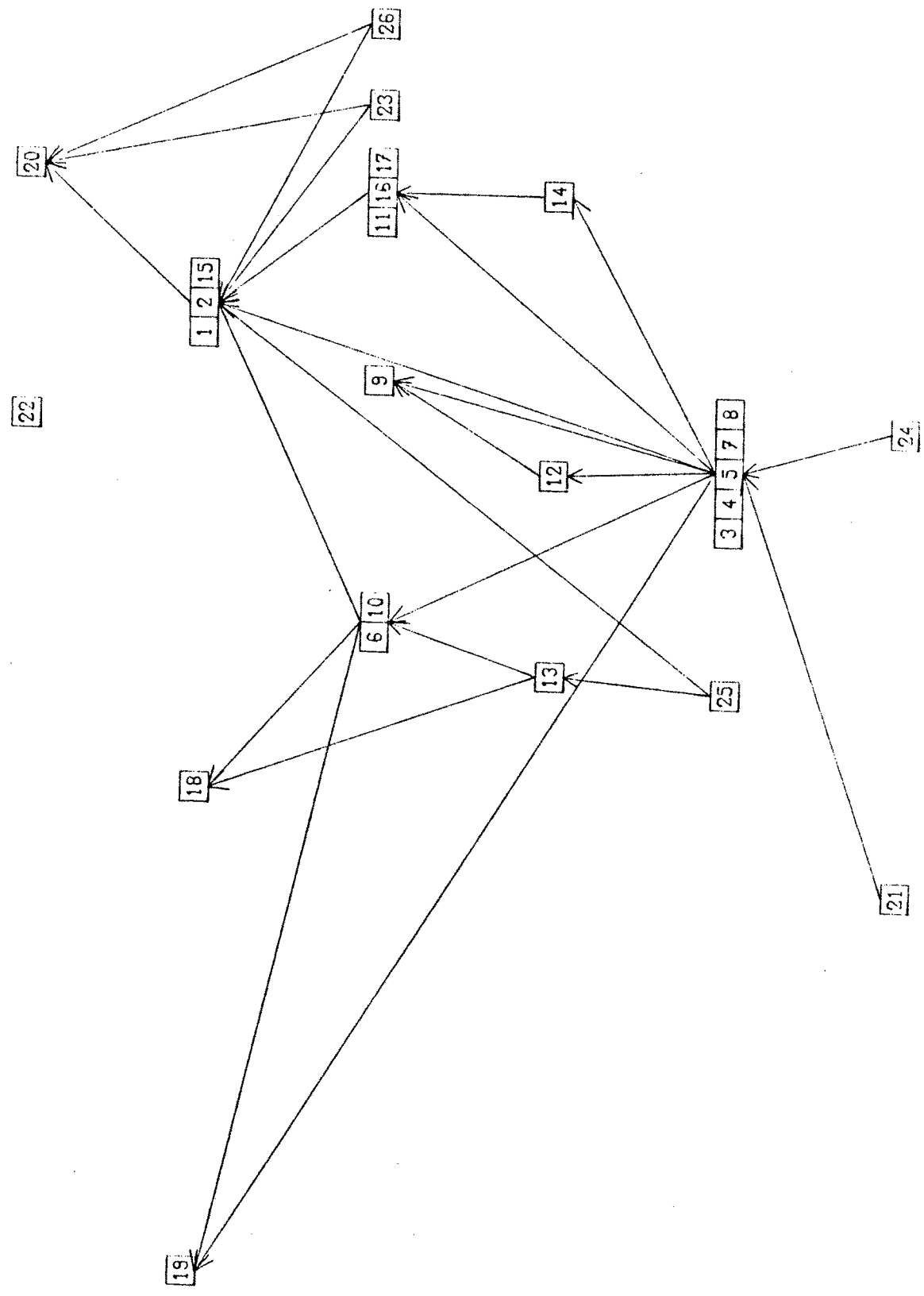
- 1 ETATS-UNIS
- 2 CANADA
- 3 ALLEMAGNE OUEST
- 4 BELGIQUE ET LUX
- 5 FRANCE
- 6 GRANDE-BRETAGNE
- 7 PAYS-BAS
- 8 ITALIE
- 9 SUISSE
- 10 AFRIQUE DU SUD
- 11 SUEDE
- 12 AUTRICHE
- 13 AUSTRALIE
- 14 DANEMARK
- 15 MEXIQUE
- 16 FINLANDE
- 17 NORVEGE
- 18 NOUVELLE-ZELANDE
- 19 NIGERIA
- 20 VENEZUELA
- 21 YUGOSLAVIE
- 22 SINGAPOUR
- 23 BRESIL
- 24 ESPAGNE
- 25 JAPON
- 26 ARGENTINE



IMPORT EXPORT VEHICULES A MOTEUR 1973  
 GRAPHE REDUIT SEUIL RELATIF MIXTE 20%

LEGENDE

- 1 ETATS-UNIS
- 2 CANADA
- 3 ALLEMAGNE OUEST
- 4 BELGIQUE ET LUX
- 5 FRANCE
- 6 GRANDE-BRETAGNE
- 7 PAYS-BAS
- 8 ITALIE
- 9 SUISSE
- 10 AFRIQUE DU SUD
- 11 SUEDE
- 12 AUTRICHE
- 13 AUSTRALIE
- 14 DANEMARK
- 15 MEXIQUE
- 16 FINLANDE
- 17 NORVEGE
- 18 NOUVELLE-ZELANDE
- 19 NIGERIA
- 20 VENEZUELA
- 21 YUGOSLAVIE
- 22 SINGAPOUR
- 23 BRESIL
- 24 ESPAGNE
- 25 JAPON
- 26 ARGENTINE

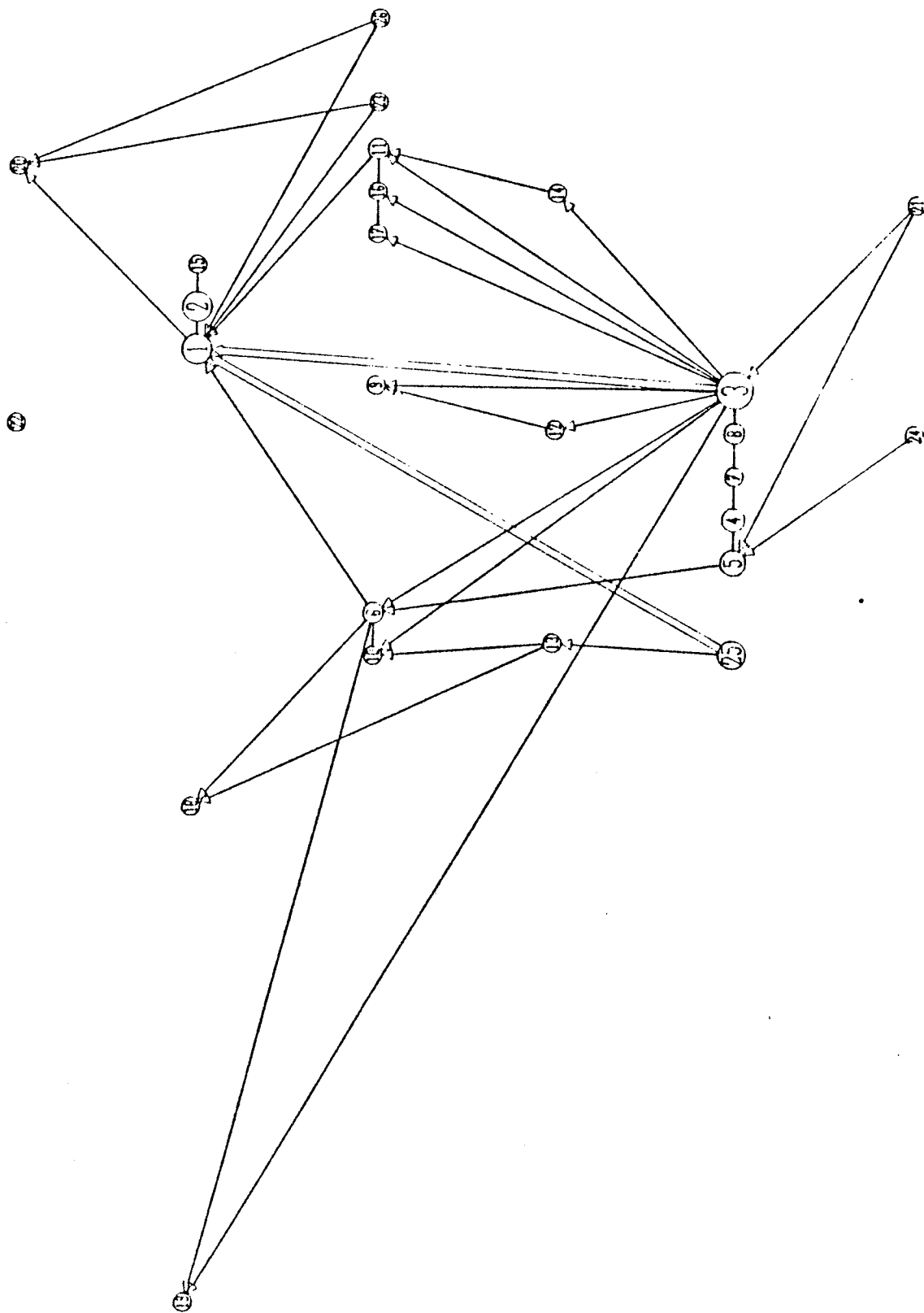


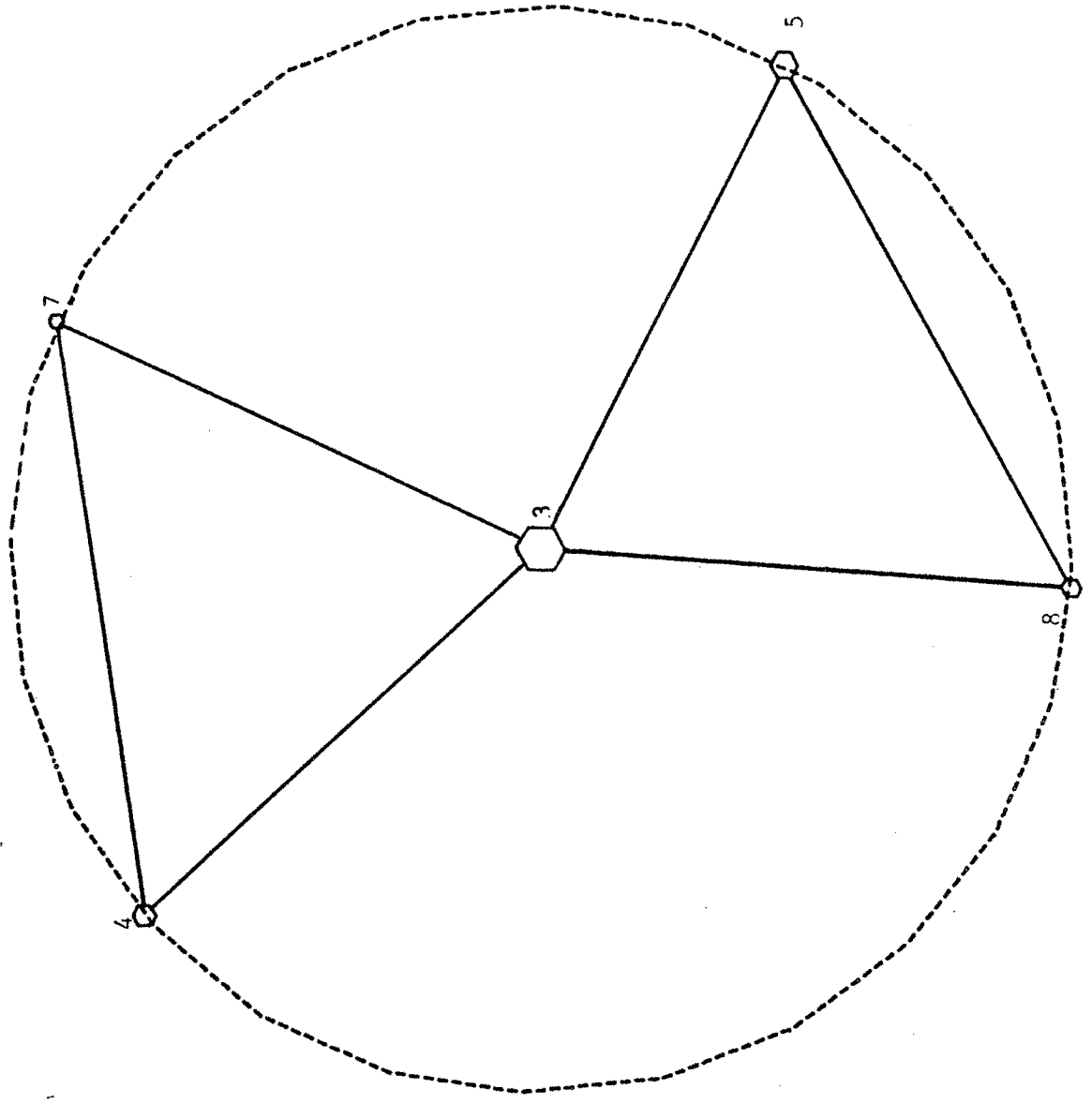
# IMPORT EXPORT VEHICULES A MOTEUR 1973

GRAPHE ECLATE SEUIL RELATIF MIXTE 20x

## LEGENDE

- 1 ETATS-UNIS
- 2 CANADA
- 3 ALLEMAGNE OUEST
- 4 BELGIQUE ET LUX
- 5 FRANCE
- 6 GRANDE-BRETAGNE
- 7 PAYS-BAS
- 8 ITALIE
- 9 SUISSE
- 10 AFRIQUE DU SUD
- 11 SUEDE
- 12 AUTRICHE
- 13 AUSTRALIE
- 14 DANEMARK
- 15 MEXIQUE
- 16 FINLANDE
- 17 NORVEGE
- 18 NOUYELLE-ZELANDE
- 19 NIGERIA
- 20 VENEZUELA
- 21 YOUGOSLAVIE
- 22 SINGAPOUR
- 23 BRESIL
- 24 ESPAGNE
- 25 JAPON
- 26 ARGENTINE

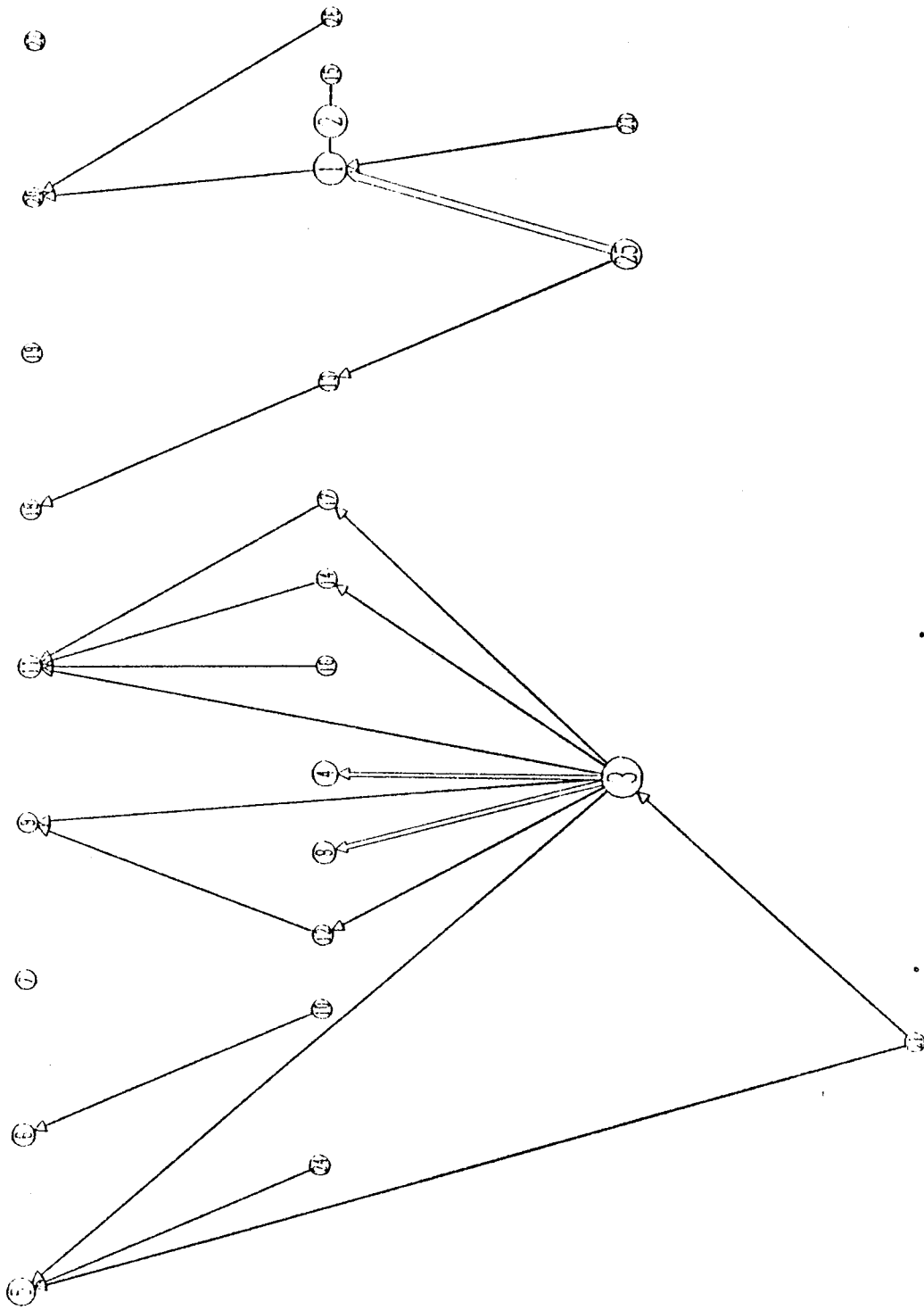




CFC PRINCIPALE SEUIL 20%

INDICATEUR DE CONTRAINDRE

IMPORT EXPORT VEHICULES A MOTEUR 1973  
 GRAPHE ECLATE SEUIL RELATIF MIXTE 35%



LEGENDE

- 1 ETATS-UNIS
- 2 CANADA
- 3 ALLEMAGNE OUEST
- 4 BELGIQUE E-LUX
- 5 FRANCE
- 6 GRANDE-BRETAGNE
- 7 PAYS-BAS
- 8 ITALIE
- 9 SUISSE
- 10 AFRIQUE DU SUD
- 11 SUEDE
- 12 AUTRICHE
- 13 AUSTRALIE
- 14 DANEMARK
- 15 MEXIQUE
- 16 FINLANDE
- 17 NORVEGE
- 18 NOUVELLE-ZELANDE
- 19 NIGERIA
- 20 VENEZUELA
- 21 YOUGOSLAVIE
- 22 SINGAPOUR
- 23 BRESIL
- 24 ESPAGNE
- 25 JAPON
- 26 ARGENTINE

1.2. Exemple d'utilisation indépendante de CØPERNIC : le réseau français des lignes téléphoniques à basse vitesse

Les données dont l'étude est présentée ici consistent en un tableau M carré  $20 \times 20$  à coefficients entiers - (données fournies par le GAPS). Les sommets du graphe associé représentent vingt villes françaises numérotées de 1 à 20 et les coefficients  $m_{ij}$  du tableau donnent le nombre de lignes téléphoniques reliant la ville  $i$  à la ville  $j$  pour une certaine gamme de vitesses de transmission (600 à 1200 bds).

Ce graphe est non orienté, puisque la matrice est symétrique, et représente une partie du réseau inter-régional de transmission de données de la Direction Générale des Télécommunications.

Nous avons travaillé avec des seuils absolus et des seuils relatifs mixtes, à différents niveaux, pour essayer de cerner des caractéristiques structurelles du réseau, en allant un peu plus loin que la simple mise en évidence du rôle privilégié de la Région Parisienne ...

Dans cette mesure, l'interprétation des résultats ne prétend pas décrire LA structure du réseau, mais veut faire ressortir quelques traits saillants qui sont d'ailleurs variables suivant le type de seuillage et le niveau de seuil, chaque type et chaque niveau apportant un éclairage différent des mêmes données de base.

Seuillage absolu :

Niveau de seuil	2	3	4	5	6
Nombre de niveaux	2	3	3	3	X
Nombre de centres	7	2	7	5	X
Nombre de sommets périphériques	13	2	3	3	X

A partir du seuil 6, le graphe n'est plus connexe, et donc plus hiérarchisable. Les images les plus intéressantes sont obtenues pour les seuils 3, 4 et 5, où l'on a 3 niveaux. Au seuil 3, les centres sont Paris et sa banlieue, ce qui n'a rien de surprenant, et l'on n'a que deux sommets périphériques : Rennes et Clermont-Ferrand. Toutes les autres villes sont au niveau intermédiaire, très interconnectées entre elles et avec les centres.

Rennes est liée à Bordeaux, Nantes et Rouen, et Clermont-Ferrand à Limoges et Lyon. Si Rennes est effectivement à la périphérie géographique du réseau, Clermont doit sa position au faible nombre de lignes qui la relie à l'extérieur. Parmi les villes intermédiaires, deux seulement apparaissent peu liées au reste du réseau : Strasbourg (Nancy et Paris) et Limoges (Orléans et Poitiers).

Au seuil 4, les villes-centres sont : Châlon, Dijon, Lyon, Marseille et Toulouse, en plus de Paris et sa banlieue. Les sommets périphériques sont : Limoges, Rennes et Strasbourg. Parmi les sommets du niveau intermédiaire, trois apparaissent déconnectés du reste de ce niveau, Clermont (qui était périphérique au seuil 3), Nancy et Montpellier.

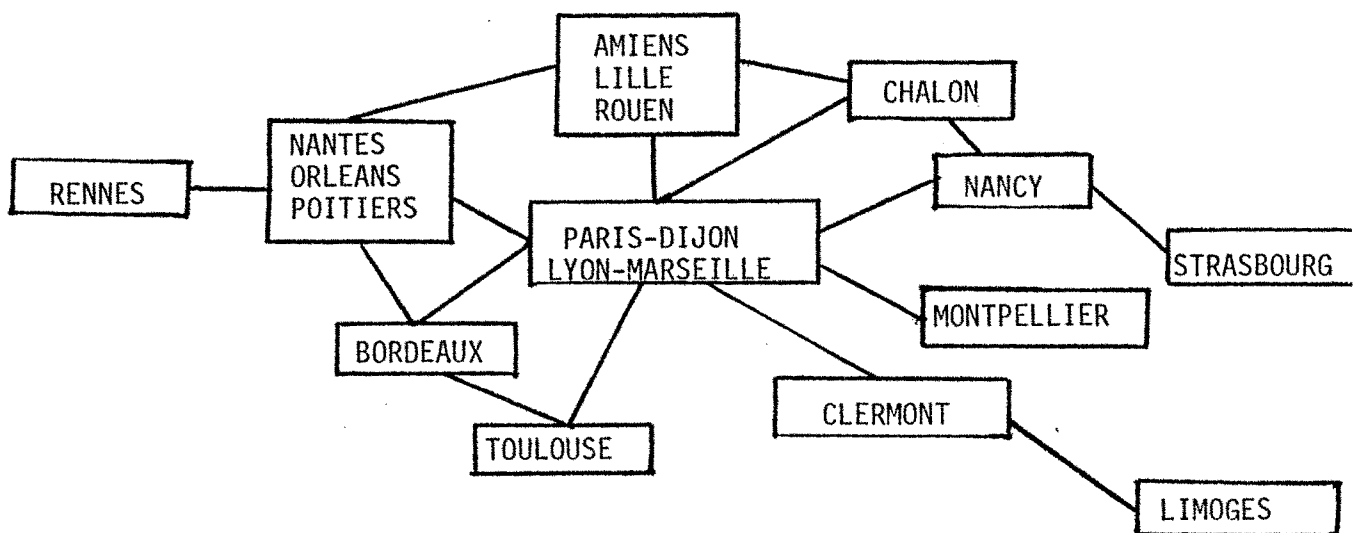
Au seuil 5, le nombre d'arcs a fortement diminué et les regroupements apparaissent clairement : Les centres sont Paris, Lyon et Marseille, soit l'axe PLM (Paris- Lyon - Méditerranée). Les sommets périphériques



n'ont pas changé : Clermont, Rennes et Strasbourg, respectivement raccordés à Limoges, Nantes et Nancy.

Au niveau intermédiaire, Montpellier est isolée et les autres se déploient de façon très symétrique de Nancy à Toulouse via le Nord et l'Ouest, avec deux groupes de trois : Amiens, Lille, Rouen et Nantes, Orléans, Poitiers ;

chacun de ces groupes étant chaîné à deux sommets : Châlon puis Nancy depuis Amiens, et Bordeaux puis Toulouse depuis Poitiers :



- Figure 35 -

Ainsi les caractéristiques suivantes se sont affinées à travers l'étude de ces trois niveaux de seuils : le cœur du réseau, son "épine dorsale", est l'axe PLM, avec une forte prédominance de Paris. Ensuite, un grand "fer à cheval" Nancy - Toulouse, et seulement deux "points gris" : Clermont et Montpellier et trois "points noirs" : Rennes, Strasbourg et Limoges.

La structure illustrée par ce diagramme n'était pas évident géographiquement : on pouvait s'attendre à trouver Rennes et Strasbourg à la périphérie, mais pas Limoges, et pourquoi pas Marseille ou Lille, tout aussi "excentrées" ?

Globalement, l'impossibilité d'obtenir plus de trois niveaux hiérarchiques et plus de trois sommets périphériques signifie que le réseau est assez bien interconnecté, et de façon équilibrée. Au seuil 6, Limoges se trouve déconnectée du réseau. Il serait peut-être intéressant de poursuivre l'étude sur les villes restantes, jusqu'à un éclatement plus significatif de la connexité.

Seuillage relatif :

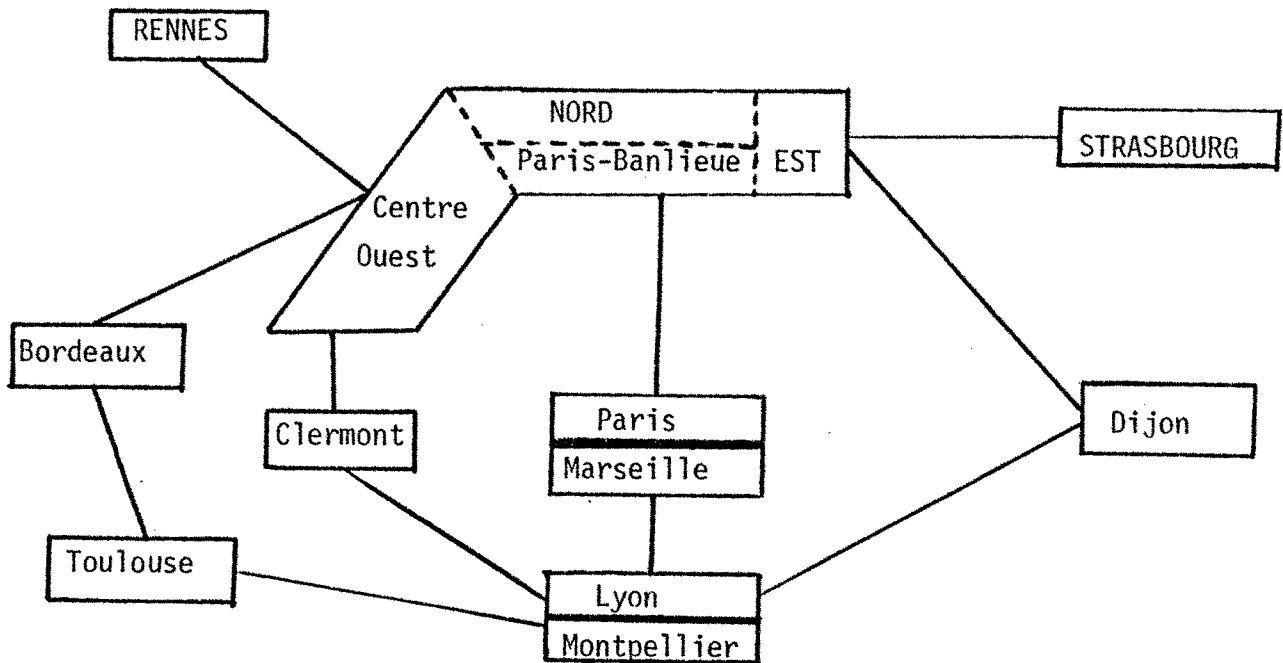
L'étude d'une série de seuils relatifs mixtes permet de compléter l'analyse précédente par une appréciation plus qualitative des liaisons, puisque le critère relatif permet de garder connectées des villes quantitativement peu liées au reste du réseau.

Niveau de seuil	5%	7.5%	10%	12.5%	15%	17.5%	20%
Nombre de niveaux	2	2	3	3	3	3	<del>X</del>
Nombre de centres	12	6	2	1	2	3	<del>X</del>
Nombre de sommets périphériques	8	14	2	9	6	6	<del>X</del>

Nous allons étudier les trois seuils 12,5%, 15% et 17% qui fournissent des répartitions des villes sur trois niveaux hiérarchiques, comme dans le cas des seuils absolus analysés précédemment.

Au seuil 12,5%, un seul centre, Paris, mais de nombreux sommets périphériques, 9 au total : Bordeaux, Clermont, Dijon, Limoges, Montpellier, Poitiers, Rennes, Strasbourg et Toulouse. Les seuls sommets isolés sur ce niveau sont Strasbourg et Dijon. On trouve donc parmi les sommets périphériques, non seulement les villes attendues (Strasbourg, Clermont) mais aussi des villes qui étaient plutôt centrales avec les seuillages absolus. Au niveau intermédiaire, on peut distinguer une composante Sud-Est (Lyon et Marseille) et un vaste groupe bien interconnecté : banlieue parisienne, Est, Ouest et Nord.

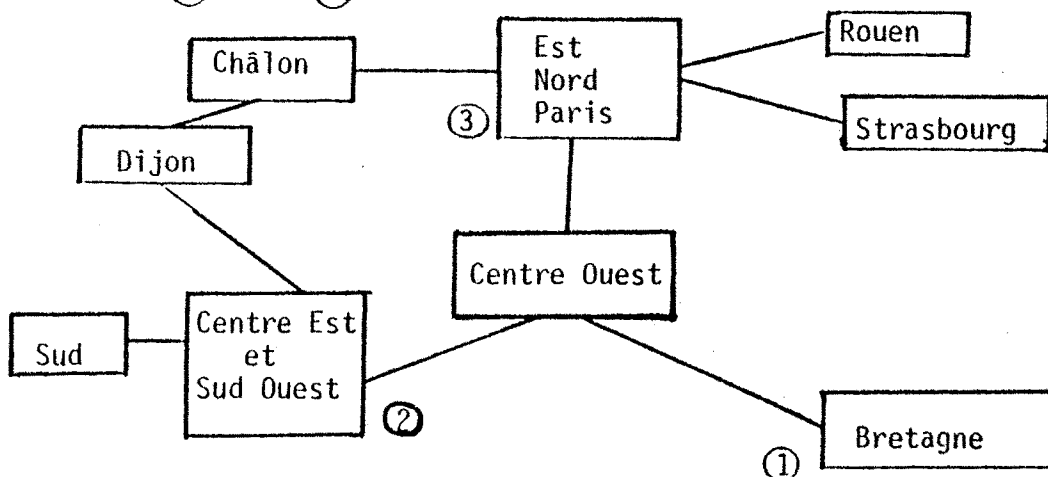
Au seuil 15%, deux centres, Paris et Marseille : on retrouve ainsi une trace de l'axe PLM détecté précédemment. Il y a seulement six sommets périphériques : Strasbourg, Clermont et Rennes qui confirment le seuillage absolu, et Dijon, Bordeaux et Toulouse. La bipolarisation au centre du graphe fait apparaître un clivage Nord-Sud avec une passerelle à l'Est et une au Sud-Ouest, (Dijon et Bordeaux-Toulouse). L'apparition de ces villes à la périphérie vient donc de ce que le seuillage relatif leur fait jouer un rôle de "relais" entre des groupes plus homogènes. On a le schéma suivant :



- Figure 36 -

Dans la mesure où le seuillage relatif ne respecte que les liaisons importantes par rapport aux autres, une ville comme Dijon ne garde que deux liaisons, Paris et Lyon, ce qui transcrit l'existence d'un axe privilégié sur lequel Dijon n'est qu'un relais. Il n'y a donc pas contradiction avec les résultats du seuillage absolu, mais transcription différente d'un même phénomène, avec en plus la mise en évidence d'un clivage Nord-Sud, déséquilibré en faveur du Nord, Paris étant un pôle bien plus puissant que Marseille.

Au seuil 17,5%, la structure a encore évolué, les trois centres sont Limoges, Orléans et Poitiers qui constituent une zone Centre-Ouest sur laquelle s'articule trois secteurs : la Bretagne (Nantes et Rennes), le Centre-Est et le Sud-Ouest (Clermont, Lyon, Bordeaux, Toulouse) avec Marseille et Montpellier rejetés à la périphérie, et Paris, le Nord (Amiens, Lille) et l'Est (Châlon, Nancy), avec Châlon-Dijon comme relais entre les secteurs ② et ③.



- Figure 37 -

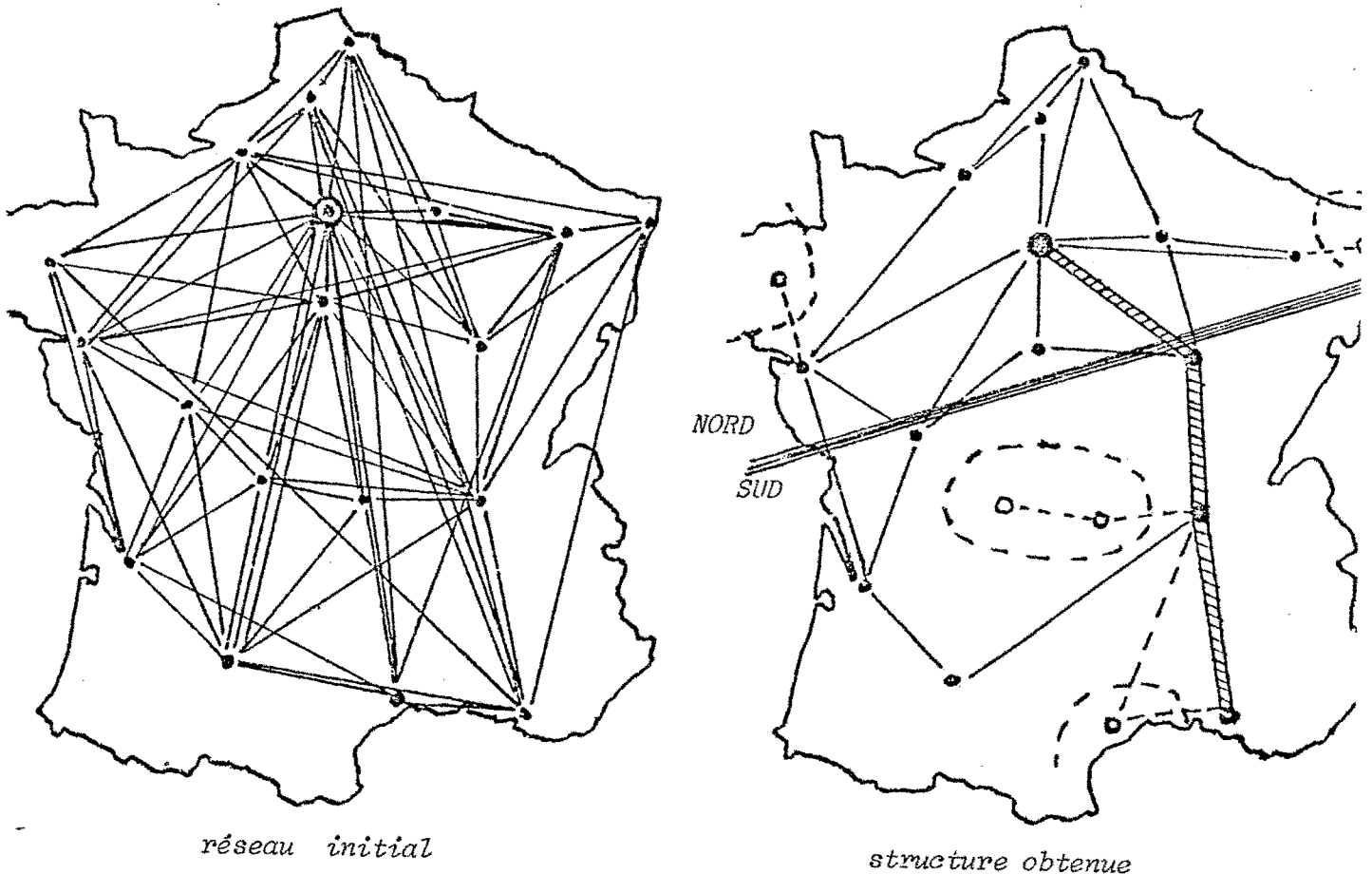
A force de ne conserver que les liaisons privilégiées entre une ville et deux ou trois autres, on obtient une structure faiblement interconnectée où la géographie reprend tous ses droits ; à part les deux villes-relais, les autres sommets périphériques correspondent effectivement à des zones excentrées : Bretagne, Alsace et Midi. L'attraction

de Paris étant ici limitée au Nord et à l'Est, le centre du graphe glisse naturellement vers le centre de gravité des trois secteurs, avec un déplacement vers l'Ouest qui répond à la passerelle Est entre les secteurs ② et ③. (Il ne reste à ce niveau de 17,5% que très peu d'arcs et au niveau 20%, le réseau est complètement détruit).

Finalement, le principal résultat des seuillages relatifs est d'affiner l'analyse faite auparavant en faisant apparaître un clivage Nord-Sud, grâce à la réduction sélective des zones d'attraction des villes pôles (Paris, Lyon et, dans une moindre mesure, Marseille et Poitiers), ce clivage se matérialisant sur l'arc PLM au niveau de Dijon qui apparaît comme une ville-relais rejetée à la périphérie.

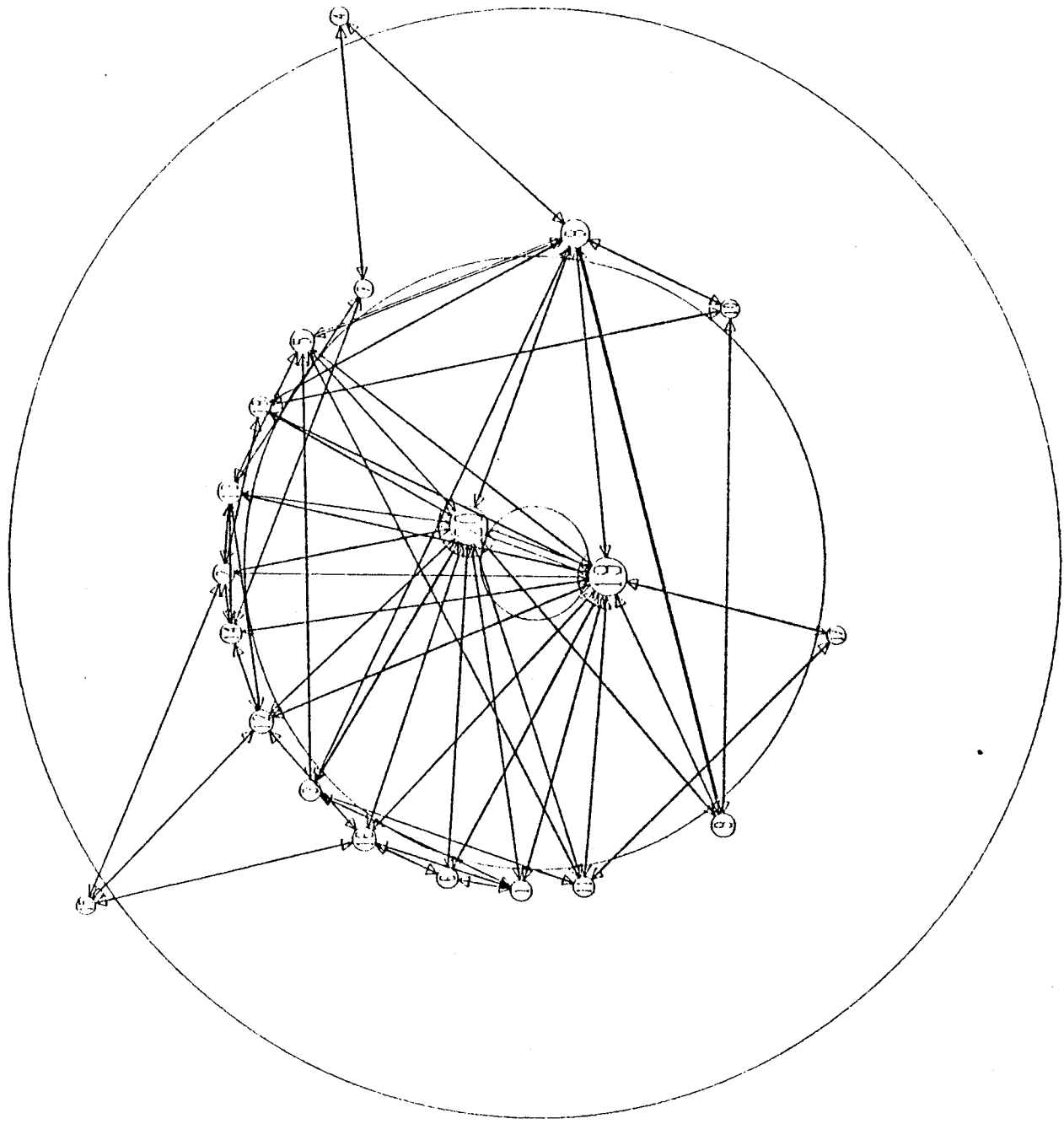
Pour le reste, on retrouve de façon très accentuée les mêmes régions périphériques enclavées : Bretagne et Alsace, et le Midi vient s'y ajouter, par un glissement de Marseille vers Montpellier dû à la réarticulation Ouest-Nord-Sud autour du Centre-Ouest.

On peut vérifier sur les deux dessins ci-dessous que la structure mise en évidence par seuillage et hiérarchisation centrifuge n'est pas une simple conséquence des positions géographiques :



- Figure 38 -

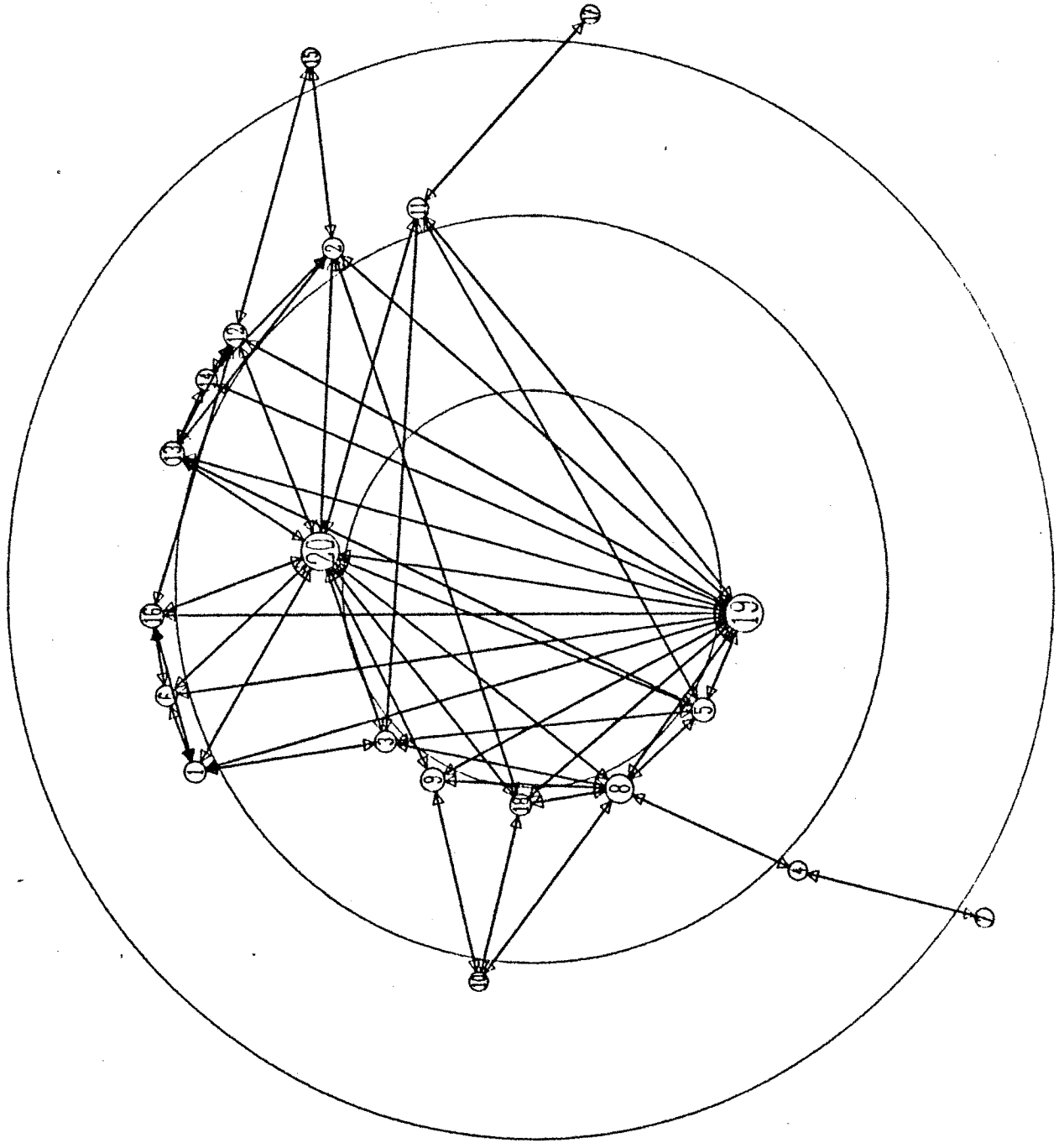
RESEAU BASSE VITESSE  
 HIERARCHIE CENTRIFUGE SEUIL ABSOLU 3



LEGENDE

- 1 AMIENS
- 2 BORDEAUX
- 3 CHALON SUR MARNE
- 4 CLERMONT-FERRAND
- 5 DIJON
- 6 LILLE
- 7 LIMOGES
- 8 LYON
- 9 MARSEILLE
- 10 MONTPELLIER
- 11 NANCY
- 12 NANTES
- 13 ORLEANS
- 14 POITIERS
- 15 RENNES
- 16 ROUEN
- 17 STRASBOURG
- 18 TOULOUSE
- 19 PAR. S - INTRAMURS
- 20 PAR. S - EXTRAMURS

RESEAU BASSE VITESSE  
 HIERARCHIE CENTRIFUGE SEUIL ABSOLU 4



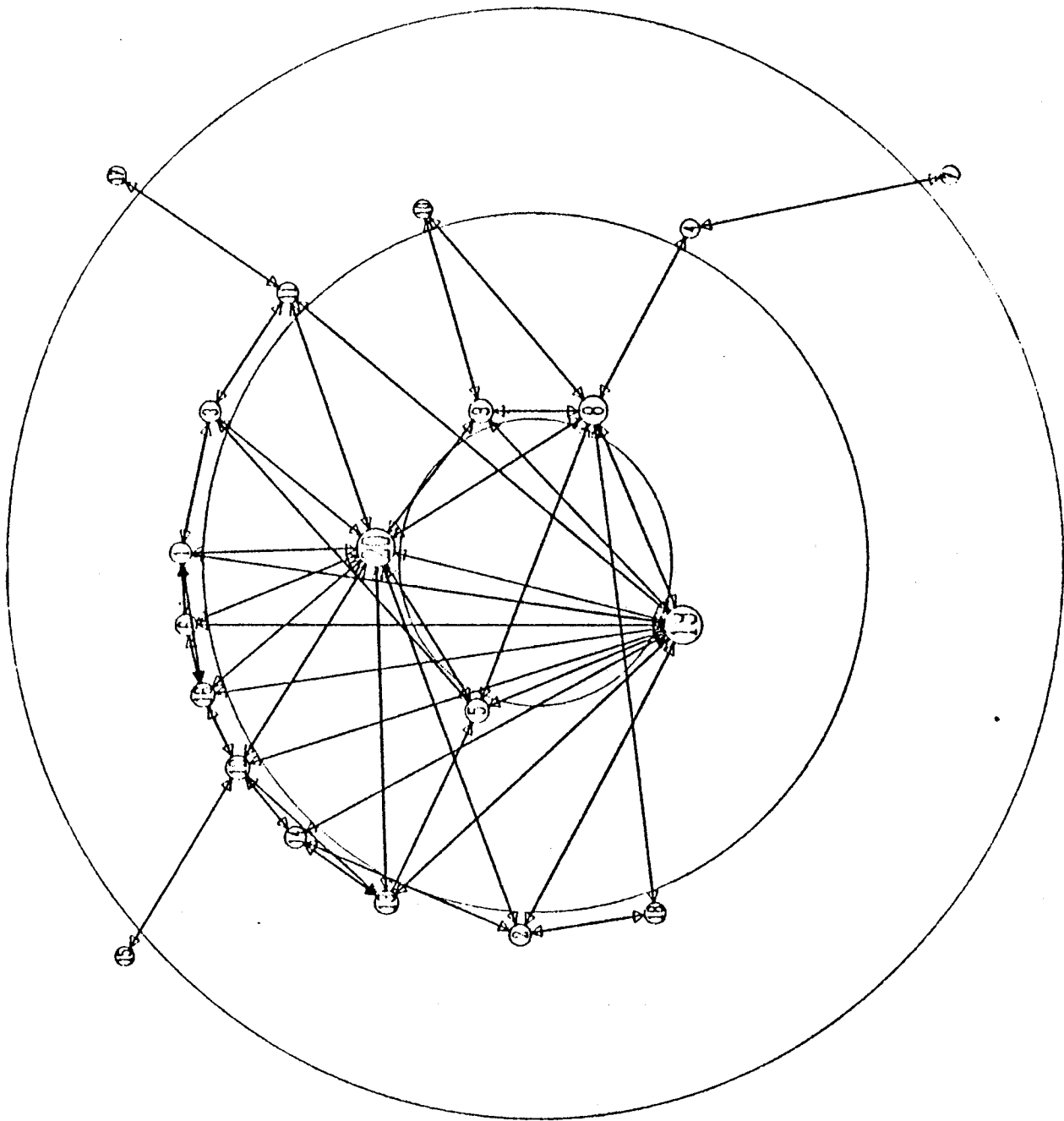
LEGENDE

- 1 AMIENS
- 2 BORDEAUX
- 3 CHALON SUR MARNE
- 4 CLERMONT-FERRAND
- 5 DIJON
- 6 LILLE
- 7 LIMOGES
- 8 LYON
- 9 MARSEILLE
- 10 MONTEPELLIER
- 11 NANCY
- 12 NANTES
- 13 ORLEANS
- 14 POITIERS
- 15 RENNES
- 16 ROUEN
- 17 STRASBOURG
- 18 TOULOUSE
- 19 PARIS-INTRAMUROS
- 20 PARIS-EXTRAMUROS

RESEAU BASSE VITESSE  
 HIERARCHIE CENTRIFUGE SEUIL ABSOLU 5

LEGENDE

- 1 AMIENS
- 2 BORDEAUX
- 3 CHALON SUR MARNE
- 4 CLERMONT-FERRAND
- 5 DIJON
- 6 LILLE
- 7 LIMOGES
- 8 LYON
- 9 MARSEILLE
- 10 MONTPELLIER
- 11 NANCY
- 12 NANTES
- 13 ORLEANS
- 14 POITIERS
- 15 RENNES
- 16 ROUEN
- 17 STRASBOURG
- 18 TOULOUSE
- 19 PARIS-INTRAMURS
- 20 PARIS-EXTRAMURS

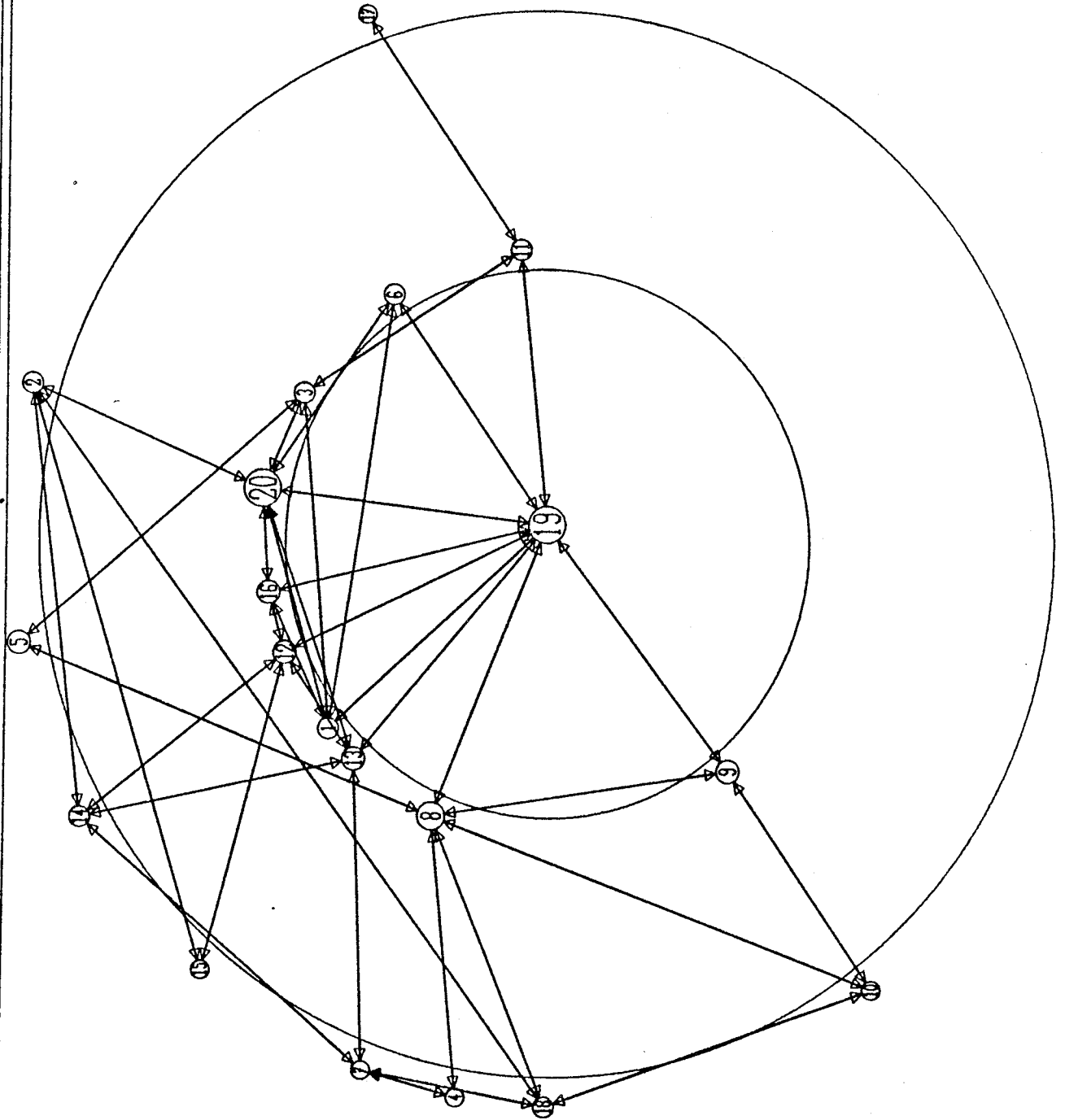




RESEAU BASSE VITESSE  
SEUIL RELATIF MIXTE 12,5%

LEGENDE

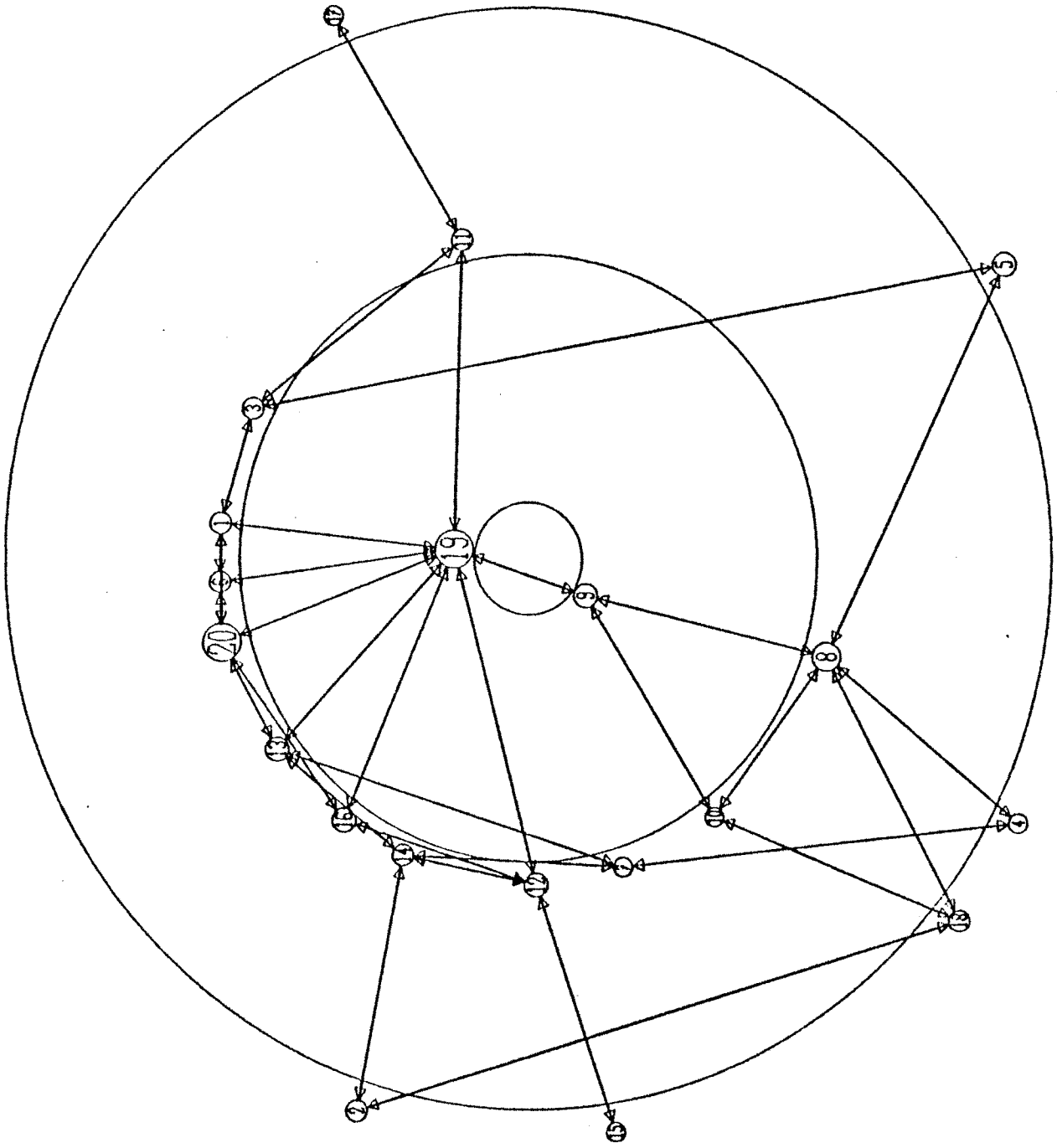
- 1 AMIENS
- 2 BORDEAUX
- 3 CHALON SUR MARNE
- 4 CLERMONT-FERRAND
- 5 DIJON
- 6 LILLE
- 7 LIMOGES
- 8 LYON
- 9 MARSEILLE
- 10 MONTPELLIER
- 11 NANCY
- 12 NANTES
- 13 ORLEANS
- 14 POITIERS
- 15 RENNES
- 16 ROUEN
- 17 STRASBOURG
- 18 TOULOUSE
- 19 PARIS-INTRAMURO:
- 20 PARIS-EXTRAMURO:



RESEAU BASSE VITESSE  
SEUIL RELATIF MIXTE 15%

LEGENDE

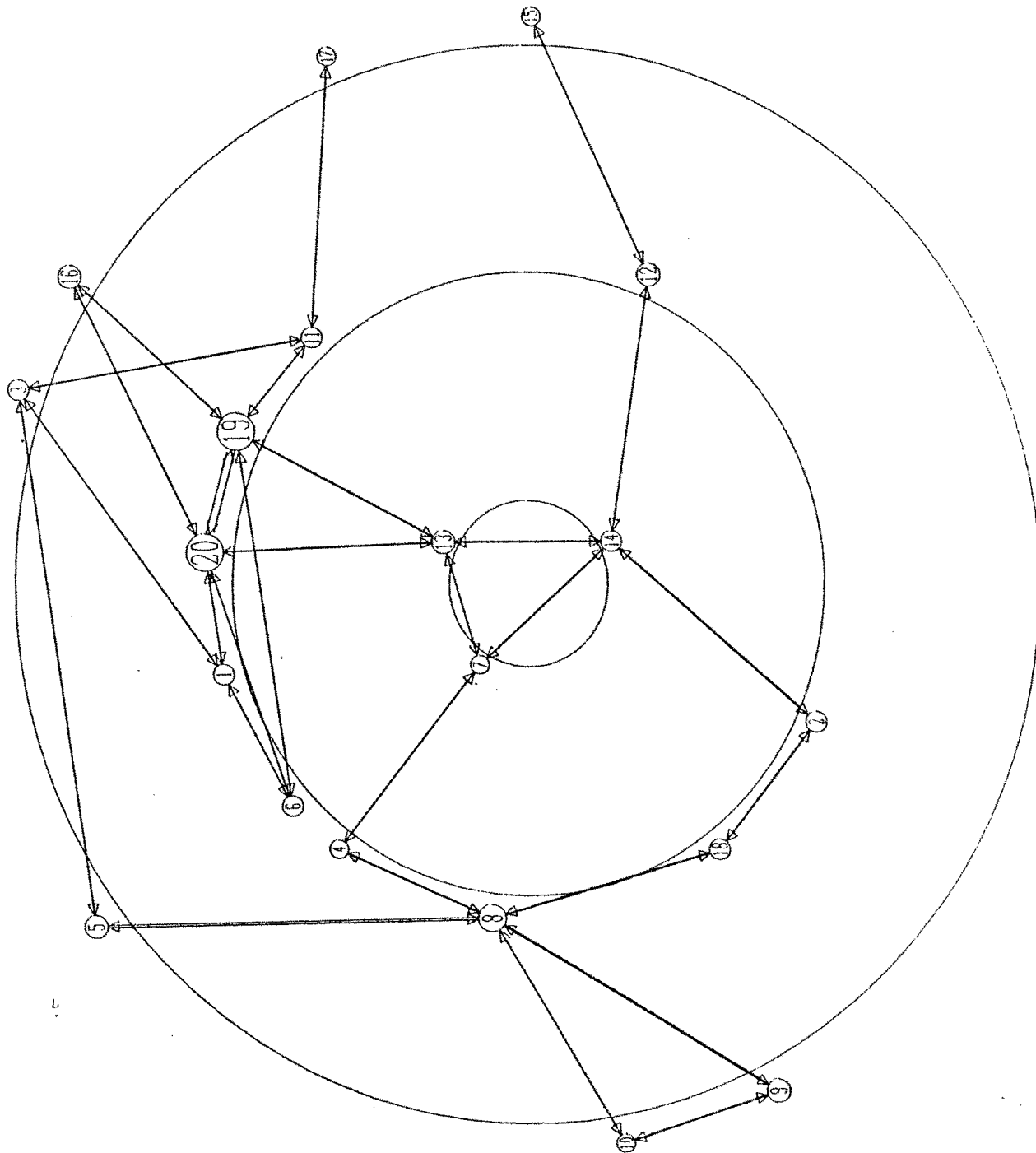
- 1 AMIENS
- 2 BORDEAUX
- 3 CHALON SUR MARNE
- 4 CLERMONT-FERRAND
- 5 DIJON
- 6 LILLE
- 7 LIMOGES
- 8 LYON
- 9 MARSEILLE
- 10 MONTPELLIER
- 11 NANCY
- 12 NANTES
- 13 ORLEANS
- 14 POITIERS
- 15 RENNES
- 16 ROUEN
- 17 STRASBOURG
- 18 TOULOUSE
- 19 PARIS-INTRAMUROS
- 20 PARIS-EXTRAMUROS



# LEGENDE

- 1 AMIENS
- 2 BORDEAUX
- 3 CHALON SUR MARNE
- 4 CLERMONT-FERRAND
- 5 DIJON
- 6 LILLE
- 7 LIMOGES
- 8 LYON
- 9 MARSEILLE
- 10 MONTEPELLIER
- 11 NANCY
- 12 NANTES
- 13 ORLEANS
- 14 POITIERS
- 15 RENNES
- 16 ROUEN
- 17 STRASBOURG
- 18 TOULOUSE
- 19 PARIS-INTRAMUROS
- 20 PARIS-EXTRAMUROS

## RESEAU BASSE VITESSE HIERARCHIE CENTRIFUGE SEUIL RELATIF 17.5



## 2. BILAN CRITIQUE POUR DE NOUVELLES DIRECTIONS DE RECHERCHE

Nous allons reprendre différents points de cette étude, dans l'ordre où ils ont été exposés ;

### 2.1. Hiérarchies dans les graphes

La définition que nous avons donnée d'un graphe hiérarchisé permet une structuration visuelle intéressante, essentiellement parce qu'en général le nombre de niveaux à identifier visuellement reste très inférieur au nombre de sommets du graphe.

En contrepartie, on n'obtient pas un classement très fin des sommets, qui peuvent être au même niveau tout en ayant des situations assez dissemblables. La vision globale du graphe facilite de ce point de vue des réajustements dans la structure hiérarchique, en particulier pour la hiérarchie des écartements, qui n'est pas aussi "naturelle" que la hiérarchie des rangs : on peut interpréter les dessins obtenus, en représentation centrifuge, comme décrivant la facilité d'accès de l'ensemble des points du réseau à partir de chaque sommet donné, ou inversement, comme une indication sur l'accessibilité de chaque sommet depuis l'ensemble du réseau. Ces deux façons de voir correspondent respectivement à la hiérarchie directe et à la hiérarchie inverse, et sont équivalentes dans le cas des graphes non orientés. Les représentations centrifuges sont bâties sur les écartements, ce qui implique que l'accessibilité des sommets n'y est pas représentée comme une moyenne sur tout le réseau, car le cas le plus défavorable est privilégié. (En effet, l'écartement est le maximum des plus courts chemins).

On peut essayer de moduler les résultats obtenus par les hiérarchies des écartements avec des calculs numériques plus fins, qui tendent à associer à chaque sommet un coefficient d'*accessibilité moyenne* (Ce genre de calculs ne permet pas de redéfinir une hiérarchie au sens du chapitre 2) Diverses méthodes ont été proposées (voir Harary [8], Lantner [11], et Rossier [19]), qui ne donnent généralement pas des résultats complètement concordants :

1. La centralité et l'anti-centralité relatives pour un sommet  $i$  :

$$CR(i) = \frac{\sum_{kj} e_{kj}}{\left( \sum_k e_{ik} \right)} \qquad ACR(i) = \frac{\sum_{kj} e_{kj}}{\left( \sum_k e_{ki} \right)}$$

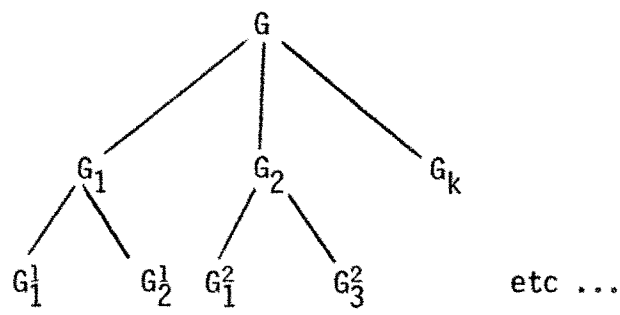
2. Les fonctions de dominance de Harary et de Lautner :

$$H(i) = CR(i) - ACR(i) \qquad L(i) = CR(i) / ACR(i)$$

3. Le degré de cohésion de Rossier  $DC(i)$ , qui est le nombre de fois où un sommet  $i$  fait partie d'un plus court chemin entre deux autres sommets quelconques;  $DC(i)$  donne des indications voisines de ce qu'on obtiendrait avec le critère :  $CR(i) + ACR(i)$ .

D'autre part, il serait intéressant d'étendre la visualisation de graphes hiérarchisés à d'autres types de structures hiérarchiques qui pourraient présenter un intérêt pratique. Nous pensons ici plus particulièrement aux structures arborescentes dérivées de la recherche des ensembles minimaux d'articulation dans des graphes orientés ou non et qui ont des applications en automatique (cf. Richetin [18]).

Ces structures s'obtiennent de la façon suivante :  
Dans un graphe  $G$ , on cherche un ensemble minimal d'articulation  $A$  qui partage  $G$  en  $G_1, G_2, \dots, G_k$  puis on itère le processus sur chacun de ces sous-graphes :



L'inconvénient majeur de ce type de décomposition est qu'il n'est pas unique : si on a le choix entre plusieurs ensembles  $A_i$  initiaux différents, on aura pour chacun un arbre de décomposition différent des autres, et de même à chaque niveau.

Néanmoins, on pourrait chercher à représenter visuellement ce genre de structures, par exemple en dessinant successivement les différentes étapes de la décomposition, et éventuellement en guidant cette décomposition à travers un dispositif interactif.

## 2.2. Croisements d'arêtes et problèmes connexes

### 2.2.1. Dessin automatique

En ce qui concerne la lisibilité des graphes, les résultats obtenus par les heuristiques de placement décrites au chapitre 3 sont assez satisfaisants sur des graphes de taille assez petite (15 à 60 sommets environ) et qui comportent peu de longs shunts.

Dans le cas de graphes de taille plus importante (60 à 100 sommets) ou contenant de nombreux arcs les résultats sont en général moins bons. Ceci se traduit concrètement par le fait que dans le premier cas une série de corrections manuelles (en général 5 à 6 sommets à déplacer) suffit à peu près toujours pour obtenir un très bon dessin. Dans le second cas, il est plus souvent nécessaire de recourir à 2 ou 3 séries de corrections pour avoir un bon dessin. Dans LØVIGREC, le problème qui apparaît le plus mal résolu est celui des longs shunts, surtout quand ils coupent des c.f.c. de grande taille. Dans CØPERNIC, le système de placement à partir d'une gestion en pile fonctionne bien, sauf quand la pile devient très profonde. Ceci se traduit sur le dessin par des arcs qui deviennent presque diamétraux, ce qui provoque de nombreux croisements. Heureusement, il suffit alors dans la plupart des cas de déplacer deux ou trois sommets pour avoir un dessin correct.

### 2.2.2. Bande-diagonalisation

Nous avons remarqué au chapitre 3 la parenté existant entre le problème de la minimisation du nombre de croisements d'arêtes dans un graphe biparti et celui de la bande-diagonalisation des matrices.

De ce fait, la méthode des degrés relatifs constitue certainement un outil heuristique très efficace pour "bande-diagonaliser" des matrices carrées ou rectangulaires.

Cet algorithme pourrait avoir des applications d'une part en analyse numérique, où l'on rencontre fréquemment de grandes matrices creuses se prêtant bien à ce genre de manipulations, et d'autre part, en sciences sociales pour classer des tableaux croisant diverses caractéristiques sur un échantillon (cf. Bertin [3] : ordonnancement de scalogrammes).

Son utilisation serait limitée par deux problèmes :

- \* l'encombrement mémoire, qui n'autorise le traitement de très grandes matrices que si elles sont creuses.

- \* la valuation booléenne imposée jusqu'ici, alors qu'on peut souhaiter utiliser un système de gradation des influences (du genre des matrices multi-seuillées).

On pourrait essayer de tester la validité de la méthode dans ce second cas, en faisant les calculs de degrés relatifs avec des pondérations variables et non plus seulement des valeurs booléennes.

### 2.2.3. Classification automatique interactive

Nous avons remarqué au chapitre 3 que les coefficients des matrices d'itération de la méthode des degrés relatifs,  $M_1(i,j)$  et  $M_2(i,j)$  pouvaient être considérés comme des intensités de liaison entre les sommets  $i$  et  $j$ , respectivement du niveau 1 et du niveau 2.

Lorsque l'on traite un graphe biparti par cette méthode pour diminuer le nombre de croisements, on tend à rapprocher les uns des autres sur chaque niveau, les sommets  $i$  et  $j$  pour lesquels les coefficients  $M_1(i,j)$

ou  $M_2(i,j)$  sont les plus élevés. On réalise donc une sorte de *classification automatique* des sommets en fonction de la structure du graphe, classification pour laquelle  $M_1$  et  $M_2$  jouent le rôle de matrices de ressemblance. Nous nous proposons d'utiliser la méthode des degrés relatifs pour bâtir *interactivement* des classifications sur deux ensembles entre les éléments desquels existent un certain nombre de relations.

L'idée de base est la suivante : si la disposition des sommets sur les niveaux est bonne et si les relations entre les niveaux sont établies correctement, le dessin du graphe biparti doit comporter relativement peu de croisements.

On travaillera en deux temps :

1. Etant donné un graphe biparti, on utilise la méthode des degrés relatifs pour classer les sommets sur les niveaux.
2. Si le dessin obtenu est jugé non satisfaisant, on modifie la structure du graphe aux endroits considérés comme critiques.

Puis on retourne à la première étape pour dessiner le graphe ainsi modifié en calculant une nouvelle classification sur chaque niveau.

La première étape permet d'assurer une bonne répartition des sommets, la deuxième de contrôler que les relations établies entre les deux niveaux ont la structure attendue et sinon d'y remédier par des modifications bien choisies.

### Le logiciel AMNESIC

Suivant le schéma esquissé au paragraphe précédent, on va maintenant décrire les caractéristiques essentielles d'un logiciel interactif pour la conception assistée de nomenclatures, dont la réalisation et l'expérimentation seront envisagées ultérieurement.

Le but assigné au logiciel AMNESIC (Analyse et Modification des Nomenclatures Eclatées par Suppression Interactive de Croisements) est de faciliter la classification dans les cas où la structure qu'on cherche



à obtenir est approximativement un partitionnement de l'ensemble des sommets du niveau 2 par rapport à ceux du niveau 1 :



Dans le cas idéal du graphe ci-dessus, on a une "nomenclature éclatée" parfaite : chaque sommet du niveau 2 n'est lié qu'à un seul sommet du niveau 1, et on a donc exactement une partition.

Nous appelons ce graphe une *nomenclature éclatée* parce que les sommets du niveau 2 sont en un certain sens des *décompositions* des sommets du niveau 1 : les  $X_i$  sont des "rubriques" dans lesquelles on range un certain nombre d'éléments  $Y_j$ . Dans la réalité, on n'aura généralement que des partitionnements approximatifs, soit que l'on hésite a priori à rattacher un sommet  $Y_j$  à tel  $X_i$  plutôt qu'à tel autre, soit que pour un  $Y_j$  donné plusieurs  $X_i$  lui sont liés du fait de la signification de la relation entre les  $X_i$  et les  $Y_j$ . Par exemple, on pourra étudier des graphes objectifs-moyens (ou matériels-composants) qui donneront des partitionnements très approximatifs.

#### a) Critères d'analyse

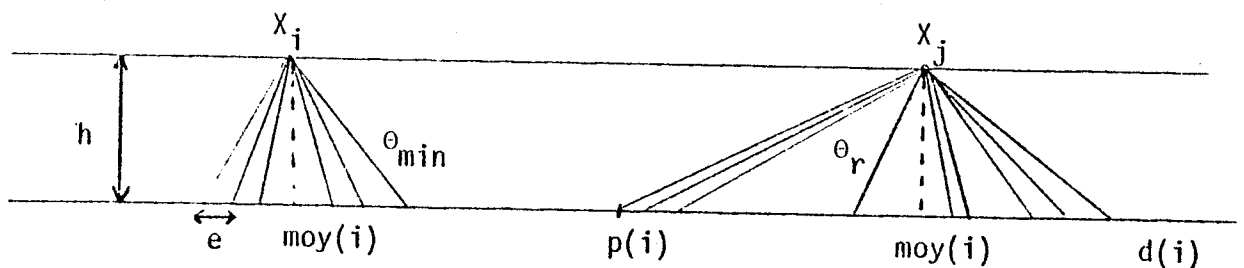
Le programme donne une liste de sommets, établie selon trois critères, pour guider l'utilisateur dans les modifications qu'il souhaite effectuer.

Les critères retenus pour établir cette liste sont :

- 1 \* le degré des sommets
- 2 \* la dispersion du faisceau d'arcs
- 3 \* l'existence de longs cycles.

L'utilisateur fixe ainsi trois normes : l'intervalle de variation du degré des sommets, l'intervalle angulaire de dispersion des faisceaux d'arcs sur la représentation, la longueur maximale tolérée pour les cycles sans qu'elle soit signalée à l'utilisateur. En réponse, il obtient la liste des sommets pour lesquels l'une ou l'autre de ces trois normes est transgressée. Ces trois critères nous semblent permettre la détection de toutes les anomalies majeures, sources de croisements d'arêtes dans le graphe.

Sur le plan du calcul, le premier critère ne crée pas de difficulté : l'utilisateur fournit  $d_{\min}$  et  $d_{\max}$  ((éventuellement deux couples de valeurs  $(d_{\min}, d_{\max})$  : un pour chaque niveau) ; le programme met dans la liste des anomalies tous les sommets qui tombent en dehors de cet intervalle. Le deuxième critère est un critère *relatif* qui doit permettre de signaler les arcs qui s'éloignent trop de la verticale ; on peut l'utiliser sur un niveau ou sur les deux. Par exemple, pour les sommets du niveau 1, en supposant que tous les sommets sont répartis à intervalle régulier sur chaque niveau, on peut calculer un angle minimum de dispersion du faisceau et le comparer à l'angle réel observé.



- Figure 39 -

Si on se ramène à la verticale de la moyenne des rangs des successeurs de  $X_j$  l'angle minimum est obtenu en supposant que tous les successeurs de  $X_j$  sont les uns à côté des autres, sans "trous".

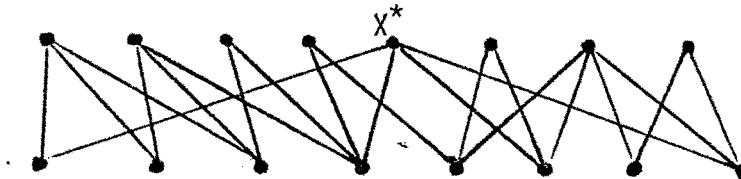
$$\text{On a : } \theta_{\min} = \text{Arctg } \frac{\Delta(i)}{2\lambda}$$

où  $\lambda$  est le rapport  $\frac{h}{e}$ , et où  $s(i)$  est le degré de  $X_i$ . L'angle réel  $\theta_r$  est obtenu à partir des rangs  $p(i)$  et  $d(i)$  du plus à gauche et du plus à droite des successeurs de  $X_i$  par :  $e' = \max(\text{moy}(i) - p(i), d(i) - \text{moy}(i))$

$$\text{D'où : } \theta_r = \text{Arctg } \frac{e'}{\lambda}$$

Connaissant  $\theta_{\min}$  et  $\theta_r$ , on peut mettre  $X_i$  dans la liste quand  $\frac{\theta_r}{\theta_{\min}}$  dépasse une valeur limite (fixée par l'utilisateur).

Le troisième critère joue un rôle essentiellement *négatif* : détecter les longs cycles peut permettre d'éviter des modifications inutiles parce que non significatives par rapport à la structure de la relation étudiée :



Sur ce graphe, on aura une mauvaise répartition angulaire pour  $X^*$  due à l'existence d'un cycle de longueur 16 (passant par tous les sommets) pour lequel  $X^*$  ne joue pas un rôle particulier. Ce critère est difficile à mettre en oeuvre pratiquement. Comme on ne cherche pas à trouver tous les longs cycles, on peut se contenter de les chercher à partir des sommets mis dans la liste des anomalies sur la base du deuxième critère, ce qui limiterait le volume des calculs à effectuer.

#### b) Modifications du graphe

Une fois déterminée la liste des sommets pour lesquels des modifications sont souhaitées, l'utilisateur dispose de quatre opérations élémentaires

pour effectuer des améliorations ; à titre de contrôle il peut visualiser sur un écran tout ou partie du graphe pour préciser les modifications :  
Les quatre opérations élémentaires sont :

- \* créer un arc
- \* supprimer un arc
- \* couper en deux un sommet
- \* fusionner deux sommets

Après enregistrement des modifications, une nouvelle représentation est calculée et dessinée à l'écran.

S'il le désire, l'utilisateur peut alors enclencher une nouvelle série de modifications en fixant de nouvelles bornes de détection des anomalies (degrés, dispersions angulaires et longs cycles).

On peut aussi envisager de mener les opérations de correction directement au vu du dessin du graphe, sans passer par la phase de calcul des anomalies par le système des trois normes.

### c) Mise en oeuvre

Du point de vue du fonctionnement concret, le programme devra être complété par diverses facilités d'utilisation : par exemple des sous-programmes affichant à la demande le libellé des sommets dans le graphe ou sauvegardant des états intermédiaires du graphe ...

Dans l'heuristique de classement des sommets, il sera utile de prévoir une décomposition préalable du graphe traité en composantes simplement connexes, car même si le graphe initial est connexe, il peut cesser de l'être après modification, sans que l'utilisateur puisse s'en apercevoir si les composantes connexes sont superposées.

### 2.3. Problèmes de seuillage

Nous avons vu au chapitre 4 qu'en dehors des critères subjectifs (dessin attrayant, conformité aux résultats attendus) le seul moyen de classer objectivement les graphes obtenus par seuillage était d'essayer d'apprécier la stabilité relative des structures hiérarchiques correspondant aux

diverses plages de seuillage. Pour celà, nous allons donner quelques éléments de réponse à deux questions complémentaires qui apparaissent dans ce problème de stabilité :

1. Qu'est-ce qu'un changement de structure hiérarchique ?
2. Comment mesurer la ressemblance entre deux graphes hiérarchisés ?

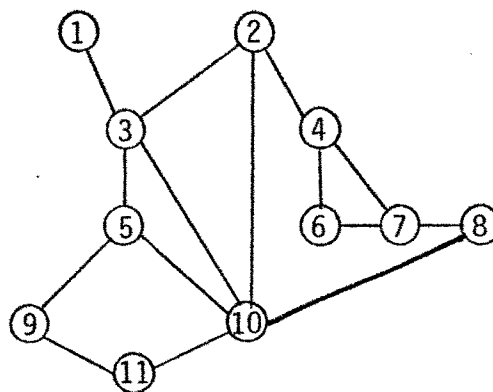
### 2.3.1. Changement qualitatif dans une structure hiérarchique

La façon la plus rigoureuse de délimiter les plages de seuillage serait de considérer que la structure change dès qu'un arc disparaît lorsqu'on élève le niveau du seuil.

En fait, cette façon de voir est sans doute trop rigide car certains arcs disparaissent sans altérer les principales caractéristiques de la structure : les niveaux peuvent rester inchangés, de même que les c.f.c. Cette observation permet de séparer les arcs en deux catégories, suivant le rôle qu'ils jouent dans la structure hiérarchique à un moment donné :

- arcs critiques
- arcs redondants.

Exemple :



- Figure 40 -

Sur la figure 40 les arcs (2,10) et (3,10) sont redondants. L'arc (1,3) est critique. Les arcs (2,4) et (8,10) sont l'un critique, l'autre redondant suivant lequel disparaît le premier quand on élève le seuil.

On peut donc en première approximation retenir que le critère de stabilité est celui des arcs critiques, qui peuvent être à tout niveau de seuil connus parfaitement, puisque l'ordre dans lequel les arcs disparaissent du graphe est déterminé par le choix de la méthode de seuillage d'une part, et par leur propre poids (relatif ou absolu) d'autre part.

Ainsi, à chaque graphe structuré hiérarchiquement on peut associer son "armature" (l'ensemble des arcs critiques) en précisant quels sont les arcs *réellement* critiques.

## 2. RESSEMBLANCE ENTRE DEUX GRAPHES HIERARCHISES

De la même manière, définir la ressemblance de deux graphes hiérarchisés sur la base d'un isomorphisme serait un critère trop exigeant. Il nous paraît plus intéressant de chercher à construire une batterie d'indices *relatifs* dont la comparaison permettrait une appréciation plus nuancée. par exemple pour deux graphes en représentation réduite éclatée on pourrait comparer les rapports suivants :

- nombre de c.f.c. / nombre de sommets
- taille de la plus grosse c.f.c. / nombre de niveaux
- nombre d'arcs / nombre de sommets
- nombre de shunts / nombre d'arcs
- etc ...

de façon à obtenir un panorama assez complet des caractéristiques de chacun des graphes, indépendamment de leurs tailles respectives, et surtout en liaison étroite avec la structure hiérarchique considérée. Ce genre de travail serait d'autant plus utile que les méthodes de dessin automatique peuvent donner des résultats d'allures très différentes pour des graphes pratiquement identiques.

A partir de ces différents critères : armature du graphe, indices de ressemblance hiérarchique, on pourra chercher à regrouper entre elles les plages de seuillage pour lesquelles ces critères donneront des résultats voisins, et dégager ainsi quelques structures types sur l'ensemble de la zone efficace de seuillage, en faisant abstraction des variations de structure les moins significatives.

## 2.4. Fonctionnement des logiciels

Les logiciels LØVIGREC et CØPERNIC ont été élaborés progressivement et nous avons cherché à respecter trois impératifs :

### 2.4.1. Souplesse d'utilisation

Cet objectif est rempli grâce à la diversité des entrées et des sorties et à la structure hiérarchisée du dialogue, qui permettent théoriquement de travailler sur des données variées, d'interrompre le traitement pour le reprendre ultérieurement au point où on en était resté et de se repérer facilement dans les différentes phases des programmes.

Le matériel disponible au CICG étant limité, particulièrement sur le plan de l'interactivité (système de partage de temps très chargé, terminal graphique peu adapté à ce type d'utilisation) nous n'avons pas pu réellement vérifier cette souplesse de fonctionnement en pratique. La configuration idéale de travail serait de disposer d'un terminal graphique d'une technologie permettant l'*effacement sélectif* de parties du dessin et offrant un niveau élevé d'interactivité.

### 2.4.2. Accessibilité

Nous avons essayé de rendre nos programmes utilisables par des non-informaticiens (voire même des non-scientifiques). Dans ce but, le fonctionnement des programmes est assuré par une seule commande initiale, la modification des noms des fichiers est prise en charge par programme, la plupart des réponses à fournir par l'utilisateur sont "oui", "non" ou un numéro. Les quelques notions élémentaires de théorie des graphes

qui figurent au chapitre 2 suffisent pour comprendre ce que représentent les dessins obtenus et pour les interpréter dans leur contexte propre, économique, informatique ou autre, comme nous avons nous-mêmes interprété les deux exemples donnés plus haut.

Sur le plan de la simplicité du maniement, le seul point qui pose vraiment un problème est la phase de préparation des données par seuillage. Nous avons déjà indiqué les problèmes théoriques posés par le choix d'une méthode de seuillage et de niveaux de seuil par rapport à l'interprétation des résultats obtenus.

Seule une certaine pratique empirique des différentes méthodes sur des exemples qu'il connaît peut ici permettre à l'utilisateur de choisir un seuil relatif en lignes à 10% plutôt qu'un seuil absolu au niveau  $x$  ou bien l'inverse. L'acquisition d'un certain savoir faire intuitif facilite la recherche des zones de seuillage les plus intéressantes, mais, comme nous l'avons dit au chapitre 4, la seule garantie d'objectivité dans l'analyse est l'exhaustivité.

#### 2.4.3. Vitesse de fonctionnement

Dans la mesure où les temps de réponse dépendent du matériel et plus précisément du degré de saturation du système de partage de temps, les observations faites en 2.4.1. sur le matériel du CICG restent valables.

En ce qui concerne les temps d'exécution intrinsèques de nos programmes ils sont variables suivant la structure des graphes traités : nombre de niveaux hiérarchiques, tailles des c.f.c. etc ...

Pour LØVIGREC, les graphes comportant environ 25 à 35 sommets donnent des images en 5 à 6 s et des graphes de grande taille (70 à 90 sommets) demandent 10 à 12 s. Le temps ainsi décompté comprend : la hiérarchisation du graphe, le calcul des coordonnées par l'heuristique, le calcul et l'affichage des points à l'écran.



Pour CØPERNIC, on a des temps correspondants de l'ordre de 5s avec des graphes à 20 sommets et de 10s avec des graphes comportant plus de 30 sommets. Par rapport aux algorithmes les plus coûteux (le double marquage des c.f.c. et le raccourcissement des arcs dans LØVIGREC et l'algorithme de Dantzig pour les plus courts chemins dans CØPERNIC sont tous en  $n^3$ ) on peut remarquer un taux de progression beaucoup moins fort des temps de calcul. A celà, deux explications : la plupart des phases d'initialisation, d'entrées sorties etc ... sont en  $n$  et la majeure partie des calculs sont proportionnels au nombre d'arcs (borné en  $n^2$ ), ce qui tend à atténuer la progressivité : on a en fait probablement une formule du type :  $t = an + bn^2 + cn^3$  où  $a$  et  $b$  sont nettement supérieurs à  $c$ .

D'autre part, les bornes théoriques des algorithmes sont calculées en prenant pour le nombre d'arcs une borne en  $n^2$ . Or, étant donné le mode d'obtention des données (seuillages), les matrices sont d'autant plus creuses que leur taille est plus grande. En fait, pour les exemples sur lesquels nous avons travaillé, le nombre d'arcs croît comme  $n$  plutôt que comme  $n^2$ , ce qui transforme les algorithmes de LØVIGREC en algorithmes bornés en  $n^2$ .

## 5. ACQUIS METHODOLOGIQUE ET EXTENSIONS ENVISAGEES

De façon générale, la réalisation pratique d'outils infographiques nous a imposé une réflexion méthodologique sur l'organisation de notre travail, vu l'importance des tâches concrètes de programmation à effectuer et la variété des problèmes rencontrés.

Le découpage des logiciels, tel que nous l'avons décrit au chapitre 5, est fondé sur des observations de bon sens :

- séparation absolue entre calcul et affichage
- modularité hiérarchisée des sous-programmes.

Néanmoins, la mise en oeuvre de ces principes dans le détail de chaque phase du programme n'a pas toujours été facile et les logiciels sont passés par plusieurs états intermédiaires avant d'acquérir la structure décrite ici.

Au-delà des tâtonnements inévitables de mise au point, deux traits importants nous semblent devoir être retenus :

a) Le "schéma de fabrication" d'un tel logiciel en trois étapes (cf. Delarche [6]).

1. définition d'un mode de hiérarchisation
2. définition d'une forme standard de représentation
3. placement heuristique des sommets

où l'étape intermédiaire joue le rôle d'interface entre les aspects mathématiques du problème et ses aspects graphiques.

b) La cohérence qui tend (ou doit tendre) à se dégager peu à peu de la confrontation entre l'organisation interne des programmes et leur fonctionnement du point de vue de l'utilisateur : ainsi, le regroupement dans une sorte de "catalogue graphique" des objets graphiques servant d'éléments de base aux dessins (hexagones, carrés, flèches, etc ...) permet de se rendre compte que ceux-ci sont en fait des combinaisons variées d'un petit nombre d'objets de base.

Du point de vue de l'utilisateur, ceci se traduit par une impression de "paysage familier" qui facilite la lecture de l'image et les distinctions nécessaires (ainsi l'opposition hexagone-carré traduit la différence conceptuelle sommet-c.f.c., y compris dans le cas où la c.f.c. est un sommet isolé).

Dans la perspective d'améliorer la prise en compte d'aspects plus dynamiques de l'analyse de systèmes, il est prévu d'adjoindre aux logiciels des sous-programmes permettant, à partir de données auxiliaires (cf. chapitre 5) fournies par l'utilisateur, de visualiser la progression en temps discrétisé d'une perturbation issue d'un sommet et qui progresse dans le réseau, jusqu'à extinction de son effet par amortissement et/ou jusqu'à un horizon donné. Dans le même ordre d'idée, il pourrait être utile de renforcer l'interactivité de manière à pouvoir altérer les données directement sur l'écran et à reconstruire une structure hiérarchique modifiée (cf. le projet de logiciel AMNESIC) pour répondre à des questions telles que : "Cet arc est-il important dans la structure actuelle ?" - "Si j'ajoute un nouvel arc entre ces deux sommets, que devient la structure hiérarchique ?" etc

### 3. CONCLUSIONS GENERALES

Globalement, nous pensons avoir atteint le but de notre travail : réaliser des logiciels de dessin automatique de graphes hiérarchisés, en vue d'une présentation synthétique et structurée de l'information disponible sur des systèmes complexes.

Sur certains points, des problèmes restent en suspens (en particulier, le seuillage et les croisements d'arêtes). Dans ces cas, nous avons manqué de temps, de recul expérimental, et sans doute aussi de perspicacité théorique, pour apporter des solutions plus définitives. A d'autres points de vue, nous avons dépassé les objectifs de départ (introduction de la représentation centrifuge, formalisation matricielle de la méthode des degrés relatifs) et ces avancées nous ont permis de resituer les problèmes dans un cadre plus large (notion générale formelle de graphe hiérarchisé, aspects de classification automatique dans le problème des croisements d'arêtes).

Indépendamment des limites de notre travail, les bornes propres à ce type de représentation des systèmes nous sont apparues clairement : Nous avons arrêté à 100 et 40 sommets respectivement les tailles des graphes traitables par LØVIGREC et CØPERNIC. En fait, l'expérimentation montre qu'au delà de 60 et 20 sommets on perd rapidement les qualités qui justifient ces représentations : la vision globale des interactions, et la lisibilité des graphes. De même, si  $n$  est le nombre de sommets du graphe, il n'est guère possible d'obtenir des dessins intéressants si le nombre d'arcs dépasse  $2n$ .

Ces logiciels n'auront d'intérêt que pour visualiser de petits systèmes comportant peu de liaisons internes, ce qui restreint considérablement leur champ d'application (on construit maintenant couramment des modèles comportant des centaines de variables et des milliers d'équations !).

Néanmoins, compte tenu de ces limitations, les logiciels ainsi conçus peuvent rendre de nombreux services, aussi bien sur le plan de la présentation de l'information que sur celui de l'exploration heuristique des structures.

Nous espérons en avoir convaincu le lecteur au fil de ce long exposé.

Grenoble, le 31 mai 1979



REFERENCES

---

- [1] Cl. BERGE : Graphes et hypergraphes ARMAND COLIN 1973.
- [2] L. von BERTALANFFY : Théorie Générale des Systèmes, DUNOD 1973.
- [3] J. BERTIN : La graphique et le traitement graphique de l'information, FLAMMARION 1977.
- [4] M.J. CARPANO : Informatique et aide à la décision. Rapport final d'exécution du contrat de recherche IRIA n° 77-005 projet SESORI n°171.
- [5] R.C. CONANT : Laws of information which govern systems, IEEE Transactions on Systems Man and Cybernetics, April 1976.
- [6] M. DELARCIE : Problèmes liés à la hiérarchisation et au dessin automatiques de graphes - Rapport de recherche IMAG n° 151.
- [7] W.E. DONATH & A.J. HOFFMAN : Lower bounds for the partitioning of graphs, IBM J. Res. and Dev. 17 (5) Septembre 1973.
- [8] HARARY, NORMAN & CARTWRIGHT : Introduction à la théorie des graphes orientés - Modèles structuraux, DUNOD 1974.
- [9] A. HOFFMAN & S. WINOGRAD : On finding all shortest distances in a directed network - Publication of the Maths Sciences Department IBM Th. J. Watson Research Center
- [10] J. HOPCROFT & R. TARJAN : Efficient planarity testing, Journal of the ACM oct. 1974.
- [11] LANTNER : Theorie de la dominance économique, DUNOD 1974
- [12] Ch. LAUGIER & F. MARTINEZ : Visualisation interactive de graphes. Contrat n° 74-232 DRME Déc. 1974 et Juin 1975.
- [13] A. LEDUC-LEBALLEUR : Conception et réalisation d'un logiciel graphique de base indépendant de son contexte. Application au logiciel GRIGRI - Thèse de Docteur-Ingénieur U.S.M. Grenoble, Juin 1977.

- [14] H. LEFEBVRE : l'idéologie structuraliste. LE SEUIL 1975
- [15] M.LUCAS : Contribution à l'étude des techniques de communication graphique avec un ordinateur. Eléments de base des logiciels graphiques interactifs - Thèse de Doctorat d'Etat U.S.M. Grenoble, Déc. 1977.
- [16] B. MERCIER & J.TREMOLIERES : Minimisation de calculs matriciels. Applications aux éléments finis et aux problèmes de croisements. Rapport de recherche IRIA Janvier 1973.
- [17] J.PIAGET : Le structuralisme "Que sais-je ?" P.U.F. 1973
- [18] M. RICHETIN : Analyse structurale des systèmes complexes en vue d'une commande hiérarchisées - Thèse de Doctorat d'Etat, Université P. Sabatier de Toulouse, 1975.
- [19] E. ROSSIER : La notion de cohésion et l'analyse des structures interdépendantes. Rapport de recherche Université de Genève, Mai 1978
- [20] B. ROY : Algèbre Moderne et Théorie des graphes, DUNOD 1969.
- [21] R.S. VARGA : Iterative Matrix Analysis, Prentice Hall 1963.
- [22] B. WALLISER : Systèmes et modèles - Introduction critique à l'analyse de systèmes, LE SEUIL 1977.
- [23] J.N. WARFIELD : Binary Matrices in System Modeling - IEEE Transactions on Systems Man & Cybernetics, Sept. 1973.
- [24] J.N. WARFIELD : Toward Interpretation of Complex Structural Models IEEE Transactions on Systems Man & Cybernetics Sept. 1974.
- [25] J.N. WARFIELD : Crossing Theory and Hierarchy Mapping - IEEE Transactions on systems Man & Cybernetics, July 1977
- [26] WILKINSON : The Algebraic eigenvalue problem, CLARENDON 1965.