



HAL
open science

Systemes experts à base de connaissances profondes : application à un poste de travail intelligent pour le comptable

Michel Page

► **To cite this version:**

Michel Page. Systemes experts à base de connaissances profondes : application à un poste de travail intelligent pour le comptable. Modélisation et simulation. Institut National Polytechnique de Grenoble - INPG, 1990. Français. NNT: . tel-00338750

HAL Id: tel-00338750

<https://theses.hal.science/tel-00338750>

Submitted on 14 Nov 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.

70123

THESE

présentée par Michel PAGE

pour obtenir le titre de Docteur de l'Institut National Polytechnique de Grenoble

(Spécialité Informatique)

SYSTEMES EXPERTS À BASE DE CONNAISSANCES PROFONDES APPLICATION À UN POSTE DE TRAVAIL "INTELLIGENT" POUR LE COMPTABLE

Thèse soutenue le 2 février 1990 devant la commission d'examens

Composition du jury:

M. Jacques MOSSIERE	Président du jury
Mme Suzanne PINSON	Rapporteurs
Mme Marie-Christine ROUSSET	
M. Michel GERVAIS	Examineurs
M. Gilles KASSEL	
M. Bruno OUDET	

Thèse préparée au Laboratoire de Génie Informatique

Remerciements

Je tiens à remercier :

- **Monsieur Jacques Mossière d'avoir accepté de présider ce jury.**
- **Mesdames Suzanne Pinson et Marie-Christine Rousset, rapporteurs, de l'intérêt qu'elles ont manifesté à mon travail.**
- **Monsieur Bruno Oudet, directeur de thèse, de la confiance qu'il m'a témoignée tout au long de son encadrement.**
- **Monsieur Gilles Kassel, examinateur, pour sa précieuse et généreuse contribution à la mise au point du mémoire.**
- **Monsieur Michel Gervais, examinateur, pour sa participation à ce jury et ses commentaires avisés.**
- **Monsieur Pierre Cassassuce, enseignant à l'UJF, d'avoir coopéré sur le projet d'analyse de surplus.**
- **Messieurs Andreï Tarkovski, Robert Bresson, Wim Wenders pour le soutien moral qu'ils m'ont fourni dans les moments difficiles.**

Résumé

La plupart des systèmes experts actuels reposent sur les connaissances "de surface" (le savoir-faire) d'un expert du domaine d'application. Plus récemment, une autre approche s'est développée. Elle vise à exploiter les connaissances "profondes" (théoriques) acquises dans le domaine d'application. La thèse étudie cette dernière approche dans le cadre du projet PIC (Poste de travail Intelligent pour le Comptable).

Les aspects méthodologiques sont développés dans la première partie. Une nouvelle classe d'applications des systèmes experts est proposée: l'interprétation comparative. Elle a pour but de mettre en évidence et expliquer la cause des différences entre deux états d'un système. Une méthode générale permettant d'aborder ce problème est présentée, ainsi que des techniques la mettant en oeuvre sur des modèles qualitatifs et numériques.

Les contributions au projet PIC sont développées dans la seconde partie. Un générateur de systèmes experts d'interprétation comparative est d'abord présenté. Il a servi à la réalisation de deux systèmes: le premier pour l'analyse de la performance d'une entreprise par la méthode des surplus, le second pour le diagnostic financier d'entreprise. Un système expert pour la déduction des écritures comptables utilisant également l'approche profonde est ensuite présenté. A la lumière de ces deux dernières applications déjà abordées par les systèmes experts utilisant des connaissances de surface, les deux approches de conception de systèmes experts sont comparées.

MOTS CLES: intelligence artificielle, systèmes experts, raisonnement qualitatif, modélisation, gestion, comptabilité, méthode des surplus, analyse financière, imputation comptable.



INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

46 avenue Felix Viallet
38031 GRENOBLE cedex

Tél. : 76.57.45.00

Année universitaire 1989

Président de l'Institut :
Monsieur Georges LESPINARD

Professeurs des Universités

BARIBAUD Michel	ENSERG	JAUSSAUD Pierre	ENSIEG
BARRAUD Alain	ENSIEG	JOST Rémy	ENSPG
BAUDELET Bernard	ENSPG	JOUBERT Jean-Claude	ENSPG
BEAUFILS Jean-Pierre	INPG	JOURDAIN Geneviève	ENSIEG
BLIMAN Samuel	ENSERG	LACOUME Jean-Louis	ENSIEG
BOIS Philippe	ENSHMG	LADET Pierre	ENSIEG
BONNETAIN Lucien	ENSEEG	LESIEUR Marcel	ENSHMG
BONNET Guy	ENSPG	LESPINARD Georges	ENSHMG
BRISSONNEAU Pierre	ENSIEG	LONGEQUEUE Jean-Pierre	ENSPG
BRUNET Yves	IUFA	LORET Benjamin	ENSHMG
CAILLERIE Denis	ENSHMG	LOUCHET François	ENSEEG
CAVAIGNAC Jean-François	ENSPG	LUCAZEAU Guy	ENSEEG
CHARTIER Germain	ENSPG	MASSE Philippe	ENSIEG
CHENEVIER Pierre	ENSERG	MASSELOT Christian	ENSIEG
CHERADAME Hervé	UFR PGP	MAZARE Guy	ENSIMAG
CHERUY Arlette	ENSIEG	MOHR Roger	ENSIMAG
CHOVET Alain	ENSERG	MOREAU René	ENSHMG
COHEN Joseph	ENSERG	MORET Roger	ENSIEG
COLINET Catherine	ENSEEG	MOSSIERE Jacques	ENSIMAG
CORNUT Bruno	ENSIEG	OBLED Charles	ENSHMG
COULOMB Jean-Louis	ENSIEG	OZIL Patrick	ENSEEG
COUMES André	ENSERG	PA ULEAU Yves	ENSEEG
CROWLEY James	ENSIMAG	PERRET Robert	ENSIEG
DARVE Félix	ENSHMG	PIAU Jean-Michel	ENSHMG
DELLA-DORA Jean	ENSIMAG	PIC Etienne	ENSERG
DEPEY Maurice	ENSERG	PLATEAU Brigitte	ENSIMAG
DEPORTES Jacques	ENSPG	POUPOT Christian	ENSERG
DEROO Daniel	ENSEEG	RAMEAU Jean-Jacques	ENSEEG
DESRE Pierre	ENSEEG	REINISCH Raymond	ENSPG
DOLMAZON Jean-Marc	ENSERG	RENAUD Maurice	UFR PGP
DURAND Francis	ENSEEG	ROBERT André	UFR PGP
DURAND Jean-Louis	ENSPG	ROBERT François	ENSIMAG
FAUTRELLE Yves	ENSHMG	SABONNADIÈRE Jean-Claude	ENSIEG
FOGGIA Albert	ENSIEG	SAUCIER Gabrièle	ENSIMAG
FONLUPT Jean	ENSIMAG	SCHLENKER Claire	ENSPG
FOULARD Claude	ENSIEG	SCHLENKER Michel	ENSPG
GANDINI Alessandro	UFR PGP	SERMET Pierre	ENSERG
GAUBERT Claude	ENSPG	SILVY Jacques	UFR PGP
GENTIL Pierre	ENSERG	SIRIEYS Pierre	ENSHMG
GENTIL Sylviane	ENSIEG	SOHM Jean-Claude	ENSEEG
GREVEN Hélène	IUFA	SOLER Jean-Louis	ENSIMAG
GUEGUEN Claude	ENSIEG	SOUQUET Jean-Louis	ENSEEG
GUERIN Bernard	ENSERG	TROMPETTE Philippe	ENSHMG
GUYOT Pierre	ENSEEG	VINCENT Henri	ENSPG
IVANES Marcel	ENSIEG	ZADWORNÝ François	ENSERG

Personnes ayant obtenu le diplôme d'HABILITATION A DIRIGER DES RECHERCHES

BECKER Monique
BINDER Zdenek
CHASSERY Jean-Marc
CHOLLET Jean-Pierre
COEY John
COLINET Catherine
COMMAULT Christian
CORNUJOLS Gérard
COULOMB Jean- Louis
COURNIL M.
DALARD Francis
DANES Florin
DEROO Daniel
DIARD Jean-Paul
DION Jean-Michel
DUGARD Luc
DURAND Madeleine
DURAND Robert
GALERIE Alain
GAUTHIER Jean-Paul
GENTIL Sylviane

GHIBAUDO Gérard
HAMAR Sylvaine
HAMAR Roger
LACHENAL D.
LADET Pierre
LATOMBE Claudine
LE HUY H.
LE GORREC Bernard
MADAR Roland
MEUNIER G.
MULLER Jean
NGUYEN TRONG Bernadette
NIEZ J.J.
PASTUREL Alain
PLA Fernand
ROGNON J.P.
ROUGER Jean
TCHUENTE Maurice
VINCENT Henri
YAVARI A.R.

Chercheurs du C.N.R.S

DIRECTEURS DE RECHERCHE CLASSE 0

LANDEAU	Ioan
NAYROLLES	Bernard

Directeurs de recherche 1ère Classe

ANSARA Ibrahim
CARRE René
FRUCHART Robert
HOPFINGER Emile

JORRAND Philippe
KRAKOWIAK Sacha
LEPROVOST Christian
VACHAUD Georges
VERJUS Jean-Pierre

Directeurs de recherche 2ème Classe

ALEMANY Antoine
ALLIBERT Colette
ALLIBERT Michel
ARMAND Michel
AUDIER Marc
BERNARD Claude
BINDER Gilbert
BONNET Roland
BORNARD Guy
CAILLET Marcel
CALMET Jacques
CHATILLON Chrétien
CLERMONT Jean-Robert
COURTOIS Bernard
DAVID René
DION Jean-Michel
DRIOLE Jean
DURAND Robert
ESCUDIER Pierre
EUSTATHOPOULOS Nicolas
GARNIER Marcel
GUELIN Pierre

JOUD Jean-Charles
KAMARINOS Georges
KLEITZ Michel
KOFMAN Walter
LEJEUNE Gérard
MADAR Roland
MERMET Jean
MICHEL Jean-Marie
MEUNIER Jacques
PEUZIN Jean-Claude
PIAU Monique
RENOUARD Dominique
SENATEUR Jean-Pierre
SIFAKIS Joseph
SIMON Jean-Paul
SUERY Michel
TEODOSIU Christian
VAUCLIN Michel
VENNEREAU Pierre
WACK Bernard
YONNET Jean-Paul

**Personnalités agréées à titre permanent à diriger
des travaux de recherche
(décision du conseil scientifique)**

E.N.S.E.E.G

HAMMOU Abdelkader
MARTIN-GARIN Régina
SARRAZIN Pierre
SIMON Jean Paul

E.N.S.E.R.G

BORRI Joseph

E.N.S.I.E.G

DESCHIZEAUX Pierre
GLANGEAUD François
PERARD Jacques
REINISCH Raymond

E.N.S.H.M.G

ROWE Alain

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

COURTIN Jacques

C.E.N.G

CADET Jean
COEURE Philippe
DELHAYE Jean-Marc
DUPUY Michel
JOUVE Hubert
NICOLAU Yvan
NIFENECKER Hervé
PERROUD Paul
PEUZIN Jean-Claude
TAIEB Maurice
VINCENDON Marc

Laboratoires extérieurs :

C.N.E.T

DEVINE Rodericq
GERBER Roland
MERCKEL Gérard
PAULEAU Yves

Situation particulière

PROFESSEURS D'UNIVERSITE

DETACHEMENT

ENSIMAG	LATOMBE	J..Claude	Détachement	21/10/1989
ENSHMG	PIERRARD	J.Marie	Détachement	30/04/1989
ENSIMAG	VEILLON	Gérard	Détachement	30/09/1990
ENSIMAG	VERJUS	J.Pierre	Détachement	30/09/1989
ENSPG	BLOCH	Daniel	Recteur à c/	21/12/1988

SURNOMBRE

INPG	CHIAVERINA	Jean	30/09/1989
ENSHMG	BOUVARD	Maurice	30/09/1991
ENSEEG	PARIAUD	J.Charles	30/09/1991

TABLE DES MATIERES

Introduction.....	1
PREMIERE PARTIE.....	5
Chapitre 1 - Les Systèmes experts de première génération.....	7
1 - Les acquis des systèmes experts de première génération.....	7
1.1 - Limitation du domaine d'application.....	8
1.2 - Principe architectural.....	8
1.3 - Représentation des connaissances.....	9
2 - Les limites des systèmes experts de première génération.....	9
2.1 - La construction de la base de connaissances.....	10
2.2 - La maintenance de la base de connaissances.....	12
2.3 - La rigidité de la base de connaissances.....	13
2.4 - La faiblesse de l'explication du raisonnement.....	14
Chapitre 2 - Les Systèmes experts à base de connaissances profondes...17	
1 - Le raisonnement sur les graphes de causalité.....	18
1.1 - Représentation des connaissances par un graphe de causalité.....	19
1.2 - Raisonnement dans un graphe de causalité.....	21
2 - Le raisonnement qualitatif.....	22
2.1 - Les modélisations proposées pour le raisonnement qualitatif.....	23

2.2 - Les applications du raisonnement qualitatif	28
Chapitre 3 - Interprétation comparative	35
1 - Interprétation comparative	36
1.1 - Modèle	36
1.2 - Interprétation comparative dans un modèle.....	38
1.3 - Applications.....	38
2 - Analyse des problèmes d'interprétation comparative.....	39
2.1 - L'évaluation comparative.....	40
2.2 - Explication comparative	42
Chapitre 4 - Explication comparative dans un modèle qualitatif.....	51
1 - Explication comparative dans un réseau d'influences	52
2 - La construction du graphe causal	53
3 - Construction de la base d'explication complète minimale.....	56
4 - Présentation de la base d'explication	61
Chapitre 5 - Explication comparative dans un modèle numérique.....	67
1 - Explication comparative dans un système d'équations.....	68
1.1 - Système d'équations contraint	68
1.2 - Système d'équations structurel	70
2 - Construction du graphe causal	73
2.1 - Recherche des variables de sortie des équations.....	73

2.2 - Construction du graphe réduit	75
3 - Construction de la base d'explication	77
4 - Présentation de la base d'explication	78
4.1 - Cas des variables numériques.....	79
4.2 - Cas des variables qualitatives	82
5 - Comparaison avec d'autres travaux	85
Conclusion de la première partie.....	87
DEUXIEME PARTIE.....	89
Introduction.....	89
Chapitre 6 - Deux applications de l'interp. comparative au projet PIC...91	
1 - CRESUS: un générateur de systèmes d'interprétation comparative	92
1.1 - Les interfaces de CRESUS.....	93
1.2 - Le noyau de CRESUS.....	94
2 - ANAFI: une maquette pour le diagnostic financier	96
2.1 - Le diagnostic financier.....	96
2.2 - Les systèmes experts et l'analyse financière	100
2.3 - Diagnostic financier et interprétation comparative dans ANAFI.....	101
2.4 - Résultats, limitations et perspectives.....	107
3 - SURPLUS: un système expert pour l'analyse de la performance d'une entreprise..	108
3.1 - La méthode des surplus.....	109
3.2 - La méthode des surplus et l'interprétation comparative	112

Table des matières

3.3 - Implantation.....	115
3.4 - Résultats et extensions.....	115
Chapitre 7 - SICA: un système expert pour l'imputation comptable....	119
1 - L'imputation comptable.....	120
2 - Systèmes experts et imputation comptable.....	121
2.1 - Le prototype de Stepniewski et Gervais	121
2.2 - Le système KIWI de Clarke.....	122
3 - Présentation de SICA	123
3.1 - La description des opérations comme échanges de flux	124
3.2 - De la description de l'échange à l'écriture comptable	126
3.3 - De la phrase à la description de l'échange	127
3.4 - Résultats et perspectives.....	140
Conclusion.....	141
ANNEXE A - Un exemple de résultat fourni par le système ANAFI....	155
ANNEXE B - Un exemple de résultat fourni par le système SURPLUS	157
ANNEXE C - Le modèle logique de SICA	161

INTRODUCTION

Selon une enquête récente ([LMI 88]), le nombre de systèmes experts opérationnels (c'est-à-dire régulièrement utilisés pour résoudre le problème auquel ils sont destinés) en France ne dépasserait pas quelques dizaines. Ce chiffre, étonnamment faible, met en évidence les problèmes que pose encore aujourd'hui la réalisation d'un projet dans ce domaine. Ces difficultés allant de pair avec le perfectionnement des outils de développement, il y a lieu de penser que c'est la **méthodologie** de conception qu'il faut remettre en cause. La plupart des systèmes experts actuels, dits de première génération, sont en effet réalisés avec la même démarche, fondée sur l'exploitation des seules connaissances **pratiques** (souvent dénommées connaissances de "surface") d'un expert du domaine d'application. Or, les limites de cette approche sont aujourd'hui connues ([DAVIS 82], [CLANCEY 83]): difficulté de construction et de maintenance d'une base de connaissances, faible réutilisabilité d'une base de connaissances pour résoudre un problème légèrement différent du problème initial, insuffisance de l'explication du raisonnement suivi par le système.

Face à ces problèmes, une seconde approche de conception des systèmes experts a été proposée plus récemment: [HART 82], [MICHIE 82]. Elle tente d'exploiter les connaissances **théoriques** (ou connaissances "profondes") du domaine d'application. Contrairement à la précédente, cette approche fait encore largement partie de la recherche, mais semble très prometteuse dans les domaines atteignant un degré de formalisation important. Son rapide développement annonce la seconde génération de systèmes experts ([STEELS 85]), capables de faire coopérer connaissances de surface et profondes pour résoudre un problème. Cependant, bien des aspects de la conception des systèmes experts utilisant des connaissances profondes restent à explorer avant qu'une telle coopération soit réellement possible.

La Gestion constitue un champ d'investigation privilégié dans cette optique. Les efforts de modélisation dans ce domaine rendent en effet possible la réalisation de systèmes experts basés sur des connaissances profondes. Le projet PIC (Poste de travail Intelligent pour le Comptable) est le résultat de cette constatation. Développé dans l'Equipe SIAD du Laboratoire de Génie Informatique à Grenoble, ce projet a pour objectif d'automatiser la chaîne des traitements comptables (figure 0.1) depuis la déduction des écritures (imputation) jusqu'à l'analyse financière et économique des documents de synthèse publiés par une entreprise. Ce but est très ambitieux car l'informatique "traditionnelle" ne permet de traiter que la tenue des comptes; les autres composantes du projet ne sont pas de nature algorithmique.

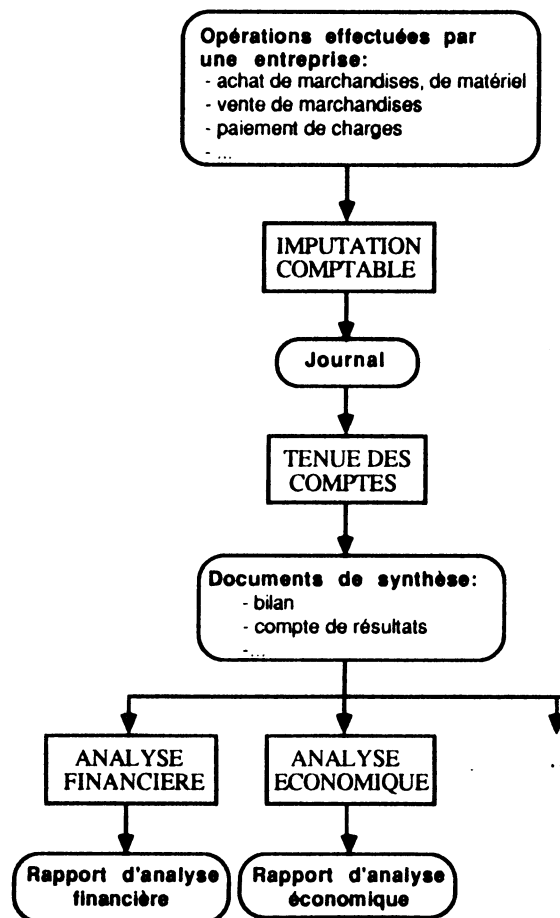


figure 0.1: les traitements comptables de la déduction des écritures (imputation) jusqu'à l'analyse des documents

Des systèmes experts de première génération ont déjà été développés, tant pour l'imputation comptable ([GERVAIS 86]) que l'analyse financière des documents ([MICHA 86], par exemple). Ces systèmes aboutissent cependant à des solutions peu satisfaisantes. Ils tentent de simuler le raisonnement de surface de l'expert, ce qui conduit souvent à l'écriture de bases de connaissances très lourdes envisageant tous les cas pouvant se présenter. Ces systèmes ne tirent aucun bénéfice des modèles développés dans leur domaine d'application. Pourtant, des recherches théoriques menées en comptabilité ont abouti à des propositions de formalisation axiomatique: [IJIRI 64]. De même, l'étude du fonctionnement d'une entreprise, qui constitue une partie importante de l'analyse financière, fait appel à un raisonnement qualitatif sur des équations comptables: [HART 86]. Il semble naturel d'utiliser ces connaissances profondes dans un système expert. Cette idée constitue le point de départ de la thèse.

Deux contributions se dégagent du travail présenté sur le plan méthodologique. Premièrement, une nouvelle classe d'applications des systèmes experts est proposée: l'**interprétation comparative**. Elle a pour but d'analyser et expliquer les différences entre deux états d'un système en utilisant des connaissances profondes sur le fonctionnement de ce système. L'objectif est la génération d'un rapport en français synthétisant les différences entre les deux états comparés et expliquant les facteurs qui ont causé ces différences. Une méthode générale est proposée pour la réalisation de systèmes experts traitant ce problème. Développée dans le cadre de modèles qualitatifs et numériques, elle est implantée dans CRESUS, un générateur de systèmes d'interprétation comparative. Deuxièmement, les deux approches de conception de systèmes experts sont comparées: l'imputation comptable et l'analyse financière, déjà abordées par des systèmes experts de première génération et traitées dans la thèse en utilisant des connaissances profondes, fournissent une base intéressante pour cette comparaison.

Sur le plan des applications, trois composantes du projet PIC ont été réalisées. La première, dénommée SICA (Système d'Imputation Comptable Automatique), est un système expert situé en amont de la tenue des comptes. Il analyse des phrases décrivant des pièces comptables et détermine les écritures à passer. Au lieu de déduire directement l'imputation à partir de la description d'une pièce comme le font les systèmes experts utilisant des connaissances de surface, l'opération économique sous-jacente est d'abord reconstituée. Celle-ci, décrite sous forme d'un échange de flux entre l'entreprise et un tiers, constitue un véritable modèle du fonctionnement comptable de l'entreprise. L'imputation s'en déduit simplement au moyen de règles logiques.

Les deux autres applications du projet PIC utilisent la notion d'interprétation comparative. Le système ANAFI aborde l'analyse financière. Ce problème consiste à comparer la situation d'une

entreprise par rapport à son secteur d'activité d'une part, et à une situation passée, d'autre part. SURPLUS, un système d'analyse des performances d'une entreprise par la méthode des surplus est basé sur le même principe. L'interprétation comparative est cette fois utilisée pour expliquer l'augmentation ou la baisse de productivité de l'entreprise. Ces deux applications ont été réalisées à l'aide de CRESUS, le générateur de systèmes d'interprétation comparative.

Le mémoire comprend deux parties développant les aspects méthodologiques puis pratiques du travail présenté. Dans la première partie, les caractéristiques fondamentales et les limites des systèmes experts utilisant des connaissances de surface sont d'abord examinées dans le chapitre 1. Les différents thèmes de recherche sur les systèmes experts utilisant des connaissances profondes sont ensuite présentés dans le chapitre 2. Le chapitre 3 introduit la notion d'interprétation comparative, qui est ensuite développée pour des modèles qualitatifs dans le chapitre 4 puis des modèles numériques dans le chapitre 5.

La seconde partie décrit les applications réalisées dans le cadre du projet PIC. CRESUS, le générateur de systèmes d'interprétation comparative puis les deux applications qu'il a permis de réaliser: ANAFI pour l'analyse financière et SURPLUS pour l'analyse des performances par la méthode des surplus, sont présentés dans le chapitre 6. SICA, le système expert pour l'imputation comptable fait l'objet du chapitre 7. Enfin, les deux approches de conception de systèmes experts sont comparées en conclusion, à la lumière de ces réalisations.

PREMIERE PARTIE

CHAPITRE 1

LES SYSTEMES EXPERTS DE PREMIERE GÉNÉRATION

La principale caractéristique des systèmes experts de première génération, illustrée par le légendaire MYCIN ([SHORTLIFFE 76]), est la priorité absolue accordée aux connaissances **pratiques**. Dans cette approche, un système expert est réalisé avec le concours d'un praticien du domaine d'application, un expert. Les concepts utilisés par ce dernier, ses raisonnements et même ses incertitudes au cours de la résolution d'un problème sont codés dans le système. Le mimétisme du savoir-faire de l'expert assure ainsi les **performances** du système. Cette démarche a largement fait ses preuves puisqu'elle utilisée pour la conception de la plupart des systèmes experts actuels. Cependant, le développement d'applications de plus en plus complexes en a également montré les limites, motivant de nouvelles recherches.

Cette thèse concernant l'utilisation des connaissances profondes dans les systèmes experts, il est hors de notre propos de présenter dans ce chapitre un état de l'art des systèmes experts de première génération. On pourra pour cela consulter [HAYES-ROTH 83] ou [WATERMAN 86]. Nous nous bornons à en résumer les principaux acquis dans la section 1, avant d'analyser en détail les difficultés qui ont conduit au développement des systèmes experts à base de connaissances profondes en section 2.

1 - Les acquis des systèmes experts de première génération

Trois points fondamentaux émergent des systèmes experts de première génération: la nécessité de délimiter le domaine d'application, la séparation des connaissances et du mécanisme qui les utilise et

enfin les propriétés souhaitables d'un langage de représentation des connaissances. Ces trois points sont succinctement développés.

1.1 - Limitation du domaine d'application

Cerner le domaine d'application pour obtenir un système performant est le premier principe utilisé pour la conception des systèmes experts de première génération. Cela signifie non seulement limiter la taille du domaine abordé, mais aussi la nature des traitements effectués. Ainsi, le domaine de prédilection de ces systèmes est la **classification heuristique** ([CLANCEY 85]), tâche consistant à choisir parmi un ensemble de solutions potentielles celle(s) qui s'adapte(nt) le mieux au problème considéré.

1.2 - Principe architectural

Le second point mis en évidence est la séparation des connaissances et de leur mécanisme d'exploitation. La figure 1.1 représentant l'architecture classique d'un système expert de première génération illustre ce point. Les connaissances fournies par l'expert sont stockées dans la base de connaissances et sont utilisées par un programme indépendant, le moteur d'inférence. Ainsi, une modification de la base n'entraîne pas celle du mécanisme d'exploitation.

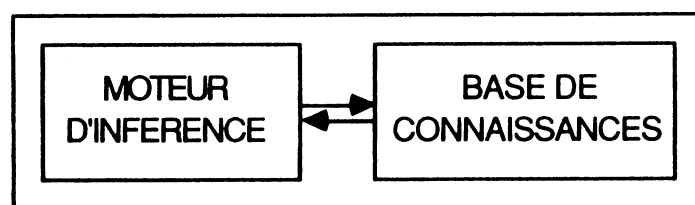


Figure 1.1: l'architecture d'un système expert de première génération

1.3 - Représentation des connaissances

Différents langages de représentation de connaissances ont été expérimentés dans les systèmes experts de première génération. Deux propriétés émergent de ces expériences: la **granularité** et l'**uniformité**.

Un langage de représentation est granulaire lorsqu'il permet de décomposer le savoir-faire de l'expert en une collection de connaissances élémentaires, chacune indépendante des autres. La modification d'une connaissance est ainsi sans répercussion sur le reste de la base. Ce point revêt une importance particulière lorsque la taille de la base de connaissances devient importante.

L'uniformité vise à représenter toutes les connaissances de l'expert dans un même formalisme. Cette propriété découle du souhait de garder le mécanisme d'exploitation des connaissances aussi simple que possible. L'utilisation de plusieurs langages de représentation multiplie en effet les traitements effectués par le moteur d'inférence et les interactions entre les diverses représentations. Ceci explique que jusqu'à une date récente, la plupart des systèmes experts utilisaient un seul formalisme, le plus souvent des règles de production (c'est-à-dire des expressions: si <CONDITIONS> alors <CONCLUSIONS>).

2 - Les limites des systèmes experts de première génération

Si la confrontation à des applications de plus en plus complexes a permis le développement des systèmes experts de première génération, elle en a également rapidement montré les limites ([DAVIS 82], [CLANCEY 83]): difficulté de la construction et de la maintenance de la base de connaissances, dégradation brutale des performances du système aux frontières de sa base de connaissances, faiblesse de l'explication du raisonnement. Ces difficultés s'illustrent parfaitement sur un exemple simple tiré d'une application développée dans la seconde partie de la thèse: le diagnostic financier.

Exemple 1

Au cours d'un diagnostic financier, l'expert calcule puis analyse un certain nombre de ratios de l'entreprise analysée. Ainsi, l'augmentation du ratio de rentabilité des capitaux, défini par l'expression résultat net / capitaux permanents, constitue en principe un indice favorable de l'évolution de la rentabilité. Cependant, causée par une diminution des capitaux, elle peut signifier que l'entreprise liquide ses immobilisations, auquel cas la santé de l'entreprise est peut-être en

danger. L'expert raisonne qualitativement sur l'équation de définition du ratio afin d'examiner laquelle de ces deux hypothèses est vérifiée. Les paragraphes suivants montrent les difficultés que pose la réalisation d'un système expert de première génération effectuant le même raisonnement.

2.1 - La construction de la base de connaissances

La construction d'une base de connaissances comporte deux aspects: l'acquisition des connaissances nécessaires à la résolution du problème et leur codage dans la base. Deux spécialistes interviennent dans ce processus (figure 1.2): le premier, expert du domaine d'application, fournit les connaissances nécessaires à la réalisation du système (en réalité, il arrive fréquemment que plusieurs experts participent à ce travail) ; le second, en Intelligence Artificielle, extrait le savoir-faire de l'expert et le code dans la base.

Le codage et surtout l'acquisition des connaissances sont si complexes que la réalisation d'un système expert se prolonge généralement sur plusieurs mois voire plusieurs années, malgré le perfectionnement des outils informatiques (éditeurs de connaissances graphiques, ...). Ceci engendre évidemment un coût important, d'autant que deux spécialistes (au moins) interviennent dans le processus.



Figure 1.2: la construction de la base de connaissances

Reconnue dès 1977 comme le "goulot d'étranglement" des systèmes experts ([FEIGENBAUM 77]), l'acquisition des connaissances est une tâche difficile. D'une part, l'expert a souvent des difficultés à expliciter son savoir-faire, ce qui entraîne de fréquents retours en arrière et modifications. D'autre part, le travail du spécialiste en Intelligence Artificielle est rendue complexe par l'absence d'une méthode d'analyse fiable comparable à celles utilisées en informatique traditionnelle.

En dehors des nombreux efforts pour développer des méthodes d'acquisition des connaissances, deux solutions ont été envisagées pour faciliter l'acquisition des connaissances dans les systèmes experts

de première génération: l'apprentissage automatique et la construction de systèmes experts à partir de documents.

L'apprentissage vise à éliminer le spécialiste en Intelligence Artificielle du processus d'acquisition. Celui-ci est remplacé par un programme qui construit automatiquement la base de connaissances à partir de cas résolus par l'expert (par exemple). L'apprentissage est une discipline en plein essor, mais encore insuffisamment avancée pour être utilisée systématiquement.

A l'inverse de l'apprentissage, la construction de systèmes experts à partir de documents élimine l'expert. Les connaissances requises pour l'expertise sont extraites d'un manuel du domaine d'application et sont codées par le spécialiste en Intelligence Artificielle dans la base de connaissances. Cette méthode, utilisée par exemple dans le système CASTER ([CLANCEY 87]), pose cependant de nouveaux problèmes. Même les manuels très spécialisés possèdent rarement la précision nécessaire à la réalisation d'un système expert. Par ailleurs, l'utilisation d'ouvrages plus théoriques requiert le raisonnement sur des modèles, impossible avec les systèmes experts de première génération.

Au problème de l'acquisition des connaissances s'ajoute la difficulté de leur codage dans la base. L'uniformité de représentation des connaissances caractérisant les systèmes experts de première génération oblige en effet le spécialiste en Intelligence Artificielle à couler toutes les connaissances de l'expert dans un même moule. Ceci peut conduire à l'écriture d'une base de connaissances complexe, même pour un problème simple, comme nous allons le voir sur l'exemple 1.

Imaginons que l'on représente les connaissances exprimées dans l'exemple 1 par des règles. Ceci peut être réalisé par le jeu de règles suivant (formalisées ici en logique des propositions, mais la logique utilisée ne change rien au problème):

r1: si rentabilité-capitaux-augmente
et résultat-net-augmente
et capitaux-permanents-augmente
alors augmentation-rentabilité-capitaux-causée-par-résultat-net

r2: si rentabilité-capitaux-augmente
et résultat-net-augmente
et capitaux-permanents-stable
alors augmentation-rentabilité-capitaux-causée-par-résultat-net

r3: si rentabilité-capitaux-augmente
et résultat-net-augmente
et capitaux-permanents-diminue
alors augmentation-rentabilité-capitaux-causée-par-résultat-net-et-capitaux-permanents

etc...

En considérant trois valeurs (augmentation, stabilité, diminution) pour qualifier l'évolution d'une variable, 13 règles sont nécessaires pour analyser la cause de l'évolution du ratio de rentabilité des capitaux, compte tenu des contraintes liant l'évolution des variables. D'autres ratios intervenant dans l'appréciation de la situation d'une entreprise sont définis par des équations où apparaissent plus de deux variables; l'utilisation du même procédé devient alors extrêmement lourde, et conduit à l'écriture d'une base de connaissances complexe.

Pour faciliter le codage des connaissances, il est nécessaire de remettre en cause l'uniformité de représentation des connaissances. Ainsi, lorsqu'il existe un modèle du domaine d'application comme dans l'exemple 1, celui-ci devrait pouvoir être décrit explicitement dans le système. Cependant, la capacité de raisonner sur ce modèle conduit hors du cadre des systèmes experts de première génération.

2.2 - La maintenance de la base de connaissances

Le second problème se posant dans les systèmes experts de première génération est la maintenance de la base de connaissances. Plusieurs facteurs rendent cette dernière nécessaire: l'évolution des connaissances du domaine d'application, une base de connaissances incomplète ou incorrecte. Cette tâche est complexe car la granularité des connaissances et leur indépendance par rapport au moteur d'inférence sont rarement complètes dans la pratique. Ainsi, l'ordre d'évaluation des termes-conditions d'une règle peut masquer ou entraîner le déclenchement d'autres règles. La modification, l'ajout ou la suppression d'une règle sont donc susceptibles d'avoir des répercussions sur toute la base de connaissances, et en particulier d'y introduire des incohérences.

Deux solutions ont été proposées pour remédier à ce problème: le contrôle de cohérence et le contrôle de validité. Le premier tente de vérifier qu'une base de connaissances ne comporte pas de contradiction logique; le second qu'elle correspond bien au savoir-faire de l'expert. Ces deux propriétés sont théoriquement vérifiables en s'assurant que la saturation de toutes les bases de faits cohérentes (resp. valides) par application de la base de règles ne produit que des bases de faits cohérentes (resp. valides). Malheureusement, l'explosion combinatoire interdit la mise en oeuvre d'une telle méthode dès que la taille de la base de connaissances devient importante. Des solutions sont néanmoins en cours de développement au prix de restrictions: sur le langage d'expression des connaissances ([ROUSSET 87]),

focalisation de la recherche des incohérences sur des types d'erreurs particuliers ([NGUYEN 85]) ou sur certaines parties de la base de connaissances ([AYEL 86]).

La vérification de la cohérence et de la validité d'une base de connaissances sont nécessaires pour assurer la maintenance d'un système expert. Elle ne constitue cependant qu'un palliatif dans l'exemple 1: l'équation de définition du ratio est, elle même, cohérente et valide. C'est la multiplication des règles nécessaires pour représenter un raisonnement sur cette équation qui introduit le risque d'incohérence et d'invalidité. Dans de tels cas, le véritable problème se situe donc ailleurs: comment concevoir des systèmes experts autorisant la représentation explicite de connaissances profondes (des équations, par exemple) et le raisonnement sur ces connaissances ?

2.3 - La rigidité de la base de connaissances

La troisième difficulté liée aux systèmes experts de première génération est encore une conséquence de la représentation uniforme des connaissances. Celle-ci confère à la base de connaissances une grande rigidité ([DAVIS 82]). Toutes les connaissances de l'expert devant être représentées dans un formalisme unique, la base de connaissances devient vouée à un raisonnement particulier, alors que les connaissances initiales de l'expert ne l'étaient pas. Il en résulte l'impossibilité d'exploiter une base de connaissances pour un problème légèrement différent de celui pour laquelle elle a été initialement conçue, comme le montre l'exemple ci-dessous.

Exemple 2:

Imaginons qu'après le diagnostic, l'analyste financier doive proposer un pronostic sur la situation de l'entreprise. Supposons qu'il dispose pour cela d'indices sur l'évolution qualitative du résultat net et des capitaux permanents, et qu'il tente de projeter l'évolution de la rentabilité des capitaux. De nouveau, ceci s'effectue par un raisonnement qualitatif sur l'équation de définition du ratio de rentabilité des capitaux défini dans l'exemple 1. Cependant, les règles écrites pour l'exemple 1 ne s'appliquent pas à ce problème, et il est nécessaire d'écrire un autre jeu de règles:

r1': si résultat-net-augmente
et capitaux-permanents-augmente
alors évolution-rentabilité-capitaux-indéterminée

r2': si résultat-net-augmente
et capitaux-permanents-stable
alors rentabilité-capitaux-augmente

r3': si résultat-net-augmente

et capitaux-permanents-diminue
alors rentabilité-capitaux-augmente

etc...

La solution consisterait une nouvelle fois à représenter explicitement les connaissances profondes et à modéliser les différents raisonnements qui leur sont applicables.

2.4 - La faiblesse de l'explication du raisonnement

L'explication du raisonnement souffre du même problème. La possibilité d'expliquer à l'utilisateur le raisonnement suivi par un système expert est pourtant fréquemment mise en avant dans la littérature. Mais il ne s'agit le plus souvent que d'une trace des règles exécutées. Ces informations sont certes utiles lors du développement du système pour contrôler la validité des inférences effectuées. Leur insuffisance pour l'utilisateur final du système expert a cependant été objectée ([CLANCEY 83], [ALTY 87]). D'une part, les règles fournies par l'expert ne sont pas nécessairement compréhensibles pour une personne étrangère au domaine d'application. D'autre part, à cause de la représentation uniforme des connaissances, des éléments essentiels peuvent se trouver éliminés de l'explication. Le dialogue imaginaire ci-dessous entre un système expert utilisant les règles de l'exemple 1 et un utilisateur illustre ce point.

Supposons que le système expert ait déjà établi les faits: résultat-net-augmente et rentabilité-capitaux-augmente.

(1) le système: quelle est l'évolution des capitaux permanents (augmente / stable / diminue) ?

(2) l'utilisateur: pourquoi ?

(3) le système: il est établi que résultat-net-augmente et rentabilité-capitaux-augmente. Donc si capitaux-permanents-augmente, en utilisant la règle r1, je déduirai le fait: augmentation-rentabilité-capitaux-causée-par-résultat-net

(4) l'utilisateur: stable

(5) le système: je déduis que augmentation-rentabilité-capitaux-causée-par-résultat-net

(6) l'utilisateur: comment ?

(7) le système: il est établi que:
- résultat-net-augmente

- rentabilité-capitaux-augmente
- capitaux-permanents-stable

En utilisant la règle r2, j'ai déduit le fait: augmentation-rentabilité-capitaux-causée-par-résultat-net

...

Cet exemple montre les deux possibilités d'explications les plus courantes dans les systèmes experts de première génération: la réponse à des questions du type "**pourquoi** me posez-vous cette question ?" (2 et 3) et celle à des questions du type "**comment** avez-vous obtenu ce fait ?" (6 et 7). Dans les deux cas, il manque un élément essentiel à l'explication fournie par le système expert: l'existence de la relation définissant le ratio rentabilité des capitaux. Or, cet élément ne peut pas figurer dans l'explication, car il est représenté implicitement dans les règles. De même, le système serait incapable d'expliquer la stratégie utilisée par l'expert, à savoir: pour analyser l'évolution d'un ratio, examiner l'évolution des variables de son équation de définition, puisque les équations ne sont pas représentées explicitement dans les règles.

Pour ce problème comme pour les précédents, la solution passe par la remise en cause de deux des hypothèses à la base des systèmes experts de première génération: l'utilisation des seules connaissances pratiques de l'expert et l'uniformité de la représentation des connaissances. Outre les connaissances pratiques, un système expert devrait pouvoir également utiliser les connaissances théoriques acquises dans le domaine d'application. De plus, ces connaissances devraient être représentées explicitement de sorte que différents raisonnements puissent leur être appliqués. C'est précisément la démarche adoptée pour la réalisation des systèmes experts à base de connaissances profondes.

CHAPITRE 2

LES SYSTEMES EXPERTS À BASE DE CONNAISSANCES PROFONDES

Face aux problèmes rencontrés avec les systèmes experts de première génération, une autre direction de recherche s'est développée. Cette dernière tente d'exploiter, non pas les connaissances de surface d'un expert, mais les connaissances **profondes** constituant la théorie du domaine d'application. Ces recherches trouvent leur origine en 1977 dans les premiers travaux concernant la mécanisation du raisonnement sur le fonctionnement de dispositifs physiques: [RIEGER 77]. Ils ont pris un essor important à partir des travaux sur le raisonnement qualitatif ([BOBROW 84]).

L'idée de raisonner sur les connaissances théoriques d'un domaine est très séduisante; elle semble la solution à la plupart des difficultés rencontrées avec les systèmes experts de première génération. Non confinées dans le cerveau de quelques experts, ces connaissances sont beaucoup plus largement diffusées, ce qui rend leur acquisition plus facile. Reposant sur des fondements théoriques et non sur des heuristiques incertaines, leur cohérence est assurée, simplifiant notablement la vérification de la correction de la base de connaissances. Enfin, indépendantes de la résolution d'un problème particulier, elles sont utilisables à différentes fins, en particulier pour l'explication du raisonnement suivi par le système expert.

Ce tableau idéal mérite cependant quelques retouches. Tout d'abord, l'applicabilité même de l'approche repose sur l'existence de connaissances théoriques dans le domaine d'application. Or, dans bien des domaines, de telles connaissances n'existent pas encore (tant est qu'elles existent un jour), si bien que ces travaux concernent essentiellement les domaines scientifiques et techniques. Par ailleurs, la représentation de ces connaissances, lorsqu'elles existent, soulève de nouveaux problèmes en Intelligence Artificielle. Comment représenter dans un système expert le fonctionnement d'un dispositif physique modélisé par des équations différentielles, par exemple ? En écrivant directement

les équations dans la base de connaissances ? La réponse n'est pas si simple; il ne faut pas oublier que le but de l'Intelligence Artificielle est de comprendre et reproduire les raisonnements effectués par les êtres humains. Or, il est hautement improbable qu'un humain raisonne sur un système d'équations différentielles pour déduire qu'une balle lancée en l'air retombera inévitablement. Ce souci d'anthropomorphisme implique donc de nouveaux modes de représentation des connaissances. Enfin, si l'uniformité de la représentation des connaissances est l'une des principales sources de problèmes dans les systèmes experts de première génération, elle n'en facilite pas moins la tâche en ne nécessitant qu'un seul mécanisme d'inférence. Disposer d'outils de représentation de connaissances profondes n'est d'aucune utilité sans les mécanismes d'inférence capables de les exploiter. Au problème de représentation, s'ajoute donc celui de l'exploitation des connaissances profondes. Devant la difficulté de ces recherches, il n'est pas surprenant que les systèmes experts à base de connaissances profondes n'aient pas encore comblé le retard qui les sépare des systèmes experts de première génération. Leurs résultats n'ont pour l'instant guère franchi les portes des laboratoires de recherche et des efforts substantiels restent nécessaires pour atteindre cet objectif. Cette thèse est, nous l'espérons, un travail dans ce sens.

Les recherches menées sur les systèmes experts à base connaissances profondes suivent deux axes principaux, associés au mode de représentation des connaissances qu'ils utilisent. Le premier concerne les travaux reposant sur une description causale du fonctionnement d'un système. Le principal outil de modélisation de ce type est le graphe de causalité. Les chercheurs suivant la seconde direction utilisent une description plus complexe reposant sur une interprétation qualitative des équations mathématiques décrivant le fonctionnement d'un système. Ces derniers travaux sont habituellement regroupés sous le nom de raisonnement qualitatif (bien que la première famille de travaux concerne également un raisonnement de type qualitatif). Ces deux axes sont successivement examinés dans les sections 1 et 2.

1 - Le raisonnement sur les graphes de causalité

Les graphes de causalité ne sont pas une invention récente; les économistes les utilisaient déjà dans les années 50 (voir [SIMON 57a] par exemple) pour analyser les propriétés de leurs modèles. L'étude de leurs extensions et des mécanismes d'inférence qui leur sont applicables revient cependant essentiellement aux chercheurs en Intelligence Artificielle.

1.1 - Représentation des connaissances par un graphe de causalité

Dans sa forme de base, un graphe de causalité est une description très simple du fonctionnement d'un système (cf. exemple de la figure 2.1). Il s'agit d'un graphe orienté dont les sommets représentent des états ou des variables du système considéré. Un arc lie deux sommets du graphe s'il existe entre eux une relation de cause à effet.

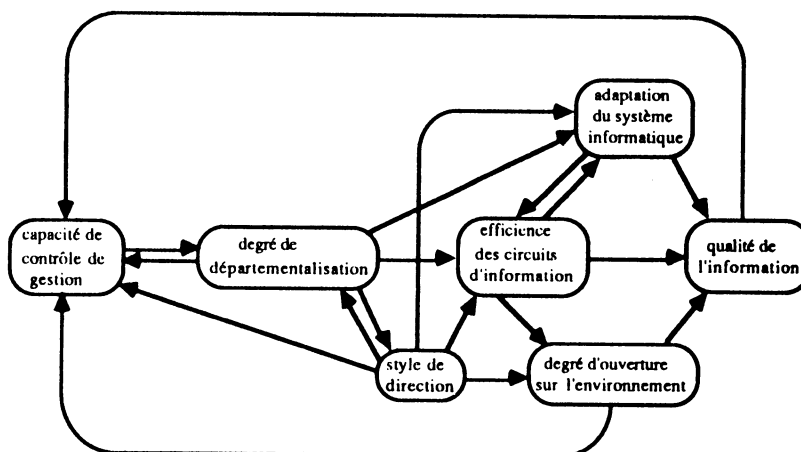


Figure 2.1: le graphe causal simplifié des variables caractérisant la compétitivité d'une entreprise d'après [LESCA 86]

Les graphes d'influences forment une classe importante des graphes de causalité (sur laquelle nous aurons largement l'occasion de revenir au cours de cette thèse). Leurs arcs sont étiquetés par un signe indiquant le sens de la relation causale: "+" si la variable initiale de l'arc influe positivement sur la variable terminale, et "-" dans le cas contraire.

La simplicité des graphes de causalité masque deux difficultés. Tout d'abord, la relation de cause à effet est imprécise: l'âge favorise l'apparition de certaines maladies, mais peut-on le

considérer pour autant comme une cause ? La seconde difficulté est liée à la faible quantité d'information contenue dans la représentation. L'existence de liens de causalité est souvent insuffisante pour raisonner sur le comportement d'un système: l'intensité des relations, leur probabilité, leur durée, etc s'avèrent fréquemment nécessaires.

Deux solutions ont été envisagées pour remédier à l'imprécision de la relation de cause à effet. La première affirme l'existence de plusieurs types de causalité et la nécessité de les distinguer ([RIEGER 77], [LONG 83], [SHIBAHARA 83]). Dix liens de causalité différents sont ainsi identifiés par C. Rieger pour modéliser le fonctionnement de mécanismes physiques. Ces dix liens se répartissent essentiellement en causalité continue (la relation de cause à effet est progressive) et instantanée (la présence de la cause un instant donné est suffisante pour provoquer l'effet). Cette représentation est beaucoup plus riche que les graphes de causalité élémentaires. Sa complexité rend toutefois la modélisation d'un système difficile. La seconde méthode proposée pour rendre la relation de cause à effet plus précise est de définir clairement sa sémantique. Cette démarche est appliquée dans [WELLMAN 87] aux graphes d'influences, où une interprétation probabiliste de la relation d'influence est proposée: la variable A influence positivement la variable B si la distribution de B sachant A augmente avec celle de A (au sens de la dominance stochastique du premier ordre). Cette définition souffre cependant d'être seulement applicable aux domaines pouvant être appréhendés statistiquement. Finalement, dans bien des travaux, la relation de causalité traduit l'existence d'une relation fonctionnelle entre des variables. Elle est alors interprétée de la manière suivante: la variable A est une cause de la variable B si B est fonction de A.

Différentes extensions ont été étudiées pour augmenter l'expressivité des graphes de causalité. La plupart utilisent un étiquetage des arcs qui leur permet de représenter diverses caractéristiques d'une relation causale. Un coefficient numérique ajouté à un arc mesure ainsi l'incertitude concernant l'existence d'une relation ([WEISS 78]). Le délai (le temps nécessaire pour que l'effet commence à se produire) et la durée peuvent être décrits de la même manière ([BLUM 83]). Lorsqu'un arc indique l'existence d'une relation fonctionnelle, il est possible de représenter ses caractéristiques de façon identique: le sens de variation de la fonction est ainsi décrit par le signe de la relation d'influence. Ces extensions se trouvent finalement bien résumées dans le système ABEL ([PATIL 81]); dans ce système, les différents attributs de la variable effet (durée, importance, ...) sont calculés à partir des attributs de la variable cause par des relations dépendant du contexte, avec la possibilité d'utilisation de valeurs par défaut (figure 2.2).

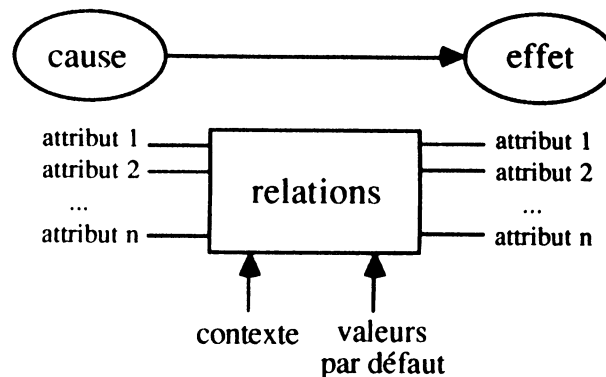


Figure 2.2: une relation de causalité dans le système ABEL, d'après [PATIL 81]

1.2 - Raisonnement dans un graphe de causalité

La propagation de cause à effet est le schéma de base du raisonnement dans un graphe de causalité. Elle permet d'analyser les conséquences d'une perturbation affectant un ou plusieurs sommets du système étudié. Dans un graphe de causalité élémentaire, la propagation consiste simplement à modifier l'état des sommets liés par un chemin au(x) sommet(s) perturbé(s); dans un graphe de causalité plus complexe comme celui de la figure 2.2, les attributs de chaque sommet affecté par la perturbation sont propagés le long des arcs du graphe. La propagation peut conduire à des indéterminations en présence d'effets contradictoires. C'est par exemple le cas dans un graphe d'influences lorsque deux paramètres variant dans le même sens affectent un troisième paramètre avec des signes d'influence contraire (nous reviendrons sur ce point dans la section suivante). Il faut noter qu'une part importante des travaux sur ce problème concerne la propagation de l'incertitude le long des arcs de causalité. Ces travaux constituent un domaine de recherche en soi, trop vaste pour être abordé ici.

Le diagnostic est la principale application des graphes de causalité. La propagation y est utilisée de différentes façons, notamment:

- pour déterminer la(es) cause(s) responsable(s) de symptômes ([WEISS 78], [KUNZ 83]). Un schéma fréquemment utilisé est le suivant: le graphe de causalité représente les mécanismes de propagation de l'ensemble des dysfonctionnements pouvant affecter le système analysé. Le graphe étant généralement sans circuit, les sommets initiaux représentent les hypothèses de dysfonctionnement tandis que les sommets terminaux représentent les symptômes. Chaque

hypothèse de dysfonctionnement (ou ensemble d'hypothèses préalablement choisi) est propagée le long du graphe afin de déterminer les symptômes qu'elle occasionne. Si l'ensemble des symptômes occasionnés correspond à l'ensemble des symptômes observés, l'hypothèse de dysfonctionnement est retenue comme une cause possible.

- pour analyser les conséquences d'une action correctrice ([KUNZ 83], [SWARTOUT 83]). Le graphe de causalité représente cette fois les mécanismes résultant de l'utilisation des différentes actions correctrices pouvant être administrées au système analysé. Les hypothèses de corrections sont propagées le long du graphe afin de déterminer si elles permettent de corriger le système défaillant sans engendrer d'autres dysfonctionnements.

Lorsque, contrairement aux deux cas précédents, les sommets dont l'état est recherché ne sont pas exclusivement des ancêtres des sommets dont l'état est connu, la propagation n'est plus suffisante: [SHIBAHARA 85] et [VAN DE VELDE 85] constituent deux exemples de ce type. Il est alors nécessaire de mettre en oeuvre un schéma d'inférence plus complexe combinant la propagation avant (i.e. la propagation telle que nous venons de la voir) et arrière (i.e. générant des hypothèses sur les causes des sommets dont l'état est connu). Un algorithme utilisant ce schéma est proposé dans [VAN DE VELDE 85] pour le diagnostic de pannes de dispositifs physiques.

Les graphes de causalité et le principal mode d'inférence qui leur est associé, la propagation, constituent un outil simple pour raisonner sur le fonctionnement d'un système. Dans bien des cas cependant, les arcs de causalité servent à représenter des relations fonctionnelles. Le raisonnement qualitatif sur les équations décrivant le fonctionnement d'un système est donc le prolongement naturel de cette démarche.

2 - Le raisonnement qualitatif

Le raisonnement qualitatif fait partie des développements récents de l'Intelligence Artificielle. Son objectif est de comprendre et reproduire les raisonnements humains sur des systèmes physiques. Cette recherche part de la constatation suivante: sans modèle mathématique sophistiqué ni valeurs numériques précises, un être humain est capable de comprendre le fonctionnement global d'un système; ceci lui permet de prédire son fonctionnement, le contrôler, le dépanner, etc. Est-il possible de concevoir une physique "naïve" comme l'appelle P. Hayes ([HAYES 79]) qui reproduise ces compétences ?

Les modèles numériques offrent bien la possibilité de simuler ou contrôler un système, mais la précision exigée par leur utilisation est sans commune mesure avec l'aspect qualitatif du raisonnement humain. C'est dans une autre direction que s'est engagé le raisonnement qualitatif. L'idée fondamentale en est l'utilisation de modèles "intelligibles" ne retenant que l'essence des phénomènes étudiés, mais conservant néanmoins un réel pouvoir de prédiction. Cette démarche a conduit à l'élaboration de nouveaux outils de modélisation, de calcul, qui ouvrent la voie à de nombreuses applications.

Cette section effectue une revue des différents travaux sur le raisonnement qualitatif. Les outils de modélisation proposés sont d'abord présentés, avant d'étudier leurs applications.

2.1 - Les modélisations proposées pour le raisonnement qualitatif

La modélisation qualitative d'un système comporte deux niveaux. Le premier, microscopique est la variable; celle-ci constitue l'entité élémentaire de la modélisation. Le second niveau est macroscopique; il décrit la structure du système et détermine notamment comment les variables du système interagissent. Comme nous allons le voir, il existe un relatif consensus sur la représentation des variables qualitatives, ce qui est loin d'être le cas pour la description structurelle.

2.1.1 - Les variables qualitatives

Comme son homologue en Physique "classique", une variable d'un modèle qualitatif représente un paramètre du système étudié. Il existe cependant une différence importante: une variable qualitative prend ses valeurs dans un ensemble fini de symboles. Ces symboles représentent des valeurs remarquables de la variable.

Une variable est caractérisée à un instant donné par sa valeur et sa dérivée; le couple (valeur, dérivée) est appelé état de la variable. La dérivée d'une variable qualitative x , notée $[dx]$, est l'un des éléments de l'ensemble $\{-, 0, +\}$. Ces symboles représentent le sens de variation de la variable: décroissant, constant ou croissant. La représentation de la valeur d'une variable diffère selon les auteurs. Pour J. De Kleer, la valeur d'une variable x , notée $[x]$, est également un élément de l'ensemble $\{-, 0, +\}$, avec pour signification: $[x] = -$ ssi $x < 0$; $[x] = 0$ ssi $x = 0$; $[x] = +$ ssi $x > 0$.

De leur côté, K. Forbus et B. Kuipers utilisent une représentation plus sophistiquée appelée espace des quantités, qui définit un ordre partiel entre les valeurs remarquables de la variable. De plus, le système de B. Kuipers est capable de découvrir de nouvelles valeurs remarquables en cours d'inférence et de les intégrer à l'espace des quantités.

Comme une part importante du raisonnement qualitatif concerne la dynamique des systèmes, il faut aussi représenter l'évolution d'une variable au cours du temps. Celle-ci est décrite par la succession des états de la variable, souvent appelée historique. Cependant, la notion de temps utilisée dans cette chronologie est différente de celle employée en physique classique: le temps est lui aussi symbolique. Chaque symbole marque un événement survenant au système, c'est-à-dire l'instant où une variable atteint une valeur remarquable. Il en résulte que la durée entre deux événements n'est ni quantifiable, ni constante.

Les variables sont insuffisantes pour décrire le fonctionnement d'un système. Un second niveau de modélisation est nécessaire pour définir comment elles interagissent: c'est le rôle de la description structurelle.

2.1.2 - La description structurelle d'un système

La plupart des recherches menées en raisonnement qualitatif concernent des systèmes dynamiques modélisés par des équations différentielles. Il existe cependant des divergences importantes entre les chercheurs sur la description structurelle qualitative devant être adoptée. On peut distinguer trois tendances principales sur ce problème. Dans la première, représentée par B. Kuipers, la description découle d'une simple traduction dans un langage qualitatif des équations différentielles régissant le fonctionnement du système. Dans la seconde, représentée par J. De Kleer, un système est un assemblage de composants interconnectés; le fonctionnement du système se déduit alors de la description des composants et des connexions qui les lient. Enfin, pour F. Forbus, le monde physique est régi par des processus qui agissent sur des objets et modifient leurs propriétés.

Ces trois points de vue sont confrontés d'une manière ou une autre à la description des interactions entre les variables qualitatives. Les équations, qui représentent ces interactions en physique classique, doivent donc être traduites au niveau qualitatif. Pour cette raison, il est nécessaire d'introduire un calcul qualitatif et les règles de ce calcul.

Les opérateurs du calcul qualitatif varient d'un auteur à l'autre, mais le point important est le suivant: le nombre des valeurs possibles d'une variable et de sa dérivée étant fini (et généralement petit), il est possible de définir chaque opération qualitative par une (ou plusieurs) table(s). Ainsi, l'addition et la multiplication qualitatives sur les valeurs des variables discrétisées dans l'ensemble {+, 0, -} sont présentées sur les tables ci-dessous:

$x \backslash y$	+	0	-
+	+	+	?
0	+	0	-
-	?	-	-

$x \backslash y$	+	0	-
+	+	0	-
0	0	0	0
-	-	0	+

L'addition qualitative: $[z] = [x] + [y]$

La multiplication qualitative: $[z] = [x] \times [y]$

Il faut noter que le résultat de l'addition de deux variables de signes opposés est ambigu (il est noté "?"): ceci est dû à la perte d'information entre le passage du numérique et le qualitatif.

Analysons comment ces opérateurs qualitatifs sont intégrés dans les différentes descriptions structurelles proposées.

2.1.2.1 - La description structurelle de Kuipers

La description structurelle d'un système proposée par B. Kuipers ([KUIPERS 84], [KUIPERS 86]) est très simple car elle découle directement des équations différentielles. Celles-ci sont transformées en contraintes qualitatives entre les paramètres du système, exprimées au moyen de six opérateurs qualitatifs: PLUS, MULT, MINUS, DERIV représentant respectivement l'addition, la multiplication, le moins unaire et la dérivation. Enfin, M^+ et M^- désignent la proportionnalité qualitative, c'est-à-dire $M^+(x,y)$ (resp. $M^-(x,y)$) si y est une fonction strictement croissante (resp. décroissante) de x . Chacun de ces opérateurs est défini par des tables comparables à celles présentées ci-dessus.

Illustrons cette description sur le système de la figure 2.3 représentant un bloc connecté à ressort qui glisse sans frottement sur un support.

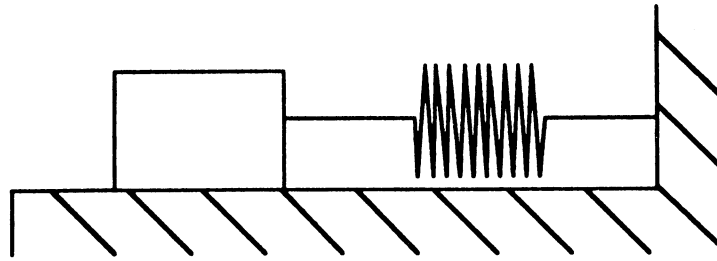


Figure 2.3: un bloc connecté à un ressort, glissant sans frottement sur un support

Supposons ce système décrit par les équations:

$$\left\{ \begin{array}{ll} F = m \gamma & \text{(loi de Newton)} \\ F = -k x & \text{(loi de Hooke)} \\ v = \frac{dx}{dt} \\ \gamma = \frac{dv}{dt} \end{array} \right.$$

Transformées dans le langage qualitatif de B. Kuipers, ces équations deviennent (puisque m et k sont des constantes positives):

DERIV(x,v);

DERIV(v,γ);

$M^-(\gamma,x)$

La modélisation structurelle de B. Kuipers est une simple traduction des équations décrivant un système dans un langage qualitatif. Ce langage est comparable à ceux utilisés pour la simulation analogique, la différence fondamentale résidant dans la sémantique des opérateurs.

2.1.2.2 - La description structurelle centrée sur la notion de composant

La description structurelle proposée par J. De Kleer ([DE KLEER 84a]) est plus complexe. Elle représente un système comme un assemblage de composants liés par des connexions. Chaque composant et connexion du système est considéré comme une instance d'une classe sur laquelle s'appliquent les lois de la physique. Ainsi, le ressort de la figure 2.3 est un élément de la classe des ressorts sur laquelle s'applique la loi de Hooke ($F = -kx$). De même, les lois physiques sont associées

aux connexions selon leur type: électrique, hydraulique, etc. Ceci laisse la possibilité de mécaniser la construction du modèle à partir de la description des composants et connexions du système étudié.

Les équations représentant les lois sont traduites dans un formalisme appelé confluences. Une confluence est une contrainte qualitative de la forme $\sum T_i = C$, où chaque T_i est soit une variable qualitative, la négation d'une variable ou le produit d'une constante et d'une variable et C une constante. Une équation modélisant un composant peut donner naissance à plusieurs confluences; celles-ci décrivent alors les différentes régions de fonctionnement du composant. Ainsi, le mobile de la figure 2.3 est soumis à la loi de Newton: $F = m \frac{dv}{dt}$. Compte tenu que m est positif, on a:

$$[dV] = [F]$$

Ceci n'est pas une confluence car F n'est pas une constante; on transforme donc cette relation en trois régions de fonctionnements ou états comme les appelle J. De Kleer:

ETAT ₁	$F > 0$	$[dV] = +$
ETAT ₂	$F = 0$	$[dV] = 0$
ETAT ₃	$F < 0$	$[dV] = -$

Cette description structurelle est plus riche que la précédente, car elle autorise plusieurs modes de fonctionnement pour un composant. Ce point est très important dans des domaines comme l'électronique, où certains composants ne sont modélisables qu'à cette condition. Par ailleurs, la construction automatique des confluences à partir des composants et connexions du système est envisageable. Cependant, comme le remarquent H. Simon et Y. Iwasaki ([IWASAKI 86a], [IWASAKI 86b]), les principes de cette modélisation ne sont pas très clairs. Selon le niveau de finesse auquel on se place, un élément peut être en effet modélisable comme une connexion ou un composant. Ainsi, le sol se comporte comme une connexion dans la figure 2.3, mais il devient un composant si l'on décide de modéliser le frottement du mobile sur le sol. Ceci entre en contradiction avec le souci de "transparence" que De Kleer entend imposer à sa modélisation.

2.1.2.3 - Les processus qualitatifs

La théorie des processus qualitatifs de K. Forbus ([FORBUS 84]) fournit une représentation différente du monde physique. Les événements y sont vus comme des processus agissant sur des

objets, modifiant leurs paramètres, causant à l'extrême leur disparition et/ou l'apparition d'autres objets.

La description structurelle d'un système comporte trois niveaux dans la théorie des processus qualitatifs.

- un **objet**¹ décrit une entité du système étudié; le ressort et le mobile en sont deux exemples dans le système de la figure 2.3 (le choix des objets pertinents dépend bien sûr des processus que l'on veut représenter). Chaque objet possède un ensemble de propriétés, comme la longueur et la force interne pour le ressort.
- une **vue** décrit l'état d'un objet; le ressort à l'état de repos en est un exemple. Une vue est caractérisée par un ensemble de conditions et de relations qualitatives. La vue est dite active si les conditions qui lui sont associées sont vérifiées; les relations qualitatives entre les propriétés de l'objet sont vérifiées. Ainsi, la vue "tendu" est active pour le ressort si sa longueur est supérieure à sa longueur de repos; la force interne est alors une fonction strictement croissante de la longueur (ce qui est noté: force-interne α + longueur).
- les **processus** décrivent les actions qui modifient les objets et leurs propriétés au cours du temps; l'élongation, la rupture du ressort en sont des exemples. Comme une vue, un processus est actif sous certaines conditions. Lorsqu'elles sont remplies le processus agit sur les paramètres des objets par l'intermédiaire de relations d'influence (cf. section 2). Ainsi, lorsque le processus d'élongation est actif, celui-ci a une influence positive sur la longueur du ressort.

Il faut noter que la causalité est représentée de manière explicite dans la théorie des processus qualitatifs. En effet, les relations qualitatives décrites sous forme de contraintes dans les deux descriptions structurelles précédentes, sont ici représentées de manière causale; que ce soit sous forme de relations d'influence ou de proportionnalités qualitatives. La causalité dans le fonctionnement d'un système est donc codée et non "découverte". Malgré cet inconvénient, la théorie des processus qualitatifs est une description très riche du monde physique. En autorisant jusqu'à la modélisation de l'apparition ou de la disparition d'objets, elle confère une très grande souplesse à la structure d'un système.

¹nous adaptons très librement la terminologie de K. Forbus

Les trois types de descriptions structurelles présentées adoptent un point de vue très différent sur la représentation du monde physique. C'est le domaine d'application qui dicte laquelle est la plus adaptée à un problème particulier.

2.2 - Les applications du raisonnement qualitatif

Par rapport aux réseaux de causalité, le raisonnement qualitatif offre la possibilité de représenter la dynamique d'un système. Aussi, la plupart des recherches se sont orientées vers la simulation qualitative. Outre son intérêt intrinsèque, celle-ci constitue également la base d'autres applications. Comme il n'est pas possible de les passer toutes ici en revue, les trois principales sont décrites après une présentation des principes de la simulation qualitative. Ces applications sont l'analyse comparative, l'interprétation de résultats numériques et le diagnostic.

2.2.1 - La simulation qualitative

Le but de la simulation qualitative est de déterminer l'évolution d'un système à partir d'un état initial. A cause de la perte d'information par rapport au numérique, il arrive cependant souvent que le comportement du système ne puisse être prédit avec certitude. Aussi, la simulation détermine toutes les évolutions possibles du système, ou autrement dit construit l'arbre des états possibles du système. Par ailleurs, la simulation qualitative offre un avantage important par rapport à la simulation quantitative: il n'est pas nécessaire de connaître les paramètres du modèle avec précision, puisque seules les valeurs remarquables des variables sont utilisées.

Etant donné les divergences existant entre les chercheurs sur la modélisation, il est évident que les techniques de simulations adoptées diffèrent. Celles-ci emploient cependant un principe commun: la satisfaction de contraintes.

Deux types de contraintes sont utilisées pour la simulation. Les relations qualitatives de la description structurelle, tout d'abord, contraignent les valeurs des variables et de leur dérivée. Ces relations ne déterminent pas en général un ensemble unique de valeurs, mais différents ensembles de valeurs compatibles. Supposons ainsi trois variables a , b et c liées par la relation: $[c] = [a] + [b]$, et telles que $[c] = +$. Grâce à la table de l'addition qualitative on déduit que les valeurs possibles du couple $([a],[b])$ sont: $(+,+)$, $(+,0)$, $(+,-)$, $(0,+)$ et $(-,+)$.

Les contraintes du deuxième type dérivent d'une traduction des résultats classiques de l'analyse sur les fonctions continues et dérivables au niveau qualitatif ([KUIPERS 86], [WILLIAMS 84]). Le théorème des valeurs intermédiaires permet ainsi d'affirmer qu'une variable ayant comme espace des quantités l'ensemble $\{+, 0, -\}$ ne peut passer en un pas de simulation de la valeur $-$ à la valeur $+$ (et inversement); le passage intermédiaire par la valeur 0 est obligatoire. Les théorèmes reliant les fonctions à leur dérivée apportent des résultats similaires: si $[x] = 0$ et $[dx] = +$ dans un état donné, on aura à l'état suivant: $[x] = +$.

A chaque pas de simulation, les états successeurs de l'état courant sont ceux satisfaisant ces deux types de contraintes. Cependant, les contraintes sont utilisées de manière différente selon les algorithmes de simulation.

Dans QSIM ([KUIPERS 86]), un pas de simulation se déroule en deux étapes. Les transitions de chaque variable à partir de l'état courant sont d'abord générées grâce à une table qui pour l'état i d'une variable (valeur, dérivée) $_i$ déduit tous les états successeurs possibles (valeur, dérivée) $_{i+1}$. Cette table découle des règles de continuité et d'intégration décrites au paragraphe précédent. Le produit cartésien des ensembles d'états successeurs de chaque variable constitue l'ensemble des états successeurs potentiels du système. Ces états sont donc ensuite filtrés afin de ne retenir que ceux respectant les contraintes de la description structurelle.

Dans la description structurelle de J. De Kleer, l'état d'un système est décrit par l'ensemble des confluences associées à ses composants. Les dérivées de chaque variable sont donc obtenues en résolvant le système de confluences. S'il existe plusieurs solutions, cela signifie que le système peut manifester plusieurs comportements possibles dans l'état considéré. Les dérivées sont ensuite intégrées en utilisant les règles d'intégration et de continuité afin de déduire les transitions possibles entre états.

L'algorithme de simulation utilisé dans la théorie des processus qualitatifs ([FORBUS 86b]) est encore différent. Les processus actifs sont dans cette description structurelle les causes d'évolution d'un système. Ils agissent sur les propriétés des objets de deux façons différentes: directement, par les relations d'influence, mais aussi indirectement lorsqu'une propriété liée à une autre par une relation de proportionnalité qualitative est elle-même directement influencée. Un pas de simulation se déroule donc entre deux étapes: les influences directes, puis indirectes sont résolues afin de déterminer les dérivées des propriétés (en calculant la somme qualitative des influences, avec

l'ambiguïté que cela comporte). Celles-ci sont ensuite intégrées dans leur espace des quantités afin de déterminer les nouvelles valeurs remarquables atteintes.

Malgré son apparence séduisante, la simulation qualitative pose deux problèmes délicats. Le premier est le non-déterminisme de la simulation: l'évolution d'un système est rarement prédite avec certitude. Ce point n'est pas toujours problématique; ainsi, en phase de conception, on souhaite précisément envisager tous les comportements possibles du système. Pour d'autres problèmes où le choix d'une action dépend du résultat de la simulation, l'ambiguïté est néanmoins rédhibitoire. De plus, Kuipers a montré ([KUIPERS 86]) que parmi les évolutions proposées par la simulation qualitative, certaines n'ont aucun sens physiquement. Plusieurs solutions ont été envisagées pour surmonter cette difficulté: l'utilisation de dérivées d'ordre supérieur ([DE KLEER 84b], [MORGAN 87]); la distinction dans la description structurelle des effets à court et à long terme ([KUIPERS 87]); l'adjonction d'informations numériques sur le modèle ([SIMMONS 83], [RAJAGOPALAN 84], [D'AMBROSIO 87]); enfin et surtout le raisonnement sur les ordres de grandeur des variables ([RAIMAN 86], [MAVROVOUNIOTIS 87]) sur lequel nous allons revenir. Ces solutions ne font cependant que réduire les ambiguïtés, et ne peuvent en aucun cas pallier à la perte d'information résultant du passage du numérique au qualitatif.

La deuxième difficulté soulevée par la simulation qualitative (et d'ailleurs tout raisonnement qualitatif) est la connexion des techniques de calcul qualitatif avec des systèmes physiques réels. Raisonner sur les valeurs remarquables des variables et leur dérivée est une chose; obtenir ces valeurs en est une autre. Que signifie par exemple " $[dx] = 0$ " lorsque l'on observe un mécanisme ? que la valeur de x doit être strictement égale entre deux instants ? que la variation de x doit se situer dans une certaine plage ? est-on sûr dans ce cas que les opérations du calcul qualitatif resteront valides ? L'effort consacré à la construction d'une physique naïve n'a guère permis d'aborder ces problèmes jusqu'à présent ([CALOUD 88], [LEPETIT 87] fournissent cependant des réflexions sur ces points). Ils sont pourtant décisifs pour la réalisation d'applications opérationnelles.

Malgré les deux critiques qui viennent d'être énoncées, la simulation constitue la base d'autres modes de raisonnement qualitatif. Nous présentons les trois principaux.

2.2.2 - L'analyse comparative

L'analyse comparative ([WELD 88a], [WELD 88b]) est un complément très important de la simulation qualitative pour expliquer le fonctionnement d'un système. Equivalent qualitatif de l'analyse de sensibilité, son objectif est de déduire la réaction d'un système à une perturbation de l'un (ou de plusieurs) de ses paramètres. Cette tâche permet dans le système de la figure 2.3, de répondre à des questions comme: "que se passe-t-il si l'on remplace le mobile par un autre de masse supérieure?".

D. Weld propose deux techniques pour résoudre ce problème: l'exagération et l'analyse différentielle qualitative. L'exagération est un raisonnement asymptotique qui consiste à effectuer deux simulations qualitatives du système: la première sans la perturbation, la seconde en la supposant infinie. Les résultats des deux simulations sont ensuite comparés pour déterminer les conséquences de la perturbation. L'exagération est une technique simple, mais elle produit des résultats incorrects lorsque le système ne fournit pas une réponse monotone à la modification des paramètres. Aussi, Weld a mis au point une seconde méthode appelée analyse différentielle qualitative. Celle-ci repose sur une notion appelée perspective, une sorte de dérivée partielle qualitative, qui permet d'exprimer les variations relatives d'une variable entre deux situations. A l'inverse de l'exagération, les résultats produits par cette méthode sont toujours corrects, mais il est possible de construire des exemples pour lesquels elle ne fournit pas de résultats.

Aucune des deux méthodes n'est totalement satisfaisante, mais Weld a montré que leur combinaison produit des résultats corrects sauf dans certains cas pathologiques.

2.2.3 - Interprétation de résultats numériques

Une autre application du raisonnement qualitatif présentant un grand intérêt est l'interprétation de résultats numériques. A partir de données numériques concernant le fonctionnement d'un système (machine, procédé, ...), il s'agit de déterminer si le système fonctionne correctement et éventuellement proposer une action correctrice. Le raisonnement qualitatif peut jouer un rôle important dans ce domaine en fournissant une description qualitative du fonctionnement du système, plus compréhensible que les données numériques brutes.

[FORBUS 86a] propose une solution basée sur la simulation qualitative. Les résultats sont segmentés en données qualitatives, puis la séquence de segments est comparée à l'arbre des états obtenus par simulation qualitative afin de déduire les processus pouvant être rendus actifs. Cette

solution n'est pas réellement satisfaisante car la simulation qualitative est inadéquate en cas d'ambiguïté: l'action à entreprendre est en effet alors indéterminée. Par ailleurs, aucune précision n'est fournie sur la discrétisation des données numériques en segments qualitatifs.

La solution totalement différente proposée dans [CALOUD 87] est beaucoup plus satisfaisante. Dans cette approche, un système est décrit par plusieurs modèles de finesse différente. Le modèle le plus fin, utilisé pour la simulation, est composé de relations causales sophistiquées. Ces dernières contiennent des informations permettant de propager l'historique d'une variable sur une autre: fonction de transfert, gain, retard, etc. La propagation est effectuée numériquement mais elle est beaucoup plus souple qu'une simulation numérique traditionnelle car elle autorise la présence d'éléments qualitatifs: ainsi, une donnée temporelle peut être un élément de l'ensemble de symboles {court, lent, moyen}. Le système est simulé en sommant les influences s'exerçant sur chaque variable de manière à déterminer leur historique. Le résultat de la simulation est ensuite transformé en une description de plus haut niveau grâce aux modèles plus agrégés, ce qui permet de fournir une interprétation des résultats dans un langage adapté à l'utilisateur.

La nécessité de combiner différents modèles semble une constante de l'interprétation de résultats numériques car ce point est également noté dans [SIMMONS 83] pour l'interprétation de coupes géologiques.

2.2.4 - Diagnostic

Le raisonnement qualitatif étant un outil de choix pour analyser le fonctionnement d'un système, il n'est pas surprenant qu'il présente également des applications dans le diagnostic. Deux types de systèmes de diagnostic peuvent être distingués selon qu'ils comportent ou non des modèles de pannes.

L'application du raisonnement qualitatif au diagnostic avec modèles de pannes est immédiate. Un exemple de ce type appliqué au diagnostic de maladies rénales est présenté dans [KUIPERS 87]. Chaque hypothèse de dysfonctionnement est simulée qualitativement; si le résultat de la simulation correspond aux symptômes observés, elle est retenue comme un diagnostic possible.

Plus complexe est l'approche sans modèle de pannes. [RAIMAN 87] présente une application de ce type au diagnostic de circuits analogiques. Le diagnostic est effectué par un raisonnement

qualitatif sur les relations modélisant le fonctionnement correct du circuit. Si la tension mesurée en un point du circuit ne peut être expliquée par le fonctionnement correct des composants connectés à ce point, l'un au moins de ces composants est défectueux. Les hypothèses de diagnostic sont obtenues en recoupant les observations sur différents points du circuits. Un point important souligné par O. Raiman est l'insuffisance du raisonnement qualitatif sur l'ensemble $\{-, 0, +\}$ pour le diagnostic de pannes dans un circuit analogique. Aussi, il a proposé une extension du raisonnement qualitatif basée sur les ordres de grandeur des variables ([RAIMAN 86]). Outre les opérateurs classiques de l'arithmétique qualitative, trois opérateurs sont ajoutés: "négligeable", "voisin" et "comparable" et intégrés dans un système formel. Une règle de ce système est présentée ci-dessous:

$$A \text{ négligeable } B \Rightarrow (A + B) \text{ voisin } B.$$

Le raisonnement sur les ordres de grandeur constitue une extension importante du calcul qualitatif; il permet notamment d'en réduire l'ambiguïté.

Comme nous venons de la voir, le raisonnement qualitatif est source de multiples applications. Il reste cependant largement tributaire de la possibilité de couplage à des systèmes physiques réels. L'objectif initial de construire une physique naïve, s'il contribue aux développements théoriques, participe fort peu à la résolution de ce problème. Les réflexions allant dans ce sens ([CALOUD 88], [KUIPERS 88]) montrent la nécessité de se rapprocher des outils de calculs numériques. Le problème est de savoir si le gain de performance résultant de cette démarche compensera la perte d'intelligibilité des connaissances qui ne manquera pas de l'accompagner.

Cette présentation des travaux sur les systèmes experts utilisant des connaissances profondes est loin d'être exhaustive. Les recherches concernant les systèmes dont les paramètres varient de manière continue ont été plus particulièrement examinées, à cause de leur généralité. D'autres travaux sont menées sur des systèmes ayant un fonctionnement discret ([KOTON 85], [WELD 86]). Nous aurons l'occasion de revenir sur ce point au chapitre 7, puisqu'un exemple de ce type a été développé pour l'imputation comptable.

CHAPITRE 3

INTERPRÉTATION COMPARATIVE

L'analyse de l'évolution d'un système, l'examen d'écart entre des observations et des prévisions, la confrontation de plusieurs hypothèses semblent à première vue des problèmes très différents. Pourtant, ils ont un point commun: la **comparaison** de plusieurs situations. Ce chapitre tente de montrer que, vus sous cet angle, ils peuvent être abordés en une même démarche par des systèmes experts utilisant des connaissances profondes. Dans ce but, une nouvelle classe d'applications pour les systèmes experts est proposée: **l'interprétation comparative**. Elle a pour but la mise en évidence et l'explication des différences entre plusieurs situations. Une méthode générale pour réaliser des systèmes d'interprétation comparative, utilisant un modèle du domaine d'application, est présentée. Des techniques permettant la mise en oeuvre de cette méthode pour des modèles qualitatifs et numériques sont décrites dans les deux chapitres suivants.

La notion d'interprétation comparative et ses applications sont d'abord présentées dans la section 1. Le traitement d'un problème d'interprétation comparative se compose de deux étapes, l'évaluation et l'explication comparatives, qui sont successivement examinées dans la section 2. Enfin, cette approche est comparée aux travaux sur le diagnostic à base de modèles en section 3.

1 - Interprétation comparative

L'interprétation est une tâche réalisée par les systèmes experts consistant à déduire une représentation abstraite de la description d'une situation (comportant généralement un grand nombre de données). Nous définissons l'interprétation comparative comme la mise en évidence et l'explication des différences entre plusieurs situations. Ces deux aspects de l'interprétation comparative se trouvent parfaitement illustrés sur l'exemple 1.

Exemple 1

Le diagnostic financier consiste à analyser la situation d'une entreprise et son évolution. Ce dernier point est abordé en comparant différents états de l'entreprise, représentées par les documents comptables qu'elle publie chaque année, afin de mettre en évidence les changements importants qui sont intervenus dans son fonctionnement. Pour apprécier ces changements, il est nécessaire d'en expliquer la cause; une baisse de la rentabilité, par exemple, n'a pas les mêmes conséquences si elle est due à des causes accidentelles ou profondes. Ces deux composantes, mise en évidence des changements intervenant dans la situation de l'entreprise et explication de ces changements, constituent les caractéristiques d'un problème d'interprétation comparative.

Une méthode générale a été développée pour réaliser des systèmes d'interprétation comparative; elle repose sur l'existence d'un **modèle** du domaine d'application. Elle est applicable à différents langages de description du fonctionnement d'un système: équations et graphes de causalité, que les variables soient qualitatives ou numériques. Nous définissons un modèle de manière à prendre en compte ces différentes possibilités.

1.1 - Modèle

On appelle modèle, un ensemble de relations liant des variables du domaine d'application tel que:

- les variables du domaine d'application sont soit numériques soit qualitatives (i.e. prenant leur valeur dans l'ensemble $\{-, 0, +\}$)
- les relations sont soit numériques (système d'équations) soit qualitatives (graphe de causalité)

Il est ainsi possible de concevoir des problèmes d'interprétation comparative où le système analysé est modélisé par un ensemble d'équations liant des variables dont les valeurs numériques

précises sont inconnues (ou inversement par un graphe de causalité liant des variables numériques de valeur connue).

On représente un modèle M par un graphe biparti non orienté $G = (X,U)$, appelé graphe des relations de M tel que:

- $X = V \cup R$ où V est l'ensemble des variables de M , et R l'ensemble des relations.
- $(x,r) \in U$ si et seulement si $\exists x \in V, \exists r \in R: x \in \text{var}(r)$, où $\text{var}(r)$ est l'ensemble des variables apparaissant dans la relation r .

Supposons pour reprendre l'exemple 1 qu'un aspect de la situation d'une entreprise, sa capacité de générer de la valeur ajoutée, soit représenté par un modèle: le système d'équations (S_1) . Le graphe associé à (S_1) est présenté sur la figure 3.1.

$$(S_1) \left\{ \begin{array}{l} \frac{\text{valeur_ajoutée}}{\text{production_exercice}} = \text{taux_efficacité} \quad (r_1) \\ \text{production_globale} - \text{consommations_intermédiaires} = \text{valeur_ajoutée} \quad (r_2) \\ \text{production_exercice} + \text{marge_commerciale} = \text{production_globale} \quad (r_3) \\ \text{apro_consommés} + \text{autres_achats} = \text{consommations_intermédiaires} \quad (r_4) \end{array} \right.$$

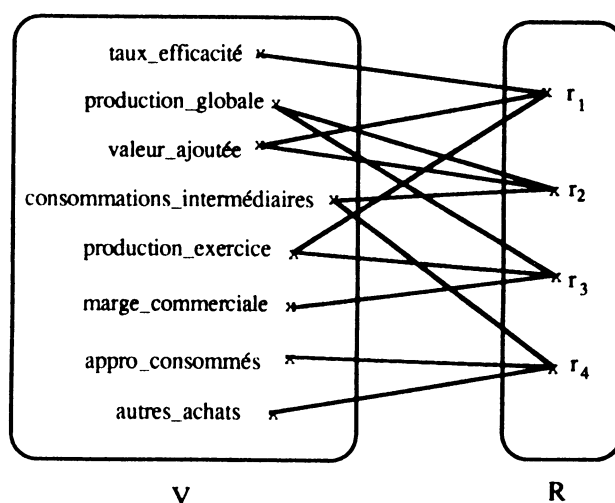


Figure 3.1: le graphe des relations du modèle (S_1)

1.2 - Interprétation comparative dans un modèle

La méthode développée pour réaliser des systèmes d'interprétation comparative repose sur l'existence d'un modèle (tel qu'il vient d'être défini) du domaine d'application. Nous utilisons en effet les relations liant les variables pour analyser les causes des différences entre les deux situations comparées. Ceci nous amène à considérer les situations comparées comme des **états** du système analysé. Chacun d'eux est défini par l'ensemble des valeurs prises par les variables du modèle.

Nous nous limitons dans cette thèse au problème déjà complexe de la comparaison de **deux états**. Cette restriction simplifie la tâche en fixant un cadre précis à la comparaison: le premier état sert de **référence** auquel on compare le second, l'**état dérivé** (figure 3.2); les deux états jouent donc des rôles dissymétriques. L'interprétation comparative est la mise en évidence et l'attribution des écarts des variables entre l'état de référence et l'état dérivé à d'autres variables, à l'aide des relations du modèle. Le but de cette opération est de générer un rapport en français présentant les écarts entre les deux états ainsi que leurs causes. Il faut noter que l'aspect linguistique est secondaire dans ce problème car la construction de l'explication repose essentiellement sur des raisonnements concernant le modèle et les valeurs des variables.

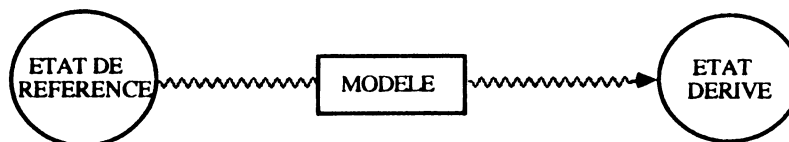


Figure 3.2: interprétation comparative d'un système, ou comment un modèle permet d'expliquer le passage du système de l'état de référence à l'état dérivé

1.3 - Applications

Même en considérant la limitation sur le nombre d'états analysés à deux, l'interprétation comparative a de nombreuses applications. Cette remarque est particulièrement vraie en Gestion, domaine dans lequel les prises de décision sont en effet fréquemment précédées par une analyse impliquant la comparaison de situations. Citons par exemple:

- **l'analyse du passé d'un système**

Cette tâche est assimilable à la comparaison de l'état d'un système entre les divers instants qui composent son historique. Lorsque la comparaison entre deux instants est suffisante, l'interprétation comparative peut remplir ce rôle. C'est le cas lorsqu'on s'intéresse à l'évolution globale du système et non à son cheminement ou que l'on dispose de peu d'observations sur l'historique du système. Cette possibilité sera illustrée en seconde partie de la thèse par une application au diagnostic financier d'entreprises.

- **le contrôle**

M. Gervais ([Gervais 83], p. 2) définit le contrôle de Gestion de la manière suivante:

"Le contrôle d'une organisation suppose que l'organisation définisse explicitement ou implicitement, un état désiré ... L'information recueillie à propos de l'état actuel sera ensuite comparée à l'état désiré et si une différence significative apparaît, une action sera entreprise."

Il est intéressant de noter que cette formulation entre directement dans le cadre de l'interprétation comparative. Ce point est à la base de la réalisation d'un système expert pour l'analyse des performances d'une entreprise par la méthode des surplus.

- **l'étude de variantes**

L'interprétation comparative peut aussi jouer un rôle dans le processus de prise de la décision. La comparaison de plusieurs états hypothétiques par rapport à un état désiré constitue en effet un outil d'aide à la décision important. Notons toutefois que la comparaison **simultanée** de toutes les hypothèses requiert l'interprétation comparative de n ($n > 2$) états.

2 - Analyse des problèmes d'interprétation comparative

Outre son aspect typologique, l'interprétation comparative présente un intérêt méthodologique, car elle est partiellement formalisable. Une méthode générale pour réaliser des systèmes experts d'interprétation comparative est proposée dans cette section. Indépendante du formalisme du modèle, elle est applicable aux modèles qualitatifs et numériques.

L'interprétation comparative est décomposable en deux étapes (figure 3.3). La première, appelée **évaluation comparative**, détermine l'ensemble des variables dont il faut expliquer l'écart. Les traitements effectués dans cette étape varient en fonction du domaine d'application ; ils ne sont pas généralisables. La seconde étape, l'**explication comparative**, attribue l'écart d'une variable à d'autres variables. Contrairement à la précédente, cette étape présente une régularité qui la rend formalisable. Après une brève présentation de l'évaluation comparative, nous analysons en détail cette deuxième étape.

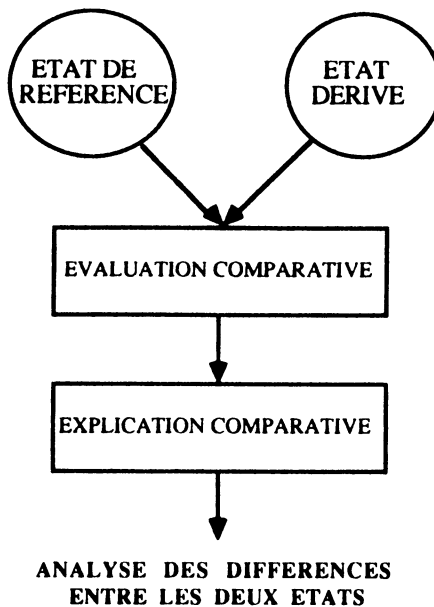


Figure 3.3: les deux étapes de l'interprétation comparative

2.1 - L'évaluation comparative

Dans un problème d'interprétation comparative, l'ensemble des variables à expliquer est rarement défini *a priori*. Ainsi, l'observation de l'impact d'une variable explicative sur un écart, peut conduire à l'explication de cette variable, elle-même demandant d'autres explications. De même, l'ordre d'analyse des écarts présente généralement un aspect dynamique: l'explication d'une variable dont l'écart doit par ailleurs être analysé en détail peut être retardée, etc... Ce rôle de pilotage de l'explication est assuré par l'évaluation comparative.

Les modalités du choix des variables expliquées et l'enchaînement des explications varient selon les problèmes traités. Les critères rendant nécessaire l'explication d'une variable peuvent en effet être très

simples (en fonction du taux de variation de la variable entre les deux états, par exemple) comme très complexes. Ils peuvent faire appel à des techniques formelles comme heuristiques. Aussi faut-il renoncer à traiter ce problème de manière générale. L'exemple 2, détaillé dans la seconde partie de la thèse, illustre la complexité du pilotage de l'explication que doit effectuer l'évaluation comparative.

Exemple 2

Poursuivons l'analyse du diagnostic financier comme un exemple d'interprétation comparative. L'état d'une entreprise est caractérisé par la valeur de plusieurs centaines de variables numériques. Il est nécessaire d'analyser l'évolution de la situation de l'entreprise à un niveau d'abstraction plus élevé que celui de ces données brutes. Dans ce but, des variables de synthèse qui agrègent les variables numériques des deux états comparés sont d'abord dégagées (figure 3.4) au cours de l'évaluation comparative. Elles sont ensuite comparées: si un écart important est constaté sur une variable de synthèse entre les deux états, une dégradation (ou une amélioration) de l'aspect de la situation qu'elle représente s'est produite; l'explication comparative des variables qu'elle agrège doit alors déclenchée, elles-mêmes pouvant nécessiter l'explication d'autres variables, etc...

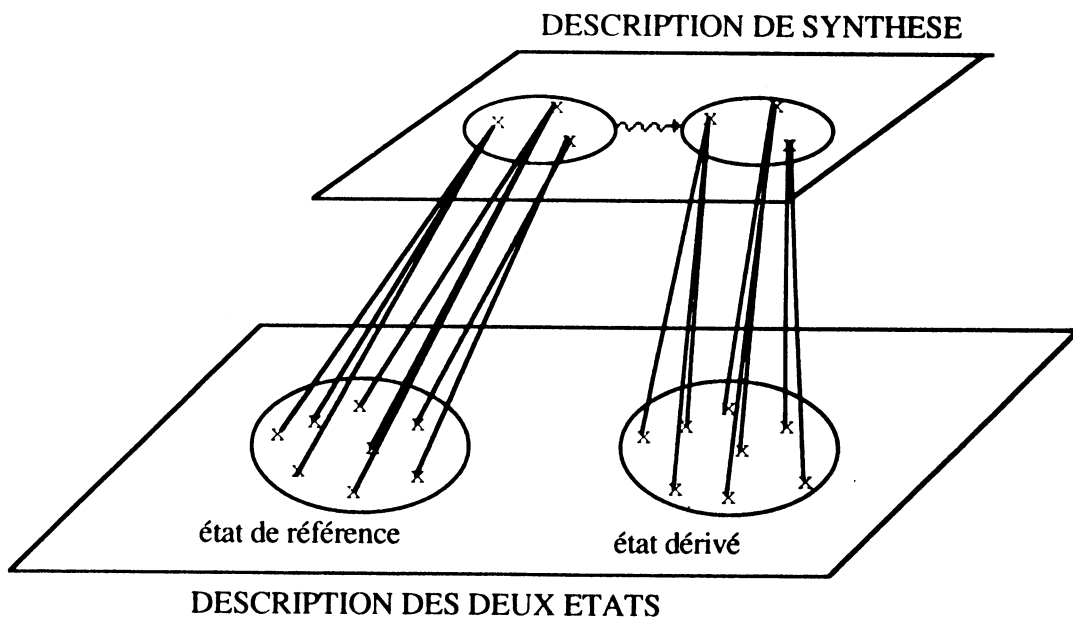


Figure 3.4: un exemple d'évaluation comparative appliqué au diagnostic financier

2.2 - Explication comparative

L'explication comparative fournit les causes de l'écart d'une variable. Elle est formalisable, contrairement à l'évaluation comparative; elle se divise elle-même en trois étapes (figure 3.5). Un graphe causal du modèle est d'abord construit. Il permet d'introduire la notion de **causalité**, nécessaire à l'attribution de l'écart de la variable expliquée à d'autres variables. La **base**, la partie du graphe causal utile à l'explication de la variable est ensuite déterminée. Enfin, la base d'explication est présentée à l'utilisateur sous forme d'un texte en français expliquant les causes de l'écart observé.

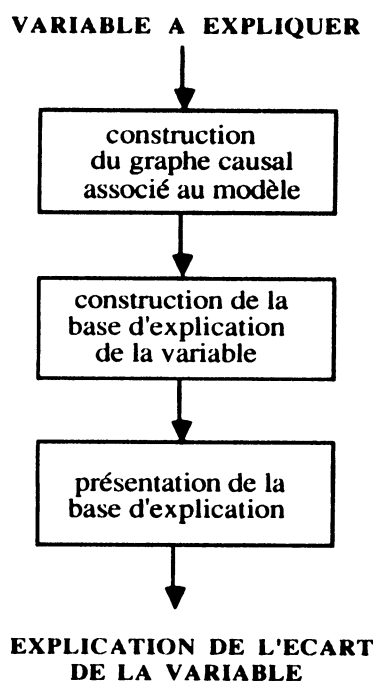
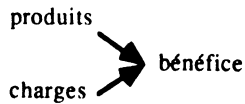


Figure 3.5: explication comparative d'une variable

2.2.1 - Le graphe causal

Dans un modèle, tel qu'il a été défini dans la section 1, les relations entre les variables sont symétriques. Or, l'explication de l'écart d'une variable requiert la notion de causalité, une relation asymétrique entre les variables. On dira ainsi à propos de l'équation produits - charges = bénéfice: "le bénéfice a diminué à cause de l'augmentation des charges", et non l'inverse. Cette explication fait apparaître les relations causales:



La causalité est l'objet de nombreux débats en Philosophie. Nous adoptons le point de vue de H. Simon ([SIMON 57a] p. 50):

"la phrase A est liée causalement à la phrase B (dans un système de phrases)" et non "ce qui est dénoté par la phrase A est la cause de ce qui est dénoté par la phrase B."

En se plaçant au niveau du langage et non du monde réel, il évite la critique objectant l'impossibilité de percevoir la causalité entre événements. Ce point de vue convient parfaitement à notre objectif d'explication.

H. Simon présente aussi une technique appelée **ordre causal**, pour déterminer les relations de causalité dans un modèle. Cette technique est applicable aux modèles numériques, qualitatifs et logiques. Elle est fondée sur l'ordre dans lequel les variables du modèle se déduisent. Dans ce but est effectuée la distinction entre variables **endogènes** (i.e. gouvernées par des lois représentées dans le système analysé), et variables **exogènes** (i.e. gouvernées par des lois extérieures au système analysé). L'ordre causal suppose connues les variables exogènes et analyse l'ordre dans lequel les variables endogènes peuvent être déterminées.

Exemple 3

Supposons les variables `appro_consommés`, `autres_achats`, `production_exercice` et `marge_commerciale` exogènes dans le système d'équations (S_1). La résolution de (S_1) peut être effectuée de la manière suivante: `appro_consommés`, `autres_achats` étant connues, `conso_intermediaires` peut être déterminée. De même, les variables `production_exercice` et `marge_commerciale` étant connues, la variable `production_globale` est calculable. La variable `valeur_ajoutée` peut alors être résolue, permettant enfin le calcul de la variable `taux_efficacité`.

Le calcul de l'ordre causal est automatisable. L'algorithme pour les modèles qualitatifs et numériques proposé dans [SIMON 57a] et [IWASAKI 86a] sera présenté dans les chapitres 4 et 5. Le lecteur intéressé par l'ordre causal dans un modèle logique pourra se reporter à [SIMON 57b].

L'ordre causal produit un graphe causal du modèle:

Soit M , un modèle et $G = (X = V \cup R, U)$ son graphe de relations. On appelle **graphe causal** de M , le graphe orienté $G_C = (S, C)$ tel que:

- S définit une partition de l'ensemble de variables de V , par rapport à la relation d'équivalence: "est calculable simultanément à ". Autrement dit, deux variables appartiennent à la même classe de S si et seulement si elles sont déterminées simultanément.
- s_1 et s_2 étant deux classes de S , $(s_1, s_2) \in C$ si et seulement si la résolution des variables s_1 précède directement celle de l'ensemble de variables s_2 .

La figure 3.6 représente le graphe causal du système d'équations (S_1), compte tenu des hypothèses figurant dans l'exemple 3.

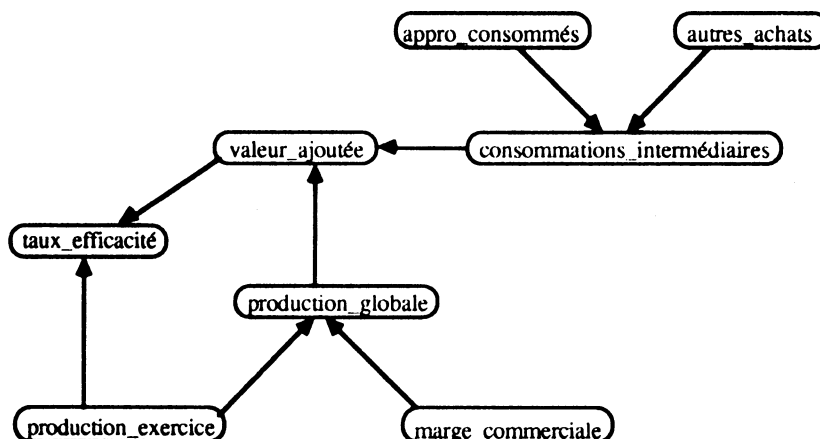


Figure 3.6: le graphe causal du système d'équations (S_1)

Cet exemple est un cas particulier car chaque classe est composée d'une seule variable.

Le graphe causal d'un modèle étant indépendant de la variable expliquée, il va de soi qu'il est construit une seule fois si l'écart de plusieurs variables doit être analysé.

Par son mécanisme d'ordre causal qui détermine les liens de causalité entre les variables à partir des relations du modèle, H. Simon fournit la clé de la première étape de l'explication comparative. La contribution de cette thèse qui se situe au niveau des deux étapes suivantes, peut donc être vue comme un prolongement de ses travaux au domaine de l'explication.

2.2.2 - Base d'explication

Les relations de causalité entre variables connues, il convient de choisir les variables explicatives. Ce choix est important: de lui dépend la qualité de l'explication fournie. Il doit permettre de construire une explication simple mais complète. A cette fin, on construit la **base d'explication** (figure 3.7) de la variable interprétée.

Définition 1

Soit M , un modèle et G_C son graphe causal. Soit x , une variable endogène de M dont la classe d'équivalence dans G_C est s . On appelle **base d'explication** de x , tout sous-graphe connexe de G_C engendré par l'union de $\{s\}$ et d'un sous-ensemble non vide des ancêtres de s dans G_C .

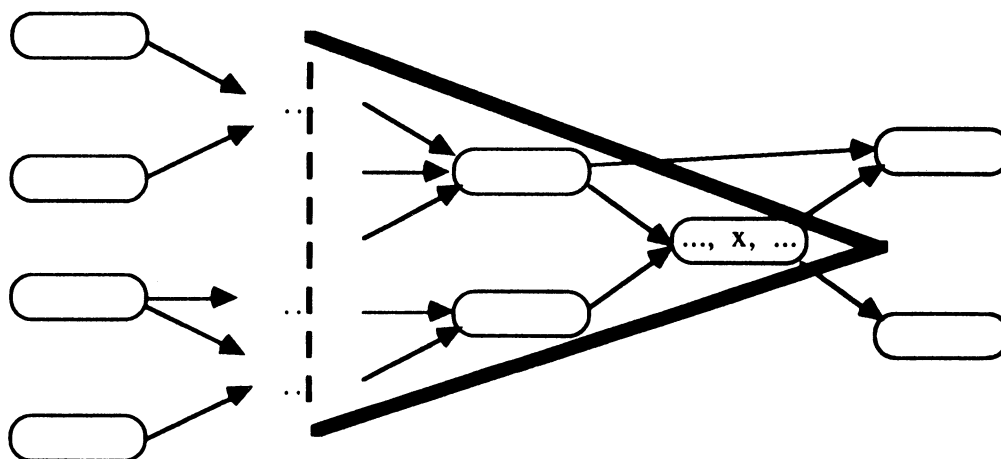
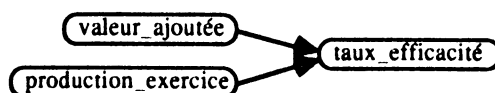


Figure 3.7 : base d'explication de la variable x

Cette définition est insuffisante, car elle peut conduire à des explications incomplètes. Le sous-graphe: $\text{valeur_ajoutée} \rightarrow \text{taux_efficacité}$ est en effet une base d'explication de la variable taux_efficacité selon la définition précédente. Or elle est incorrecte car elle conduit à des explications ne mentionnant pas l'effet de la variable $\text{production_exercice}$. La condition suivante doit donc être ajoutée: si une variable a l'une de ses causes dans la base, ait toutes ses autres causes doivent également y figurer. De plus, l'indépendance causale des variables explicatives est également nécessaire. En effet, la base:



n'a pas le défaut de la base précédente, mais elle n'en est pas moins inadéquate. Elle fait ressortir deux causes indépendantes à l'écart de la variable `taux_effacité` alors que ces causes sont liées. Ces remarques nous amènent à formuler la définition suivante:

Définition 2

Soit M un modèle et $G_C = (S, C)$ son graphe causal. La base d'explication $G_B = (E, B)$ d'une variable x est **complète** si et seulement si elle vérifie les deux conditions suivantes:

- $\forall s_1 \in E, \forall s_2 \in E, \forall s_3 \in S, (s_1, s_2) \in B \wedge (s_3, s_2) \in C \Rightarrow (s_3, s_2) \in B$
- $\forall s_1 \in E, \forall s_2 \in E, \forall s_3 \in S, s_3 \in \text{ancêtres}_C(s_2) \cap \text{descendants}_C(s_1) \Rightarrow s_3 \in E$, où $\text{ancêtres}_C(s_2) \cap \text{descendants}_C(s_1)$ est l'ensemble des points situés sur un chemin allant de s_1 à s_2 dans G_C .

La première condition assure que **toutes** les causes d'une variable figurent dans sa base. La seconde condition vérifie que si deux variables de la base sont liées par un chemin de causalité, alors les variables intermédiaires du chemin appartiennent également à la base.

Il existe généralement plusieurs bases complètes pour une variable. Aussi, le choix d'une base d'explication doit-il se porter dans cet ensemble. Ce choix est guidé par l'objectif de simplicité de l'explication; il nous amène à considérer, pour expliquer une variable, sa plus petite base complète.

Définition 3

Une base d'explication complète d'une variable x , $G_B = (E, B)$ est **minimale** si et seulement si :

- Pour tout $E' \subsetneq E$, $\Gamma_{E'}(G_B)$ n'est pas une base d'explication complète de x . $\Gamma_{E'}(G_B)$ est le sous-graphe de G_B engendré par E' .

Il est intéressant de noter le résultat suivant: toute variable endogène admet une et une seule base d'explication complète minimale. Le graphe causal étant sans circuit, cette propriété peut aisément être démontrée par récurrence sur le nombre de sommets.

La figure 3.8 présente la base d'explication complète minimale de la variable `taux_effacité` dans le système d'équation (S_1) . La variable `production_globale` se trouvant sur le chemin $\{\text{production_exercice}, \text{production_globale}, \text{valeur_ajoutée}\}$, elle doit figurer dans la base. De ce fait, la variable `valeur_ajoutée` a maintenant une de ses causes dans la base; l'autre,

consommations_intermédiaires doit donc être ajoutée. De même, l'autre cause de la variable production_globale: marge_commerciale, doit être ajoutée. La base est alors complète et minimale.

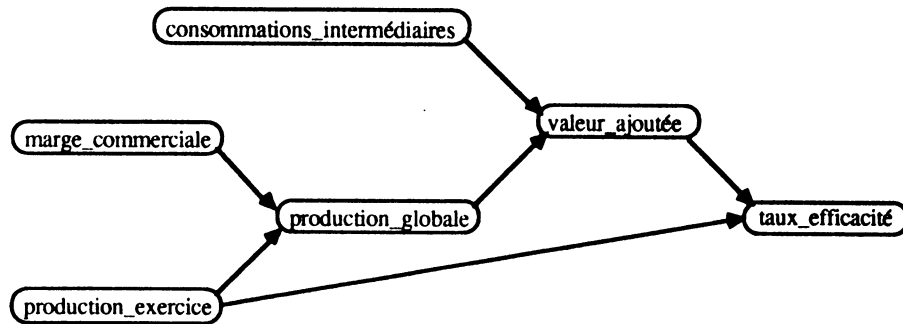


figure 3.8: la base d'explication complète minimale de la variable taux_effacité

Un algorithme de calcul de la base d'explication complète minimale d'une variable sera présenté aux chapitre 4.

2.2.3 - Présentation de la base d'explication

Au cours de cette étape, un texte en français analysant les causes de l'écart de la variable est généré à partir de sa base. Cette tâche peut être illustrée sur l'exemple suivant: imaginons ainsi un commerçant dont les dépenses et les recettes augmentent tandis que son bénéfice diminue. Une interprétation (naïve) de sa situation est: "le bénéfice a diminué à cause de l'augmentation des dépenses, malgré l'augmentation des recettes". Cet exemple met en évidence la nécessité de distinguer deux types de variables explicatives: celles ayant un rôle moteur (dépenses) sur la variable expliquée, et celles jouant le rôle de frein (recettes).

L'écart d'une variable entre les deux états est déterminé par l'ensemble de ses variables explicatives. Or, pour effectuer la distinction précédente, il est nécessaire de connaître l'effet **individuel** d'une variable. Ce problème n'est pas sans rappeler ceux de l'analyse de sensibilité en Mathématiques et de l'analyse comparative ([WELD 88]) en Intelligence Artificielle. Nous résolvons ce problème en introduisant deux relations appelées "confirme" et "infirme", définies de la manière suivante:

Définition 4

Soit y une variable endogène et x une variable appartenant à sa base d'explication. On suppose l'ensemble des valeurs de y totalement ordonné (par la relation notée " $<$ "). On dira que:

- x **confirme** y si et seulement si $[(y^d < y^r) \wedge (y^{\Delta x} < y^r)] \vee [(y^r < y^d) \wedge (y^r < y^{\Delta x})]$
- x **infirme** y si et seulement si $[(y^d < y^r) \wedge (y^r < y^{\Delta x})] \vee [(y^r < y^d) \wedge (y^{\Delta x} < y^r)]$

y^r et y^d sont les valeurs de y dans l'état de référence et dérivé ; $y^{\Delta x}$ est la valeur qu'aurait y dans l'état dérivé si toutes les variables précédant x dans le graphe causal gardaient leur valeur de l'état de référence.

L'effet d'une variable est donc décidé en fonction de la réponse à la question: si x était la seule variable modifiée entre les deux états, dans quel sens varierait la variable y ? si y varie dans le même sens que celui observé lorsque toutes les variables explicatives sont modifiées, on considère alors que x est une cause motrice de y , on dit que x confirme y . Dans le cas contraire, on estime que x contribue à limiter l'écart de y : x infirme y . Notons que la relation d'ordre total nécessaire à cette distinction, n'implique pas que les variables soient numériques. En particulier, on pourra considérer pour le raisonnement qualitatif l'ensemble $\{-, 0, +\}$ ordonné par la relation " $<$ " telle que: $- < 0$ et $0 < +$.

Dans l'exemple précédent, si les recettes avaient augmenté seules, le bénéfice aurait également augmenté; l'augmentation des recettes infirme donc la diminution du bénéfice. Par contre, si les dépenses avaient augmenté seules, le bénéfice aurait diminué, comme c'est le cas dans la situation observée; l'augmentation des dépenses confirme donc la diminution du bénéfice.

Nous verrons dans les deux chapitres suivants que les techniques de simulation de l'effet d'une variable explicative sont très différentes pour les modèles qualitatifs et numériques.

Une fois cette distinction opérée, le texte de l'explication d'une variable est généré par un parcours de sa base d'explication, en utilisant le vocabulaire du domaine d'application.

3 - Comparaison avec d'autres travaux

L'interprétation comparative présente une grande similitude avec le diagnostic à base de modèles. Ainsi, le schéma du diagnostic de J. De Kleer et B. Williams (figure 3.9), n'est pas sans rappeler celui de la figure 3.2. Dans les deux cas, il s'agit de comparer deux états d'un système et d'expliquer la cause

de leurs différences. Le lien entre interprétation comparative et diagnostic est renforcé par l'utilisation du même mode d'inférence: le **raisonnement abductif**.

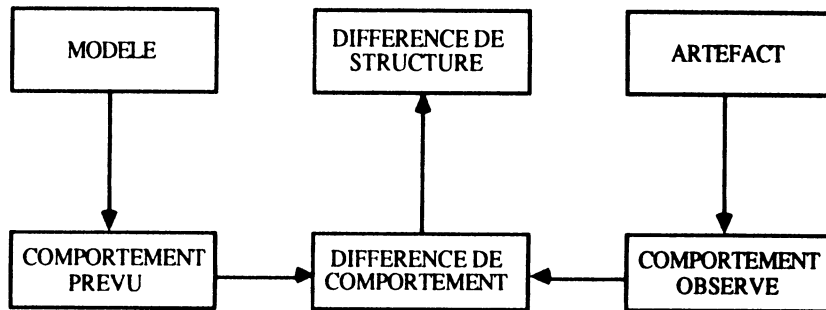


Figure 3.9: le diagnostic, d'après [DE KLEER 87]

Un raisonnement abductif est du type : "de $P \Rightarrow Q$ et Q , je peux déduire: P est possible". Ce mode d'inférence est particulièrement adapté à la construction d'un ensemble d'hypothèses permettant d'expliquer des observations. En diagnostic, il permet de formuler un ensemble d'hypothèses de pannes comme le montre l'exemple 4.

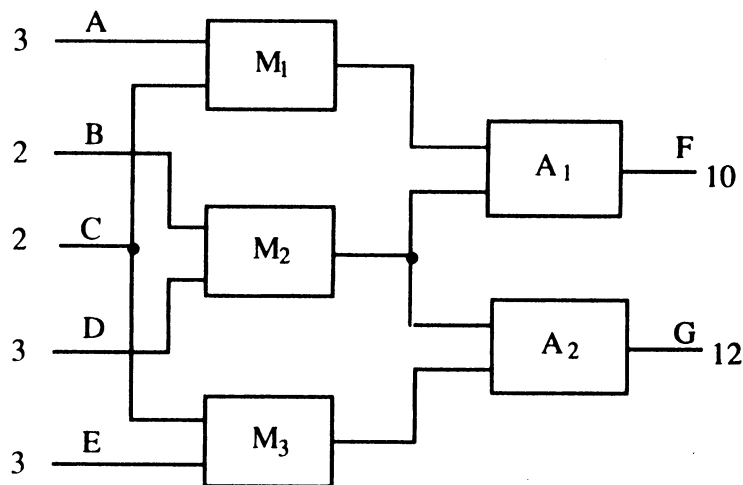


Figure 3.10 : un circuit familier, d'après [DAVIS 84]

Exemple 4

Considérons le circuit de la figure 3.10, où M_1 , M_2 et M_3 sont des multiplieurs tandis que A_1 et A_2 sont des additionneurs. Ayant $F = A \times C + B \times D$, la sortie F devrait être égale à 12. La sortie observée étant 10, le circuit fonctionne incorrectement. Déterminer les pannes possibles avec un raisonnement

abductif revient à poser la question suivante: quels sont les ensembles de composants qui, s'ils étaient défectueux expliqueraient la valeur incorrecte de F ? La réponse à cette question est: $\{A_1\}$, $\{M_1\}$, $\{A_2, M_2\}$ et $\{M, M_3\}$ (M_2 ne peut pas être incriminé seul sinon la sortie G serait incorrecte).

Poursuivant l'analogie, au cours de l'explication comparative, on pose la question: quelles sont les variables qui, si elles agissaient seules expliqueraient l'écart observé ?

Il existe cependant deux différences importantes entre diagnostic et interprétation comparative. Premièrement, les différences entre les deux états sont connues *a priori* en diagnostic: elles se manifestent sous forme de symptômes, tandis qu'elles doivent être mises en évidence en interprétation comparative: c'est le rôle de l'évaluation comparative. Deuxièmement, le diagnostic recherche les causes d'un écart dans une altération du modèle. Ainsi, les hypothèses de pannes expliquent le non respect de l'égalité $F = A \times C + B \times D$. A l'inverse, l'explication comparative suppose le modèle identique entre les deux états et recherche la cause de l'écart dans une modification des données. Les deux techniques apparaissent donc, non pas totalement similaires, mais complémentaires.

Une méthode générale pour l'explication des différences entre deux situations utilisant le raisonnement abductif a été présentée dans ce chapitre. Nous proposons dans les deux chapitres suivants des techniques permettant de l'appliquer à des modèles qualitatifs puis numériques.

CHAPITRE 4

EXPLICATION COMPARATIVE DANS UN MODELE QUALITATIF

Un point rarement mis en évidence dans la littérature en Intelligence Artificielle est que le raisonnement qualitatif résulte de deux besoins différents: la manipulation de **données qualitatives** à partir de modèles numériques et le raisonnement sur des **relations qualitatives**.

Dans des domaines bien formalisés où les systèmes étudiés ont des comportements **prédictibles**, les modèles quantitatifs sont bien adaptés. Si le raisonnement qualitatif y est nécessaire, c'est pour manipuler des données d'un niveau d'abstraction plus élevé que les valeurs numériques utilisées par ces modèles. Dans cette démarche, un modèle qualitatif découle d'une traduction directe des équations différentielles décrivant le système.

Dans les disciplines moins bien formalisées, la situation est différente. Si les modèles quantitatifs ont également investi ces domaines, leur adéquation y est beaucoup moins évidente. Ainsi, aucun modèle économique n'a pu prévoir la crise de 1929, les chocs pétroliers ou le krach boursier de 1988. Il ne faut donc pas s'étonner de l'intérêt porté dans ces domaines aux représentations qualitatives. Quand l'existence même de relations est parfois incertaine, il semble en effet sage d'abandonner la précision illusoire des modèles numériques au profit d'une modélisation de l'essence des phénomènes. Cet intérêt pour le raisonnement qualitatif relève donc d'un ordre différent de celui en Physique qualitative, car il résulte du besoin de manipuler des relations plus abstraites que les équations numériques. C'est ainsi qu'en Sociologie ([HEISE 75]), en Gestion ([LESCA 86]), en Ecologie ([PUCCIA 85]) et d'autres domaines sont utilisés des modèles basés sur des relations qualitatives.

Parmi ces modèles, les réseaux d'influences (ou graphes d'influences) occupent une place centrale. Sous les différentes aspects qu'ils revêtent, ils permettent un large échantillon de raisonnements. Dans leur forme la plus simple, ils indiquent seulement les variables du domaine qui

interagissent. Plus complexes, ils représentent la condition, la probabilité, l'intensité,... de ces interactions. Ce chapitre est consacré à l'implantation de l'explication comparative dans ce type de modèles.

Les caractéristiques de l'explication comparative dans un type très simple de réseaux d'influences sont d'abord exposées en section 1. Puis les trois étapes de son traitement sont examinées dans les sections suivantes. Enfin, les extensions des techniques présentées à des réseaux d'influences plus complexes sont analysées en conclusion.

1 - Explication comparative dans un réseau d'influences

Comme nous l'avons vu au chapitre 2, un réseau d'influences exprime sous sa forme la plus simple, des dépendances qualitatives entre des variables par un ensemble de relations:

$$I+ (x y) \text{ ou } I- (x y)$$

signifiant: la variable x influence positivement (resp. négativement) la variable y si une augmentation de x entraîne une augmentation (resp. diminution) de y , et une diminution de x entraîne une diminution (resp. augmentation) de y .

Si les relations sont qualitatives, les variables d'un réseau d'influences peuvent être soit toutes numériques, toutes symboliques, ou des deux types. Dans les deux derniers cas, le domaine des symboles de chaque variable doit être totalement ordonné, afin que les termes augmentation et diminution aient une signification. Cette relation d'ordre sera notée " $<$ " aussi bien pour les nombres que pour les symboles.

L'écart d'une variable x dont l'explication comparative recherche les causes, est noté $\Delta_q x$ pour un réseau d'influences. C'est un symbole appartenant à l'ensemble $\{+, 0, -\}$, défini de la manière suivante:

$$\begin{aligned}\Delta_q x &= + \text{ si } x^r < x^d \\ \Delta_q x &= 0 \text{ si } x^d = x^r \\ \Delta_q x &= - \text{ si } x^d < x^r\end{aligned}$$

où x^r et x^d sont les valeurs de x dans l'état de référence et dérivé.

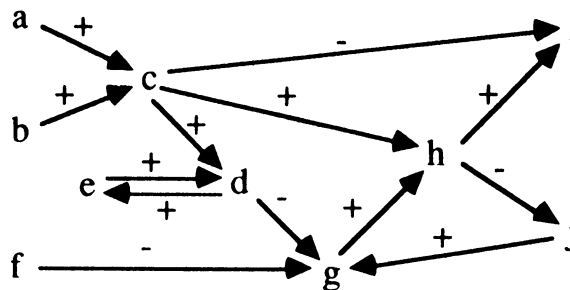


Figure 4.1 : un exemple de réseau d'influences

L'étude des trois étapes de l'explication comparative dans un réseau d'influences s'appuie tout au long de ce chapitre sur l'exemple de la figure 4.1. Il faut noter dans ce réseau la présence de deux circuits: le premier entre les variables g, h et j; le second entre d et e. Un circuit traduit une boucle de rétroaction entre des variables. Ce phénomène est extrêmement fréquent dans les mécanismes complexes. Ainsi, le budget publicité a une influence positive sur les ventes, qui ont une influence positive sur les bénéfices, qui en retour ont une influence positive sur le budget publicité.

2 - La construction du graphe causal

La première étape de l'explication comparative est la construction d'un graphe causal associé au modèle (cf. chapitre 3). Ce graphe met en évidence les relations causales existant entre les variables. Dans les réseaux d'influences, cette tâche est simplifiée car les arcs d'influences sont des liens de causalité. Il suffit donc de rechercher les circuits dans le réseau pour en déterminer l'ordre causal.

La distinction entre variables endogènes et exogènes, qui sert à propager les valeurs des variables à travers le modèle est également particulière. Pour un modèle numérique, nous verrons qu'elle doit être opérée explicitement par le concepteur du modèle. A l'inverse, dans un réseau d'influences, elle est implicite: une variable exogène constitue une source du graphe (un sommet de demi-degré¹ intérieur nul); les autres variables sont endogènes. Ainsi, les variables a, b et f du réseau de la figure 4.1 sont exogènes, tandis que c, d, e, g, h, i et j sont endogènes.

¹La terminologie de la Théorie des Graphes que nous utilisons est issue de [BERGE 83]

Le sens des arcs d'influences indique l'ordre de résolution des variables. Les variables déterminées simultanément constituent donc des composantes fortement connexes (CFC) du réseau d'influences. Pour cette raison, le graphe causal correspond donc au **graphe réduit** du réseau. On rappelle que le graphe réduit G_r d'un graphe orienté G est le quotient de G par la relation de forte connexité. Les sommets de G_r sont donc les CFC du réseau d'influences (i.e. les ensembles de variables liées par un circuit) et il existe un arc entre deux CFC s_i et s_j de G_r , s'il existe au moins un arc d'influence entre une variable de s_i et une variable de s_j .

La construction du graphe réduit du réseau d'influences peut s'effectuer en deux étapes. Les CFC sont tout d'abord calculées, puis les arcs entre les CFC sont ajoutés.

Il existe de nombreux algorithmes pour le calcul des CFC d'un graphe. Celui de Tarjan ([TARJAN 72]) fait partie des plus performants; sa complexité est de l'ordre du nombre d'arcs du graphe. Etant utilisé dans CRESUS, il est présenté ci-dessous.

L'algorithme de Tarjan détermine les CFC par un parcours "en profondeur d'abord" du graphe. Les sommets visités sont numérotés dans l'ordre du parcours par un premier indice: $\text{num}(v)$. A chaque sommet est associé un second indice: $\text{inf}(v)$, qui est le plus petit numéro des sommets accessibles par un arc depuis l'un des successeurs de v , dans l'arborescence générée par le parcours. Tarjan montre que le calcul d'une CFC se ramène à la recherche de sa racine, l'équivalent du point d'articulation pour un graphe non orienté. Il démontre ensuite que r est la racine d'une CFC si et seulement si $\text{num}(r) = \text{inf}(r)$. Les autres sommets de la CFC sont les descendants v de r non encore intégrés dans une CFC tels que: $\text{num}(v) > \text{num}(r)$. D'où l'algorithme:

```

procédure cfc(v) /*recherche la CFC du sommet v*/
  début
    i := i + 1; num(v) := i; inf(v) := i;
    empiler(v);
    pour tout w ∈ successeurs(v) faire
      début
        si non(num(w)) alors
          début
            cfc(w);
            inf(v) := min(inf(v), inf(w))
          fin
        sinon
          si num(v) < num(w) alors
            si w ∈ pile alors inf(v) := min(inf(v), num(w))
          fin;
      si inf(v) = num(v) alors /*v est la racine d'une CFC*/
        début

```

```

composante-fortement-connexe(v) :=  $\emptyset$  ;
dépiler(w);
tant que num(w)  $\geq$  num(v) faire           /*on ajoute les autres sommets de la CFC*/
  début
    composante-fortement-connexe(v) := composante-fortement-connexe(v)  $\cup$  {w};
    si pile  $\neq \emptyset$  alors dépiler(w) sinon num(w) := 0;
  fin;
si num(w)  $\neq$  0 alors empiler(w)
fin
fin;

```

La procédure est lancée par la séquence d'instructions:

```

début
  pile :=  $\emptyset$  ;
  i := 0;
  pour tout sommet w faire
    si non(num(w)) alors cfc(w)
fin;

```

L'algorithme pour déterminer l'ensemble des arcs du graphe réduit est immédiat:

```

début
  arcs :=  $\emptyset$ ;
  pour toute CFC u faire
    début
      pour toute CFC v faire
        début
          si u  $\neq$  v et ( $\exists x \in u$  et  $\exists y \in v : I+(x y)$  ou  $I-(x y)$ ) alors arcs := arcs  $\cup$  {(u,v)}
        fin
      fin
    fin
fin;

```

L'application successive de ces deux algorithmes sur le réseau d'influences de la figure 4.1 produit le graphe causal de la figure 4.2.

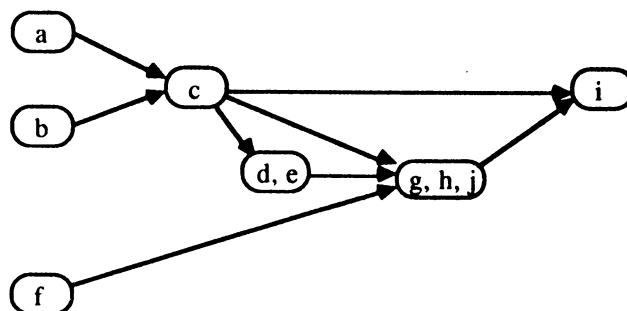


Figure 4.2: le graphe causal du réseau d'influences de la figure 4.1

3 - Construction de la base d'explication complète minimale

Une fois le graphe causal construit, la base d'explication complète minimale (BECM) de la variable expliquée est recherchée. Rappelons que la BECM est le petit sous-graphe du graphe causal conduisant à une explication complète. Un algorithme a été développé pour la construire; il est présenté dans cette section.

La procédure *becm* ci-dessous construit itérativement l'ensemble dénommé "sommets" des CFC de la BECM. L'ensemble des arcs de la BECM est obtenu à la fin de la procédure à partir du sous-graphe engendré par l'ensemble des sommets de la base. L'algorithme construit à chaque pas de l'itération un sous-ensemble de sommets de la base appelé "support", contenant uniquement les CFC non «expansées» (i.e. dont aucune cause n'est dans la base). Il suffit de vérifier à chaque itération l'indépendance causale entre les éléments du support: si elle est assurée, la base est complète. Sinon, un chemin lie deux sommets du support, et les CFC intermédiaires de ce chemin et toutes leurs causes doivent être ajoutées à la base. Cet algorithme peut s'écrire de la manière suivante:

```

procédure becm(x)      /*détermine la base d'explication complète minimale de la variable x */
  début
    u := composante-fortement-connexe(x);
    support := prédécesseursC(u);      /*on considère d'abord les prédécesseurs directs de x*/
    sommets := {u} ∪ prédécesseursC(u);      /*sommets mémorise les sommets de la becm*/
    si ∃v ∈ support, w ∈ support: v ∈ ancêtresC(w)
      alors chemin := {(v,w)}                      /*il y a un chemin entre v et w*/
      sinon chemin := ∅;                          /*pas de chemin entre les éléments du support*/
    tant que chemin ≠ ∅ faire
      début
        intermédiaires := (ancêtresC(w) ∪ {w}) ∩ descendantsC(v);          /*cf. rem. 2*/
        support := (support - {w}) ∪ [(∪w ∈ intermédiaires prédécesseursC(w)) - sommets];      /*cf. rem. 3*/
        sommets := sommets ∪ (∪w ∈ intermédiaires prédécesseursC(w));
      si ∃v ∈ support, w ∈ support: v ∈ ancêtresC(w);
      alors chemin := {(v,w)}                      /*reste-t-il des chemins dans le support?*/
      sinon chemin := ∅
    fin;
    arcs := {(v,w): v ∈ sommets, w ∈ sommets}
  fin;
  
```

Cet algorithme appelle quelques remarques:

1. Les relations sur les graphes indicées par la lettre "C" sont appliquées dans le graphe causal et non dans la base d'explication.
2. Lorsqu'un chemin lie deux CFC v et w du support, toutes les CFC intermédiaires de tous les chemins entre v et w sont ajoutées à la base. Cet ensemble de CFC est l'intersection des ancêtres de l'extrémité terminale du chemin et des descendants de l'extrémité initiale.
3. A chaque itération, le nouveau support intègre les prédécesseurs des CFC intermédiaires du chemin. Deux cas sont à considérer pour un prédécesseur: ou bien il n'appartient pas encore à la base; il n'a donc jamais été expansé et doit être ajouté au support. Ou bien, il appartient déjà à la base; dans ce cas, un arc incident **extérieur** lui a été ajouté, ce qui ne modifie en rien son appartenance au support par rapport à l'itération précédente. Le calcul du support dépend de l'ensemble de sommets obtenus à l'itération précédente, il doit donc être effectué avant celui du nouvel ensemble de sommets. Enfin, l'extrémité terminale du chemin doit être enlevée du support, d'où l'expression de calcul du support.
4. La procédure devient terminale. En effet, au moins une CFC est expansée à chaque itération: l'extrémité terminale du chemin. De plus, le support est construit de sorte que chaque CFC est expansée une seule fois. Le nombre d'itérations est donc majoré par le nombre d'ancêtres de la CFC de la variable expliquée dans le graphe causal.
5. La base calculée est complète. Les prédécesseurs d'une CFC sont toujours ajoutés simultanément, il ne peut donc pas manquer de cause à une variable dans la base. Par ailleurs, si un chemin lie deux CFC, il appartient nécessairement à la base construite par l'algorithme. Pour s'en convaincre, il suffit de montrer que tous les sommets intermédiaires de ce chemin appartiennent à la base (les arcs de la base sont obtenus à la fin de l'algorithme en prenant tous les arcs du graphe causal liant les sommets de la base). Supposons qu'il existe un chemin liant deux sommets v et w de la base, et z un sommet du graphe causal appartenant à ce chemin mais pas à la base. Montrons que cette hypothèse aboutit à une contradiction. A la fin de l'algorithme, ou bien w appartient au support (il n'a jamais été expansé), ou bien tous ces prédécesseurs sont dans la base (il a été expansé). Si w n'appartient pas au support, l'un au moins de ces prédécesseurs, w_1 , est sur le chemin $\{z, \dots, w\}$. Le même raisonnement est applicable à w_1 . Donc, ou bien w_1 appartient au support, ou bien tous ces prédécesseurs sont dans la base. Comme z n'appartient pas au support, ce processus ne peut être répété indéfiniment. Il existe donc w_i sur le chemin $\{z, \dots, w\}$ tel que w_i appartient au

support. Par ailleurs, un raisonnement similaire sur v montre que v ou l'un de ses ancêtres appartient au support. Or, ceci est impossible, sinon il existerait un chemin entre ce sommet et w_i , deux éléments du support, non examiné par l'algorithme. D'où la contradiction.

6. La base calculée est minimale: seuls les sommets nécessaires à la complétude de la base sont ajoutés à chaque étape.
7. Evaluons la complexité de l'algorithme. On peut supposer les ancêtres et les descendants des CFC du graphe causal déterminés une fois pour toutes dans un pré-calcul. Le temps d'exécution d'une itération de la boucle dans la procédure `becm` devient alors constant: k . Comme le nombre d'itérations est majoré par le nombre n de sommets du réseau (cf. remarque 4), le temps d'exécution de la boucle est majoré par kn .

Illustrons l'algorithme par le calcul de la BECM de la variable h du réseau d'influences de la figure 4.1.

becm(h)

initialement:

$u := [g, h, j]$ (les CFC sont notées entre $[]$ pour éviter les confusions avec les ensembles de CFC)

support := $\{[c], [d, e], [f]\}$

sommets := $\{[c], [d, e], [f], [g, h, j]\}$

chemin := $([c], [d, e])$

première itération:

intermédiaires := $(\{[a], [b], [c]\} \cup \{[d, e]\}) \cap \{[d, e], [g, h, j], [i]\} = \{[d, e]\}$

support := $\{[c], [f]\} \cup \emptyset = \{[c], [f]\}$

sommets := $\{[c], [d, e], [f], [g, h, j]\} \cup \emptyset = \{[c], [d, e], [f], [g, h, j]\}$

chemin := \emptyset

fin:

arcs := $\{([c], [d, e]), ([c], [g, h, j]), ([d, e], [g, h, j]), ([f], [g, h, j])\}$

d'où la BECM de la variable h présentée sur la figure 4.3.

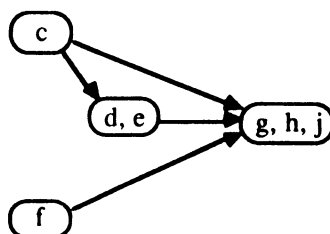


Figure 4.3: la BECM de la variable h du réseau d'influences de la figure 4.1

Il subsiste un problème lorsque la BECM d'une variable est calculée: comment traiter les boucles de rétroaction, c'est-à-dire les variables à l'intérieur d'une même CFC ? Deux stratégies opposées sont envisageables. On peut tout d'abord considérer que ces variables étant interdépendantes, elles ne doivent pas s'expliquer mutuellement. Cette décision conduit à expliquer les variables d'une CFC par les variables des CFC prédécesseurs dans le graphe causal. Le point de vue inverse est également défendable: puisque le concepteur du réseau a spécifié des arcs particuliers à l'intérieur d'une CFC, il est logique de les utiliser pour l'explication. Dans ce cas, les variables s'expliquent mutuellement à l'intérieur d'une CFC.

Nous aurions souhaité adopter la seconde solution. Ce choix se justifie par le fait que dans le langage courant, les explications négligent fréquemment les boucles de rétroaction. On dit par exemple: "les ventes ont progressé grâce à l'augmentation du budget publicitaire", ignorant par là l'effet en retour de l'augmentation des ventes sur le budget publicitaire. L'implantation de cette solution est néanmoins délicate; aussi une solution intermédiaire entre les deux alternatives ci-dessus a dû être adoptée.

Expliquer mutuellement les variables liées par une boucle de rétroaction nécessite de supprimer certains arcs à l'intérieur d'une CFC, inutiles pour l'explication (comme la séquence d'arcs allant des ventes au budget publicité). Le problème qui se pose alors est le choix des arcs pouvant être supprimés. Illustrons ce point sur le réseau de la figure 4.4 qui comporte un circuit entre les variables x_1 , x_2 , x_3 et x_4 (les signes des arcs d'influences ont été omis pour simplifier). Quels arcs peut on éliminer pour construire la base d'explication de la variable y ? La réponse à cette question réside partiellement dans la remarque suivante: l'arc (x_3, x_4) est inutile pour l'explication de la variable y car aucun chemin élémentaire allant d'une variable précédant la CFC $\{x_1, x_2, x_3, x_4\}$ à la variable y , n'emprunte sa variable terminale: x_4 . Son seul rôle est donc le retour de la boucle de rétroaction de la variable x_3 à x_1 . Il peut donc être supprimé. A l'inverse, les autres arcs du circuit sont nécessaires à l'explication: (x_1, x_2) et (x_2, x_3) parce qu'ils appartiennent au chemin $\{x, x_1, x_2, x_3, y\}$ et (x_4, x_1) parce qu'il est un prédécesseur de x_1 (quand x_1 sera expansée, x_4 sera introduit dans la base).

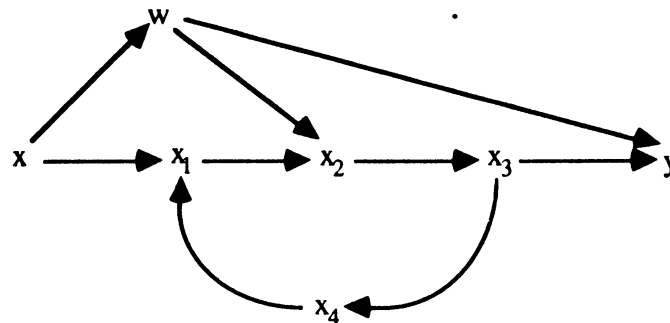


Figure 4.4: le problème de la suppression des arcs pour l'explication des variables à l'intérieur d'une CFC

En généralisant cette idée, on obtient la condition de suppression d'un arc. Soient C une CFC d'un réseau et P l'ensemble des prédécesseurs des variables de C . Soient y la variable à expliquer et (a,b) un arc entre deux variables de C . Si pour toute variable p de P , aucun chemin élémentaire de p à y ne passe par b , alors l'arc (a,b) peut être supprimé du réseau.

Cependant cette remarque ne résout pas totalement le problème car, après la suppression des arcs vérifiant la propriété ci-dessus, il peut rester des circuits dans le réseau. Ainsi, s'il existait un arc (x_4,x_2) dans l'exemple de la figure 4.4, l'arc (x_3,x_4) ne pourrait pas être supprimé (car x_4 serait alors nécessaire à l'expansion de x_2), et le circuit (x_1,x_2,x_3,x_4) subsisterait. C'est pourquoi, une solution intermédiaire pour l'explication mutuelle des variables d'une CFC doit être adoptée. La technique précédente est d'abord appliquée afin de supprimer le maximum de boucles de rétroaction, puis la BECM est calculée à partir du graphe causal résultant en appliquant la procédure becm.

Appliquons cette méthode au réseau de la figure 4.1. L'arc (d,e) peut être supprimé car aucun chemin élémentaire allant de c (l'unique prédécesseur de la CFC $[d,e]$) à h ne passe par e . Pour la même raison, l'arc (h,j) peut être éliminé. Il ne subsiste donc plus de boucle de rétroaction dans le réseau. L'application de la procédure becm sur le réseau ainsi modifié produit alors pour la variable h , la BECM présentée sur la figure 4.5. Le texte de l'explication est ensuite généré à partir de cette BECM.

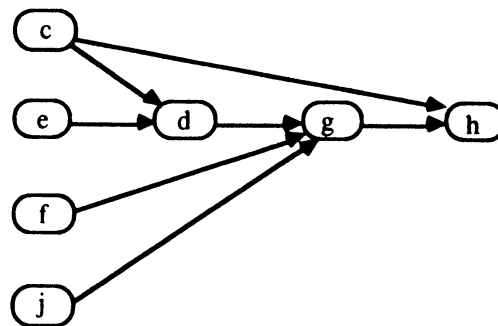


Figure 4.5: la BECM modifiée de la variable h

4 - Présentation de la base d'explication

La dernière étape de l'explication d'une variable est la présentation de sa base modifiée sous forme d'un texte. Nous avons vu au chapitre 3 que cette tâche repose sur la construction de deux relations, "confirme" et "infirme", indiquant l'effet d'une variable sur une autre. Dans un réseau d'influences, cet effet est calculable à partir des signes des arcs d'influence. Pour cela, on définit la **signature** d'un chemin. On appelle signature d'un arc (x,y) , le réel noté $\sigma(x,y)$ valant +1 si $I+(x,y)$ et -1 si $I-(x,y)$. La signature d'un chemin $C = \{x_1, x_2, \dots, x_n\}$ est définie par:

$$\sigma(C) = \prod_{i=1}^{n-1} \sigma(x_i, x_{i+1})$$

La signature détermine l'effet global de la variable initiale du chemin sur la variable terminale car une séquence d'arcs $x_1 \xrightarrow{\sigma_1} \dots \xrightarrow{\sigma_{n-1}} x_n$ de signature σ a un effet équivalent à un seul arc $x_1 \xrightarrow{\sigma} x_n$. Cependant, l'existence de plusieurs chemins entre deux variables complique la situation. Deux cas peuvent se présenter:

- tous les chemins sont de même signature σ . Dans ce cas, l'effet de la variable explicative sur la variable expliquée se résume à un seul arc de signature σ .
- il existe des chemins de signatures différentes. Comme il est impossible de déterminer lesquels ont le plus d'impact, l'effet est indéterminé. On retrouve dans ce cas l'ambiguïté de la simulation qualitative inhérente à la pauvreté d'information de l'espace $\{+,0,-\}$. Il existe cependant une exception notable à ce cas: celui des boucles de rétroaction. Lorsqu'une boucle de rétroaction de

signature négative se trouve sur un chemin, elle introduit toujours des chemins de signatures contraires entre les deux extrémités. Considérons l'exemple de la figure 4.5, comprenant une boucle de rétroaction entre les variables x_1 , x_2 , x_3 et x_4 . Il existe une infinité de chemins allant de x à y : $\{x, x_1, x_2, x_3, y\}$, $\{x, x_1, x_2, x_3, x_4, x_1, x_2, x_3, y\}$, etc. Si la signature du premier est positive, celle du second est en revanche négative. L'indétermination peut néanmoins être levée; la raison en est que l'explication comparative analyse des états d'équilibre. Si l'effet de retour d'une boucle de rétroaction sur une variable était supérieur à l'effet direct, des états instables seraient induits. Il est donc légitime de négliger l'effet des boucles de rétroaction. Ceci revient à ne prendre en compte que les chemins élémentaires pour le calcul de la signature. Ainsi, sur la figure 4.6, la variable x a un effet positif sur y car il existe un seul chemin élémentaire entre ces deux variables.

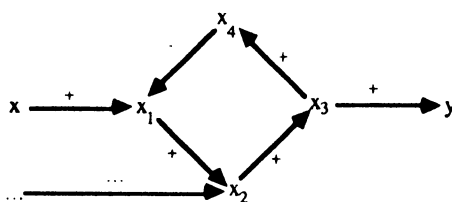


Figure 4.6: cas des boucles de rétroaction

A partir de la notion de signature, les deux relations confirme et infirme définies dans le cas général au chapitre 3, s'écrivent pour les réseaux d'influences:

x **confirme** y si et seulement si, pour tout chemin élémentaire C allant de x à y ,

$$\left\{ \begin{array}{l} \alpha(C) = +1 \\ (\Delta_Q y = + \text{ et } \Delta_Q x = +) \text{ ou } (\Delta_Q y = - \text{ et } \Delta_Q x = -) \end{array} \right. \quad \text{ou} \quad \left\{ \begin{array}{l} \alpha(C) = -1 \\ (\Delta_Q y = + \text{ et } \Delta_Q x = -) \text{ ou } (\Delta_Q y = - \text{ et } \Delta_Q x = +) \end{array} \right.$$

x **infirme** y si et seulement si, pour tout chemin élémentaire C allant de x à y ,

$$\left\{ \begin{array}{l} \alpha(C) = -1 \\ (\Delta_Q y = + \text{ et } \Delta_Q x = +) \text{ ou } (\Delta_Q y = - \text{ et } \Delta_Q x = -) \end{array} \right. \quad \text{ou} \quad \left\{ \begin{array}{l} \alpha(C) = +1 \\ (\Delta_Q y = + \text{ et } \Delta_Q x = -) \text{ ou } (\Delta_Q y = - \text{ et } \Delta_Q x = +) \end{array} \right.$$

Sur l'exemple de la figure 4.1, les variables prenant les valeurs du tableau 1, f infirme h car $\{f, g, h\}$ est le seul chemin élémentaire entre f et h , $\sigma(\{f, g, h\}) = -1$, $\Delta_Q f = -$ et $\Delta_Q h = -$. Par contre, l'effet de c sur h est indéterminé car $\alpha\{c, h\} = +1$ et $\alpha\{c, d, g, h\} = -1$. Dans ce cas, l'effet global ne peut pas être précisé et l'explication générée mentionne alors chacun des chemins (cf. l'explication de la variable c sur l'exemple ci-dessous).

Le texte de l'explication est généré à partir d'un parcours de la BECM de la variable expliquée. Le texte produit dépend du domaine d'application car il est nécessaire d'employer le vocabulaire de l'utilisateur.

variable	écart
a	+
b	0
c	+
d	+
e	+
f	-
g	-
h	-
i	-
j	+

Tableau 1

Le réseau d'influences de la figure 4.1 et les écarts qualitatifs du tableau 1 ont été introduits dans CRESUS, pour deux situations différentes. La première pour des variables qualitatives. La comparaison est supposée effectuée entre deux hypothèses dénommées "hypothèse1" et "hypothèse2", représentant respectivement l'état de référence et l'état dérivé du problème. CRESUS explique l'écart de la variable h de la manière suivante:

C et E étant plus élevé dans l'hypothèse 2 que dans l'hypothèse 1, D est plus élevé. D étant plus élevé, G est plus faible bien que F soit plus faible et J plus élevé. G étant plus faible, H est plus faible, malgré l'effet direct de C qui est plus élevé.

La seconde comparaison utilise le même réseau d'influences et des écarts qualitatifs identiques. La comparaison est cette fois temporelle: l'état de référence est l'année 1987 et l'état dérivé, 1988. Les variables sont numériques: les valeurs du tableau 2 sont considérées, avec pour unité commune le kF.

	1987	1988
a	10	11
b	5	5
c	15	17
d	9	23
e	3	5
f	6	5
g	23	20
h	104	100
i	13	12
j	11	12

Tableau 2

CRESUS explique l'écart de la variable h de la manière suivante:

L'augmentation de C de 13,3 % entre 1987 et 1988 et l'augmentation de E de 66,7 % ont entraîné l'augmentation de D de 9 à 23 kF. L'augmentation de D a provoqué la diminution de G de 15,0 % malgré la diminution de F de 16,7 % et l'augmentation de J de 9,1 %. La diminution de G a entraîné la diminution de H de 4,0 %, malgré l'effet direct de l'augmentation de C.

La technique présentée dans ce chapitre a été utilisée sur des modèles de taille réduite. Il nous est donc difficile d'évaluer sa robustesse lorsque la complexité du modèle augmente. Quelques remarques peuvent néanmoins être formulées à ce sujet. Tout d'abord, il faut distinguer deux types de complexité. La première est due à une augmentation du nombre de variables dans le modèle; la seconde à une augmentation du nombre d'interconnexions entre les variables. Il y a lieu de penser que la technique présentée s'accommode assez bien de la première source de complexité. Celle-ci consiste à introduire dans le modèle les mécanismes agissant sur des variables exogènes (rendant ainsi ces dernières endogènes). Cet ajout n'a pour effet d'augmenter la profondeur des explications et n'en dégrade donc pas la compréhensibilité. Il est par contre acquis que l'augmentation de la connexité du réseau d'influences dégrade très rapidement la compréhensibilité des explications. La taille des bases d'explication des variables devient en effet très importante, et les chemins de signature contraire se multiplient; les explications deviennent donc très complexes.

Les techniques décrites jusqu'à présent dans ce chapitre sont applicables aux réseaux d'influences les plus simples. Nous en analysons, pour conclure ce chapitre, les extensions à des modèles plus complexes.

Les réseaux d'influences conditionnels sont constitués d'un ensemble d'expressions du type:

[si <CONDITION> alors] I+ (x y)
ou [si <CONDITION> alors] I- (x y)

(les crochets indiquent que la condition précédant la relation d'influence est optionnelle).

Ces réseaux permettent de décrire des relations telles que: l'endettement d'une entreprise a un effet favorable sur sa rentabilité ... sauf si la rentabilité est négative. Les techniques présentées s'appliquent de manière strictement identique à ces réseaux. Les conditions précédant les relations d'influences sont en effet évaluées avant l'explication comparative de façon à se ramener au cas des réseaux simples.

Les arcs d'un réseau d'influences probabiliste sont quantifiés par des coefficients indiquant la vraisemblance de la relation. Ces réseaux sont fréquemment utilisés dans les domaines comme la Médecine, où la forme des relations entre les variables est mal connue, mais dont l'existence peut être vérifiée statistiquement. L'explication comparative est également applicable à ce type de modèles. Les explications doivent alors être modulées par la vraisemblance des relations: "l'augmentation de la pression artérielle est **probablement** due à ...".

Par contre, les réseaux d'influences temporels, dont les arcs indiquent le temps nécessaire pour la réalisation de la relation d'influence, sont hors de portée de l'explication comparative. Ils s'apparentent en effet aux modèles dynamiques et nécessitent la connaissance des états antérieurs aux deux états comparés.

Quant aux réseaux d'influences dont les arcs sont quantifiés par un coefficient indiquant l'intensité de la relation (HEISE 75]), ils sont équivalents à des systèmes d'équations linéaires. Ils sont étudiés à ce titre dans le chapitre suivant.

CHAPITRE 5

EXPLICATION COMPARATIVE DANS UN MODELE NUMÉRIQUE

L'explication comparative dans un système d'équations présente deux différences importantes par rapport aux réseaux d'influences. La première concerne la construction du graphe causal. Elle est plus complexe car elle doit mettre en évidence les relations de causalité entre les variables. La seconde est liée à la présentation du graphe causal. Alors qu'elle se présente de manière identique dans les réseaux d'influences pour les variables numériques et qualitatives, il en va autrement dans un système d'équations. Lorsque les variables sont qualitatives, les équations doivent être transformées en relations d'influences. Les techniques présentées au chapitre précédent lui sont alors appliquées. La même méthode pourrait être appliquée aux variables numériques en transformant au préalable les valeurs numériques en valeurs qualitatives. Ce n'est pourtant pas la démarche adoptée, car la perte d'information résultant du passage du numérique au qualitatif est trop importante. Au lieu de cela, l'explication est construite à partir des valeurs numériques en introduisant une notion voisine de la dérivée partielle appelée **responsabilité**. Cette dernière permet de quantifier l'effet d'une variable sur une autre; elle est donc plus puissante que la signature. De ce gain d'information résulte des explications plus précises.

Ce chapitre suit une présentation identique au précédent. Les caractéristiques de l'explication comparative dans un système d'équations sont d'abord présentées. Puis les trois étapes qui la composent sont examinées dans les sections suivantes. Notre approche est enfin comparée à d'autres travaux dans le domaine.

1 - Explication comparative dans un système d'équations

La plupart des travaux sur le raisonnement qualitatif sont orientés vers la simulation. Ce pôle de recherche justifie l'étude d'une classe de modèles permettant de représenter la dynamique: les systèmes d'équations différentielles. L'ambition de notre travail est plus limitée quant aux modèles traités car l'explication comparative ne s'applique qu'à des états d'équilibre. Aussi notre intérêt s'est-il porté sur des modèles permettant de décrire ces états: les systèmes d'équations algébriques.

Un système d'équations algébrique (S) est caractérisé par:

- un ensemble de n variables endogènes x_1, x_2, \dots, x_n .
- un ensemble de m variables exogènes u_1, u_2, \dots, u_m .
- un ensemble d'équations algébriques (ici sous forme implicite):

$$(S) \begin{cases} f_1(x_1, \dots, x_n, u_1, \dots, u_m) = 0 \\ f_2(x_1, \dots, x_n, u_1, \dots, u_m) = 0 \\ \dots \\ f_k(x_1, \dots, x_n, u_1, \dots, u_m) = 0 \end{cases}$$

Le cadre de la modélisation étant fixé, il est nécessaire d'introduire des contraintes supplémentaires pour que l'explication comparative soit applicable. H. Simon a en effet montré dans [SIMON 57a] que l'ordre causal est conditionné par deux autres hypothèses: le système d'équations doit être "**contraint**" (self-contained) et **structurel**.

1.1 - Système d'équations contraint

L'hypothèse de système contraint est une condition nécessaire et suffisante pour pouvoir déterminer l'ordre causal d'un système d'équations (nous reviendrons sur ce point en section 2). Basée sur la syntaxe des équations, elle est formulée par les deux points suivants:

- le système comporte autant d'équations que de variables endogènes.
- dans tout sous-système de k équations, au moins k variables endogènes différentes apparaissent

Il faut remarquer que ces deux conditions n'entraînent ni l'existence, ni l'unicité de la solution du système d'équations. Il existe ainsi des systèmes admettant un ordre causal mais n'ayant aucune (ou plusieurs) solution(s); le système d'équations à variables réelles (S_1) ci-dessous en est un exemple. Inversement, un système peut comporter une solution unique sans admettre d'ordre causal. C'est le cas du système composé de la seule équation $x^2 + y^2 = 0$ où x et y sont endogènes. L'ordre causal est donc lié à la possibilité de mettre en place un algorithme de résolution du système d'équations, et non à sa résolution effective.

$$(S_1) \begin{cases} x \text{ exogène} \\ x^2 + y^2 = -1 \end{cases} \quad x \rightarrow y$$

Les deux conditions de l'hypothèse de système contraint reposent sur la distinction entre variables endogènes et exogènes. Il n'existe cependant aucun moyen formel pour effectuer cette distinction. Elle incombe donc au concepteur du modèle qui doit préciser outre les équations, les variables exogènes et les variables endogènes. Plusieurs facteurs peuvent l'aider au cours de cette tâche. Il faut en effet se souvenir que les mécanismes fixant la valeur d'une variable exogène sont extérieurs au système étudié. Ceci est par exemple vérifié lorsque:

- le système étudié est inclus physiquement ou conceptuellement dans un système plus vaste qui fixe la valeur de la variable. Ainsi, pour une entreprise, le taux d'inflation est exogène puisqu'il est fixé à partir du système économique tout entier.
- la variable est contrôlée expérimentalement ou fait l'objet d'une décision indépendamment du reste du système. L'analyse des conséquences d'une décision dans un modèle normatif utilise ce principe: les variables de décision (contrôlées) sont exogènes et les variables objectifs (sur lesquelles on veut agir indirectement) endogènes.
- des éléments agissent sur la variable, mais d'intensité suffisamment faible pour être ignorés. La variable est alors exogène par approximation: pour étudier un problème, une entreprise peut décider de considérer le prix d'achat de ses matières premières comme exogène, même si par ailleurs elle a conscience de son pouvoir de négociation avec les fournisseurs.

Comme le montrent ces exemples, la distinction entre variables endogènes et exogènes est délicate. Elle est empirique et non définitive. Une variable exogène à une étape de modélisation, peut devenir ultérieurement endogène. C'est directement le cas lorsque les mécanismes qui gouvernent la

variable sont ajoutés au modèle, ou indirectement lors de l'ajout d'un paramètre ayant une influence sur cette variable.

1.2 - Système d'équations structurel

La structuralité du modèle est la seconde hypothèse de l'explication comparative dans un système d'équations. Elle concerne, non pas l'existence de l'ordre causal comme la précédente, mais la possibilité de lui attribuer une signification. Illustrons ce problème par un exemple.

Les systèmes (S₂) et (S₃) ci-dessous ont la même solution. Pourtant, l'ordre causal des variables b et c est inversé entre les deux systèmes. Ceci n'est pas surprenant puisque la causalité dépend de la présence ou de l'absence d'une variable dans une équation. Il importe donc de fixer les règles d'écriture des modèles afin que la notion de causalité et donc celle d'explication comparative aient un sens.

$$(S_2) \begin{cases} a \text{ exogène} \\ b = 2a + 1 \\ c = a + b \end{cases} \quad (S_3) \begin{cases} a \text{ exogène} \\ c = 3a + 1 \\ b = c - a \end{cases}$$

$a \rightarrow b \rightarrow c$ $a \rightarrow c \rightarrow b$

L'hypothèse de structuralité a précisément pour but de fixer la forme admissible d'un système d'équations modélisant un domaine d'application. Un système est structurel si chacune de ses équations est structurelle. Une équation est structurelle si elle décrit un mécanisme et un seul du domaine d'application.

Tout comme l'hypothèse de système contraint, la structuralité repose sur une notion empirique: celle de mécanisme. Sa vérification incombe donc également au concepteur du modèle. Le terme "mécanisme", tel que l'entendent H. Simon et Y. Iwasaki dans [SIMON 57a] et [IWASAKI 86a], est lié à celui de loi comportementale. Ce point de vue nous semble cependant trop restreint, aussi faut-il préciser notre position.

La littérature sur la causalité dans les modèles numériques (notamment en Economie) distingue deux types d'équations. Les premières appelées équations de définition décrivent des équivalences entre

variables ou des définitions. L'équation (5.1) qui définit les ventes d'un bien par le produit du nombre d'unités vendues par le prix de vente unitaire en est un exemple:

$$\text{ventes} = \text{nb_unités_vendues} \times \text{prix_de_vente} \quad (5.1)$$

Les secondes appelées équations de comportement modélisent les mécanismes gouvernant les variables du domaine d'application. Ainsi, l'équation de comportement (5.2) approxime le nombre d'unités vendues par une fonction f de son prix de vente :

$$\text{nb_unités_vendues} = f(\text{prix_de_vente}) \quad (5.2)$$

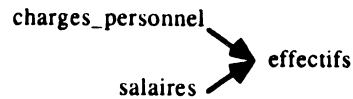
H. Simon et Y. Iwasaki restreignent la notion de causalité aux seules équations de comportement. Ils écrivent ainsi dans [IWASAKI 86b], p. 65:

"Le seul cas où la variation d'une variable se propage instantanément à une autre variable est lorsque la relation entre deux variables n'est pas causale mais définitionnelle, lorsqu'une variable est définie en terme de l'autre. Cependant, il est évident que de telles relations n'ont rien à voir avec la causalité."

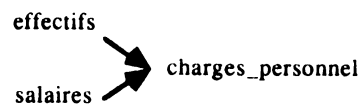
Nous ne partageons pas l'opinion de H. Simon et Y. Iwasaki à propos de la dernière affirmation. L'ordre causal est basé sur la distinction entre variables endogènes et exogènes et sur l'ordre de résolution des variables. Il n'existe donc pas de différence de ce point de vue, entre équations comportementales et définitionnelles. La restriction de la causalité aux équations comportementales a peut-être été introduite pour éviter la critique suivante: dans une équation de définition, la causalité est implicite; elle va des variables utilisées dans la définition vers la variable définie. Cette objection n'est cependant pas fondée. Considérons en effet l'équation de définition (5.3):

$$\text{charges_personnel} = \text{effectifs} \times \text{salaires} \quad (5.3)$$

Imaginons une entreprise cherchant à augmenter ses effectifs (variable endogène) sans pour autant alourdir ses charges de personnel. Le seul moyen à sa disposition est une diminution des salaires (variable exogène). Une explication ultérieure de la situation pourrait être: "les effectifs ont augmenté grâce à la diminution des salaires, les charges de personnel étant restées fixes". Cette explication fait apparaître les relations causales:



Envisageons une autre situation où l'entreprise chercherait à diminuer ses charges de personnel en contrôlant les effectifs et les salaires (variables exogènes). Avec les mêmes valeurs des variables, l'explication appropriée serait alors: "les charges de personnel sont restées stables à cause de l'augmentation des effectifs compensée par la diminution des salaires". Cette explication fait apparaître les relations causales:



Cet exemple montre bien qu'une équation définitionnelle n'impose pas le sens de la causalité entre les variables. Il n'y a donc pas lieu de restreindre la causalité aux équations comportementales. L'ordre causal est étendu aux équations définitionnelles pour l'explication comparative. A cette fin, nous remplaçons l'hypothèse de structuralité formulée par H. Simon par: une équation est structurelle si elle représente un mécanisme ou une définition du domaine d'application.

$$\begin{array}{l}
 \left. \begin{array}{l}
 \frac{\text{valeur_ajoutée}}{\text{production_exercice}} = \text{taux_efficacité} \quad (5.4) \\
 \text{production_globale} - \text{consommations_intermédiaires} = \text{valeur_ajoutée} \quad (5.5) \\
 \text{production_exercice} + \text{marge_commerciale} = \text{production_globale} \quad (5.6) \\
 \text{appro_consommés} + \text{autres_achats} = \text{consommations_intermédiaires} \quad (5.7)
 \end{array} \right\} (S_4)
 \end{array}$$

Pour résumer les deux hypothèses de l'explication comparative dans un système d'équations, considérons le système d'équations (S₄), déjà présenté au chapitre 3. Il est contraint et structurel. Il est contraint car, en supposant les variables appro_consommés, autres_achats, production_exercice et marge_commerciale exogènes, il comporte quatre équations pour quatre variables endogènes. De plus, chaque équation comporte une variable endogène différente, donc tout ensemble de k (1 ≤ k ≤ 4) équations comporte au moins k variables endogènes différentes. Il est structurel car chaque équation représente une définition comptable. L'explication comparative lui est donc applicable. C'est ce que nous allons nous employer à faire dans les sections suivantes.

Auparavant, il faut définir ce qu'est l'écart d'une variable dans un système d'équations. Il est défini par l'accroissement de la variable x entre l'état de référence et l'état dérivé: $\Delta x = x^d - x^r$, si la variable x est numérique. Dans le cas contraire, il est défini de manière identique aux réseaux d'influences par le signe de l'accroissement.

2 - Construction du graphe causal

La construction du graphe causal est plus complexe dans un système d'équations algébriques que dans un réseau d'influences. Une étape supplémentaire est nécessaire pour transformer les équations en relations asymétriques, c'est-à-dire introduire la notion de causalité. Ceci est réalisé en associant à chaque équation une **variable de sortie** différente, ce qui revient à déterminer pour chaque équation la variable qu'elle permet de calculer. Cette opération a un double objectif. Premièrement, elle constitue une autre formulation de l'hypothèse du système contraint. Elle permet donc de vérifier la possibilité de déterminer l'ordre causal du système. Deuxièmement, en faisant jouer aux variables d'une équation des rôles différents, elle introduit la relation asymétrique qui sert à construire le graphe causal du système.

2.1 - Recherche des variables de sortie des équations

Nous avons vu au chapitre 3, qu'un système d'équations peut être représenté par un graphe biparti non orienté $G = (X \cup R, U)$, où X est l'ensemble des variables du modèle, R l'ensemble des relations (des équations, ici) et U l'ensemble des couples (x,r) tels que la variable x apparaît dans la relation r . Associer à chaque équation une variable de sortie revient, dans la terminologie des graphes, à déterminer un **couplage parfait** du sous-graphe de G engendré par l'ensemble des variables endogènes et des relations du modèle.

2.1.1 - Couplage parfait d'un graphe

Un **couplage** d'un graphe $G = (X,U)$ est un sous-ensemble d'arêtes K de U tel que deux arêtes quelconques de K ne soient pas adjacentes (i.e. n'aient aucun sommet commun). Un sommet $s \in X$ est dit **saturé** par le couplage K s'il existe une arête de K incidente à s . Un couplage est **parfait** s'il sature tous les sommets du graphe. Par un abus de langage, nous écrirons "couplage du système S " au lieu de "couplage parfait du sous-graphe engendré par l'ensemble des relations et variables endogènes du graphe

des relations de S'' . La figure 5.1 présente ainsi le couplage de (S_4) dont la figure 3.1 du chapitre 3 illustre le graphe des relations.

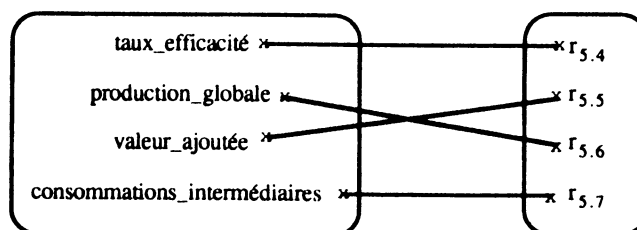


Figure 5.1: le couplage du système d'équations (S_4)

2.1.2 - Algorithme de couplage d'un système d'équations

Les algorithmes de calcul du couplage maximal dans un graphe (i.e. laissant un nombre minimum de sommets insaturés) font appel à des notions assez complexes de la théorie des graphes. Celui d'Hopcroft, qui est brièvement présenté, n'échappe pas à la règle (pour plus de détails, le lecteur pourra se reporter à [HOPCROFT 73]).

L'algorithme d'Hopcroft est basé sur la notion de **chaîne augmentante**. Soit $G = (X, U)$ un graphe et K un couplage de ce graphe. Une chaîne C est augmentante relativement à K si et seulement si:

- les arêtes qui composent C sont alternativement dans $U - K$ et dans K
- les extrémités de C sont insaturées.

Si K est un couplage de G d'ordre n (comprenant n arêtes) et C une chaîne augmentante relativement à K , alors $K \Delta C$ est un couplage d'ordre $n+1$ (Δ est l'opérateur de différence symétrique: $K \Delta C = (K - C) \cup (C - K)$). De plus, la suite:

$$\left\{ \begin{array}{l} K_0 = \emptyset \\ K_{i+1} = K_i \Delta C_i \\ \text{où } C_i \text{ est une chaîne augmentante relativement à } K_i, \text{ de longueur minimale} \end{array} \right.$$

converge vers le couplage maximum du graphe, ce qui fournit l'idée de l'algorithme.

Les algorithmes utilisant cette technique pour des graphes quelconques (non bipartis) ont une complexité de l'ordre de n^3 . Hopcroft montre cependant que dans un graphe biparti, (comme le graphe

des relations d'un système d'équations), l'algorithme est plus efficace en calculant l'ensemble des C_i en une seule étape. La complexité obtenue est alors de l'ordre de $n^{5/2}$. L'algorithme peut être écrit de la manière suivante:

procédure couplage-max

début

$K := \emptyset$;

répéter

début

$l(K)$:= longueur de la plus petite chaîne augmentante relativement à K ;

S := ensemble maximal des chaînes $\{C_1, C_2, \dots, C_n\}$ tel que:

a- pour tout i , C_i est une chaîne augmentante relativement à K et $|C_i| = l(K)$

b- les ensembles de sommets des C_i sont disjoints;

si $S \neq \emptyset$ alors $K := K \Delta C_1 \Delta C_2 \dots \Delta C_n$

fin

jusqu'à $S = \emptyset$

fin;

2.2 - Construction du graphe réduit

L'existence d'un (il peut en exister plusieurs) couplage d'un système est une condition nécessaire et suffisante de l'existence de l'ordre causal. Il est en effet démontré dans [IVANESCU 66] qu'en l'absence de couplage, le système contient un sous-ensemble d'équations ayant moins de variables que d'équations; il est donc non contraint. Inversement, s'il existe un couplage, l'ordre causal existe également. Le graphe causal du système peut en effet être déterminé à partir du couplage en deux étapes. Un graphe orienté associé au couplage, assez semblable à un réseau d'influences, est d'abord construit. Le graphe réduit de ce graphe fournit alors l'ordre causal. Ces deux opérations étant toujours possibles, l'existence d'un couplage entraîne donc celle de l'ordre causal.

Le graphe associé au couplage est construit de la manière suivante. Les variables (endogènes et exogènes) sont les sommets du graphe. Un arc est tiré de la variable x à la variable y s'il existe une équation où apparaissent les variables x et y et où y est la variable de sortie. Ainsi, le graphe associé au couplage du système (S_4) est présenté sur la figure 5.2.

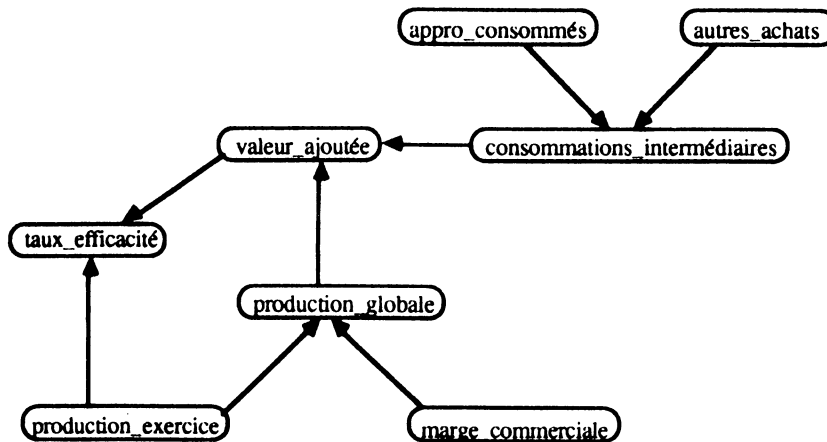


Figure 5.2: le graphe obtenu à partir du couplage du système (S₄)

L'ordre causal est obtenu en calculant le graphe réduit du graphe précédent. Celui du système (S₄), est identique au graphe de la figure 5.2 car chaque CFC comporte une seule variable.

Il faut noter que le couplage d'un système d'équations, lorsqu'il existe, n'est généralement pas unique. Ainsi, le système (S₅) ci-dessous comprenant quatre variables endogènes a, b, c, d, admet deux couplages différents présentés sur la figure 5.3.

$$(S_5) \begin{cases} f_1(a,b) = 0 & (5.8) \\ f_2(a,b) = 0 & (5.9) \\ f_3(b,c) = 0 & (5.10) \\ f_4(a,b,d) = 0 & (5.11) \end{cases}$$

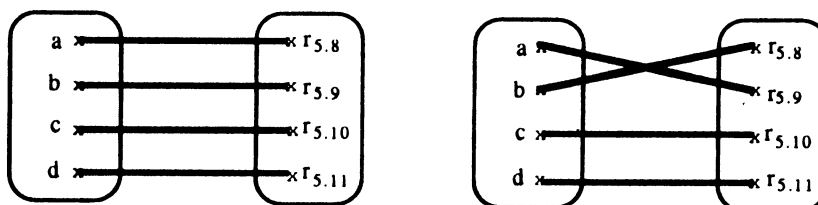


Figure 5.3: les deux couplages possibles du système (S₅)

Il est cependant prouvé dans [STEWART 62] que l'ordre causal ne dépend pas du couplage obtenu. Les différentes variantes portent en effet sur les variables résolues simultanément, et ne modifient donc

pas le graphe causal. Ainsi, les deux couplages de la figure 5.2 conduisent au même graphe causal présenté sur la figure 5.4.

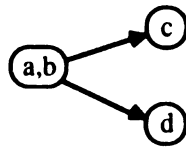


Figure 5.4: le graphe causal du système (S₅)

3 - Construction de la base d'explication

La construction de la base d'explication complète minimale d'une variable est identique à celle dans les réseaux d'influences lorsque le système ne comporte pas de boucle de rétroaction. Ainsi, la BECM de la variable `taux_effacité` du système d'équations (S₄) est présentée sur la figure 5.5.

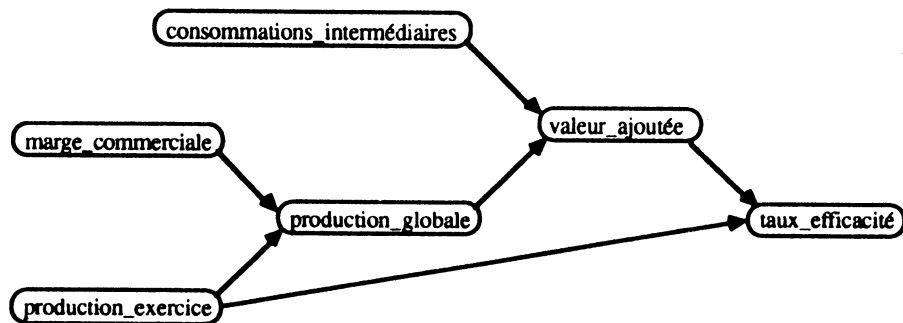


Figure 5.5: la BECM de la variable `taux_effacité`

La situation est néanmoins différente quand il existe des boucles de rétroaction. Nous avons vu au chapitre 4 que deux stratégies sont possibles pour l'explication des variables liées par une boucle de rétroaction: soit l'explication mutuelle de ces variables, soit les considérer comme indissociables. Un compromis entre ces deux alternatives a été adopté dans les réseaux d'influences. Cette décision n'est pas valide dans les systèmes d'équations. La raison en est que le couplage d'un système n'est pas unique lorsque le système comporte des boucles de rétroaction. Or, les variantes du couplage portent sur les variables à l'intérieur d'une même CFC. L'utilisation de la solution adoptée pour les réseaux d'influences conduirait donc ici à des explications différentes selon le couplage obtenu. Par conséquent, la seconde

alternative doit être choisie: la BECM d'une variable est construite à partir du graphe causal, et les variables d'une même CFC ne peuvent s'expliquer mutuellement.

Ce traitement est acceptable lorsque les CFC sont peu nombreuses et de petite taille. Dans le cas inverse, il soulève un problème délicat: l'analyse de l'écart d'une variable fait appel à un nombre trop important de variables et perd son pouvoir explicatif. Dans le cas extrême, le système comporte une seule CFC, et aucune explication ne devient possible. On retrouve ici un problème classique en Economie où les systèmes peuvent atteindre plusieurs milliers d'équations. Afin d'en maîtriser la complexité, il est nécessaire de les décomposer en sous-systèmes, comme dans l'explication comparative. Cependant, les boucles de rétroaction rendent généralement cette décomposition insuffisante. Il faut donc utiliser d'autres techniques: c'est l'objet de l'analyse structurale de modèles (voir [MALLOUM 80], par exemple).

Les économistes résolvent ce problème en déterminant un ensemble de variables dites de bouclage, telles qu'une fois supprimées du modèle il n'existe plus de boucle de rétroaction. Comme le nombre de variables de bouclage est généralement très inférieur au nombre de variables endogènes, la structure "essentielle" du modèle est ainsi déterminée. Les possibilités d'analyse du modèle sont alors restaurées.

Différents algorithmes ont été développés pour effectuer cette tâche (cf. [NEPOMIASTCHY 78] par exemple). Ne s'étant pas révélés nécessaires dans les applications que nous avons abordées, leur implantation dans CRESUS n'a pas été réalisée.

4 - Présentation de la base d'explication

Cette étape revêt des aspects totalement différents selon que les variables du modèle sont numériques ou qualitatives. Dans le premier cas, les valeurs numériques permettent de quantifier l'effet d'une variable sur une autre et donc de fournir des explications plus précises que dans les réseaux d'influences. Dans le second, il n'y a pas d'autre recours que de transformer le système d'équations en un réseau d'influences.

4.1 - Cas des variables numériques

Lorsque les valeurs numériques de l'état de référence et dérivé sont connues, il est possible de quantifier l'effet d'une variable sur une autre. Pour cela, une grandeur appelée **responsabilité** a été introduite. La responsabilité de la variable x sur la variable y est le réel noté: $r(x/y)$, défini par :

$$\left\{ \begin{array}{l} r(x/y) = \frac{y^{\Delta x} - y^r}{y^d - y^r} \text{ si } y^d \neq y^r \\ r(x/y) = 0 \text{ si } y^d = y^r \end{array} \right.$$

où y^r et y^d sont les valeurs de y dans l'état de référence et dérivé. $y^{\Delta x}$ est la valeur qu'aurait y si toutes les variables précédant x dans le graphe causal gardaient leur valeur de l'état de référence et x prenait sa valeur de l'état dérivé.

Autrement dit, la responsabilité de x sur y est la variation relative de y due à la seule variation de x . Il faut noter l'analogie entre responsabilité et dérivée partielle. Nous verrons cependant que cette analogie est loin d'être totale.

Le signe de la responsabilité est suffisant pour déterminer l'effet d'une variable, car les relations "confirme" et "infirme" définies au chapitre 3 s'écrivent:

- x confirme y si et seulement si $r(x/y) > 0$
- x infirme y si et seulement si $r(x/y) < 0$

Non seulement le signe de la responsabilité détermine l'effet d'une variable, mais sa valeur en fournit une mesure. Cette propriété est utilisée pour comparer les effets de plusieurs variables, afin de générer des explications plus précises. Considérons à nouveau le système (S_4) avec les valeurs du tableau 1. On a:

$$r(\text{marge_commerciale} / \text{production_globale}) = \frac{(103,03+1100,78) - (101,64+1100,78)}{(1300,78 - 1202,42)} = \frac{1,39}{98,36} \approx 0,01$$

$$r(\text{production_exercice} / \text{production_globale}) = \frac{(101,64+1197,75) - (101,64+1100,78)}{(1300,78 - 1202,42)} = \frac{96,97}{98,36} \approx 0,99$$

La marge commerciale ayant une responsabilité minimale sur l'augmentation de la production globale par rapport à la variation de produit de l'exercice, elle n'est pas mentionnée dans l'explication produite.

Par ailleurs, la responsabilité est utilisée pour ordonner les variables dans le texte de l'explication: les causes principales sont d'abord présentées. Elle sert également à moduler les expressions utilisées ("l'augmentation de x et à un degré moindre la diminution de y ont entraîné...").

variables	1987	1988
taux_efficacité	0,105	0,114
valeur_ajoutée	115,69	136,64
production_globale	1202,42	1300,78
production_exercice	1100,78	1197,75
marge_commerciale	101,64	103,03
consommations_intermédiaires	1086,73	1164,14
appro_consommés	1014,21	1090,51
autres_achats	72,52	73,63

Tableau 1

Le système d'équations (S₄) a été introduit dans CRESUS avec les variables exogènes mentionnées plus haut et les valeurs du tableau 1. L'explication suivante pour la comparaison temporelle du taux d'efficacité est obtenue:

L'augmentation de la production de l'exercice de 8,8 % entre 1987 et 1988 a entraîné l'augmentation de la production globale de 8,2 %. L'augmentation de la production globale a provoqué l'augmentation de la valeur ajoutée de 18,1 % malgré l'augmentation des consommations intermédiaires de 7,1 %. L'augmentation de la valeur ajoutée a entraîné l'augmentation du taux d'efficacité de 8,6 %.

Malheureusement, la responsabilité présente des imperfections. Analysons par exemple le problème d'interprétation comparative de la figure 5.6.

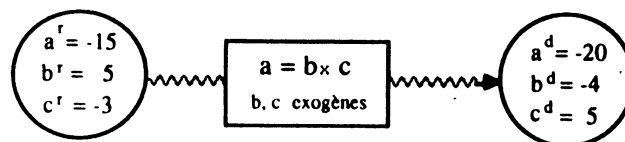


Figure 5.6: les imperfections de la responsabilité

Le calcul des responsabilités des variables b et c sur la variable a donne les résultats suivants:

$$r(b/a) = \frac{(-4 \times -3) - (5 \times -3)}{-20 - -15} = -\frac{27}{5}$$

$$r(c/a) = \frac{(5 \times 5) - (-3 \times 5)}{-20 - -15} = -\frac{40}{5}$$

Les variables b et c infirment a; l'écart de a ne semble donc avoir aucune cause. Pour comprendre les raisons de ce problème, il faut revenir sur l'analogie entre responsabilité et dérivée partielle. Considérons une équation mise sous forme explicite $y = f(x_1, x_2, \dots, x_n)$. Supposons que les variables x_1, x_2, \dots, x_n soient indépendantes causalement. On a:

$$\begin{aligned} \frac{\partial y}{\partial x_i} &= \lim_{x_i^d \rightarrow x_i^r} \frac{f(x_1^r, \dots, x_i^d, \dots, x_n^r) - f(x_1^r, \dots, x_i^r, \dots, x_n^r)}{x_i^d - x_i^r} \\ &= \lim_{x_i^d \rightarrow x_i^r} \frac{f(x_1^r, \dots, x_i^d, \dots, x_n^r) - f(x_1^r, \dots, x_i^r, \dots, x_n^r)}{x_i^d - x_i^r} \\ &= \lim_{x_i^d \rightarrow x_i^r} \frac{y^d - y^r}{x_i^d - x_i^r} \\ &= \lim_{x_i^d \rightarrow x_i^r} \frac{y^d - y^r}{x_i^d - x_i^r} r(x_i/y) \end{aligned}$$

En supposant x_i^d proche de x_i^r , on a: $r(x_i/y) \approx \frac{\Delta x_i}{\Delta y} \frac{\partial y}{\partial x_i}$

où Δx_i et Δy sont les accroissements respectifs de x_i et y entre l'état de référence et l'état dérivé

Or, la différentielle totale dy de la variable y s'exprime par la relation: $dy = \sum_{i=1}^n \frac{\partial y}{\partial x_i} dx_i$.

Donc si toutes les variables x_i sont telles que x_i^d est suffisamment proche de x_i^r , on a:

$$\Delta y \approx \sum_{i=1}^n \frac{\partial y}{\partial x_i} \Delta x_i, \text{ et par conséquent: } \sum_{i=1}^n r(x_i/y) \approx 1.$$

C'est sur cette dernière approximation que repose la possibilité de comparer les responsabilités des variables x_i . Cependant, lorsque les écarts entre valeurs dans l'état de référence et dérivé augmentent, cette approximation devient invalide. En particulier, lorsqu'une dérivée partielle par rapport à une variable explicative change de signe entre les deux états, comme sur l'exemple de la figure 5.6, les responsabilités perdent toute signification.

Aucune solution totalement satisfaisante n'a pu être apportée à ce problème. Pour que les responsabilités calculées soient significatives, il faut que la dérivée partielle par rapport à chaque variable explicative garde un signe constant entre l'état de référence et l'état dérivé. La vérification de cette condition est cependant très coûteuse. Pour la rendre opérationnelle, il faut en effet simuler le système d'équations en chaque point où l'on désire obtenir les dérivées partielles.

La solution adoptée consiste à calculer et comparer les signes des dérivées partielles uniquement dans l'état de référence et dérivé. Si les signes de chaque dérivée partielle dans l'état de référence et dérivé sont identiques, les responsabilités calculées sont jugées significatives. Dans le cas inverse, l'effet devient indéterminé comme dans les réseaux d'influences.

Le système (S₄) avec les valeurs du tableau 1 est un exemple où les responsabilités sont valides. Toutes les dérivées partielles par rapport aux variables explicatives ont en effet un signe constant. Par contre, dans l'exemple 5.6, $\frac{\partial c}{\partial a}|_{\text{référence}} = 5$ et $\frac{\partial c}{\partial a}|_{\text{dérivé}} = -4$; les effets de a et b sur c sont donc indéterminés. L'explication utilise dans ce cas une formule neutre ne précisant pas l'effet des variables explicatives: "le passage de b de 5 à -4 et de c de -3 à 5 ont causé le passage de a de -15 à -20".

Il est bien sûr aisé de construire des exemples mettant cette solution en défaut. Empiriquement, plusieurs facteurs concourent à rendre les explications correctes dans les applications que nous avons traitées: variables pour la plupart définies positives, équations simples utilisant essentiellement des opérateurs arithmétiques. L'applicabilité de cette technique dans des domaines où ces conditions ne sont pas vérifiées reste cependant à démontrer.

4.2 - Cas des variables qualitatives¹

Un système d'équations contient plus d'informations qu'un réseau d'influences puisqu'il précise la forme des relations entre les variables. Néanmoins, ce surcroît d'information n'est pas exploitable lorsque les valeurs numériques des états de référence et dérivé sont inconnues. Les écarts des variables étant qualitatifs, les techniques présentées dans le paragraphe précédent sont en effet impossibles à mettre en oeuvre. Pour cette raison, le système d'équations est transformé en un réseau d'influences de sorte à se ramener au cas du chapitre 3. Cette opération est cependant extrêmement délicate, car elle nécessite des

¹La version actuelle de CRESUS ne traite pas ce cas

délicate, car elle nécessite des manipulations symboliques sur les équations. Quand elle est possible, elle est effectuée en deux étapes. Les équations sont d'abord traduites en confluences ([DE KLEER 84], cf. chapitre 2), qui sont ensuite transformées en relations d'influences.

4.2.1 - Des équations aux confluences

Rappelons qu'une confluence est une expression qualitative de la forme: $\sum T_i = C$, où T_i est soit une variable soit le produit d'une constante par une variable, et C une constante. Pour obtenir une confluence à partir d'une équation sous forme explicite $y = f(x_1, x_2, \dots, x_n)$, la variable de sortie y de l'équation est d'abord différenciée par la relation:

$$dy = \sum_{i=1}^n \frac{\partial y}{\partial x_i} dx_i$$

Puis, la relation est traduite au niveau qualitatif, en utilisant l'algèbre qualitative ([DE KLEER 84]).

Exemple 1

Considérons l'équation (5.12), où a est la variable de sortie.

$$a + b + c = 0 \quad (5.12)$$

Mise sous forme explicite, cette équation devient:

$$a = -b - c \quad (5.13)$$

En différenciant, on obtient:

$$da = -db - dc \quad (5.14)$$

Le passage au niveau qualitatif produit donc:

$$[da] = -[db] - [dc] \quad (5.15)$$

Exemple 2

Considérons l'équation (5.16), où a est la variable de sortie.

$$a - bc = 0 \quad (5.16)$$

Sous forme explicite, elle s'écrit:

$$a = b c \quad (5.17)$$

On a donc: $da = c.db + b.dc \quad (5.18)$

D'où: $[da] = [c] [db] + [b] [dc] \quad (5.19)$

où $[x]$ est la valeur qualitative de x dans l'ensemble $\{+, 0, -\}$

L'automatisation de cette tâche est difficile car elle nécessite des manipulations symboliques sur les équations pour:

- obtenir la forme explicite de la variable de sortie
- calculer formellement les dérivées partielles de y par rapport à chacune des variables x_i
- simplifier les expressions qualitatives obtenues

4.2.2 - Des confluences aux relations d'influences

La traduction des confluences en relations d'influences est très simple: à une expression $[dy] = \sum_{i=1}^n T_i [dx_i]$, sont associées n relations d'influences conditionnelles $I+(x_i y)$ ou $I-(x_i y)$. Les conditions sont obtenues à partir des expressions T_i . La relation (5.15) de l'exemple 1 est transformée en:

$$\left\{ \begin{array}{l} I- (b a) \\ I- (c a) \end{array} \right.$$

Dans l'exemple 2, à partir de la relation (5.19), on obtient:

$$\left\{ \begin{array}{l} \text{si } [c] = + \text{ alors } I+ (b a) \\ \text{si } [c] = - \text{ alors } I- (b a) \\ \text{si } [b] = + \text{ alors } I+ (c a) \\ \text{si } [b] = - \text{ alors } I- (c a) \end{array} \right.$$

Lorsque le système d'équations est transformé en un réseau d'influences, les procédures décrites au chapitre 3 peuvent lui être appliquées.

5 - Comparaison avec d'autres travaux

L'explication comparative dans un système d'équations avec des variables numériques s'inscrit dans la même démarche que les travaux effectués dans ROME ([KOSY 86]), le générateur de systèmes "intelligents" d'aide à la décision développé à Carnegie Mellon. Une composante de ROME, dénommée ERGO ([KING 86]), explique des résultats numériques produits par un modèle financier. Ces explications n'ont cependant pas le même objectif qu'en interprétation comparative. Il s'agit de fournir une aide interactive à la modélisation et à l'interprétation de résultats.

Les explications sont construites dans ERGO à partir de la notion d'"effet", très voisine de celle de responsabilité. Considérant une variable y définie par l'équation $y = f(a,b,c,...)$ dont on veut expliquer l'écart entre deux instants t_1 et t_2 , l'effet d'un ensemble X de variables explicatives sur la variable expliquée y , noté $\epsilon(X,y)$ est défini par: $\epsilon(X,y) = y_2 - f(Z)$, où $f(Z)$ est la valeur qu'aurait y si toutes les variables exceptées celles dans X avaient changé (en restreignant X à un élément, c'est donc la définition **duale** de la responsabilité). Cependant, les hypothèses adoptées limitent considérablement la généralité de ERGO.

Tout d'abord, l'ordre causal n'est pas déduit du modèle, mais imposé par la forme des équations; $y = f(a,b,c,...)$ est interprétée: $a, b, c...$ causent y . Ensuite, ERGO ne traite que des modèles récursifs (i.e. sans boucle de rétroaction). Enfin et surtout, pour expliquer une variable, ERGO ne prend en compte que les variables explicatives apparaissant dans l'équation qui la définit, et les traite comme des variables indépendantes. S'il existe un chemin de causalité entre les variables explicatives, cette méthode peut conduire à des explications incorrectes. Ainsi, sur l'exemple du système (S₄), ERGO ne considérant que la seule équation (5.4) pour expliquer la variable `taux_efficacité`, pourrait être amené à conclure: "le taux d'activité a augmenté à cause de la diminution de la production de l'exercice" !.

CONCLUSION DE LA PREMIERE PARTIE

Nous avons vu au chapitre 2 que la plupart des raisonnements applicables à un modèle qualitatif dérivent, soit de la propagation (dans les graphes de causalité), soit de la simulation qualitative. Il est donc important de souligner au terme de cette première partie que l'interprétation comparative enrichit la panoplie des raisonnements qualitatifs. Il est par ailleurs nécessaire de rappeler les limites et hypothèses du travail présenté afin de suggérer différents prolongements du travail entrepris. Les principales limites sont au nombre de trois: inadéquation de l'explication comparative lorsque les variables du domaine d'application sont fortement interconnectées, absence de moyen formel pour vérifier la validité de la responsabilité d'une variable sur une autre dans un système d'équations, et difficulté pour implanter l'interprétation comparative sur des variables qualitatives dans un système d'équations. Les principales hypothèses sont également au nombre de trois.

La première est la restriction à deux états comparés. Elle a été introduite afin de fixer un cadre précis à la comparaison; c'est elle qui rend possible la distinction entre état de référence et état dérivé sur laquelle est fondée la méthode présentée au chapitre 3. Deux niveaux de relaxation de cette hypothèse peuvent être distingués. Le premier est la comparaison de n ($n \geq 2$) états, où toutes les comparaisons deux à deux ne sont pas indispensables; c'est par exemple le cas pour l'interprétation de séries chronologiques. Les techniques proposées dans [CALOUD 88] pour l'interprétation qualitative de résultats numériques fournis par un procédé continu constituent un élément de réponse à ce problème: les n valeurs successives de chaque variable doivent être segmentées puis l'explication comparative déclenchée sur les différents segments. La comparaison de n états indépendants, nécessaire pour analyser les conséquences résultant de n décisions, par exemple, est encore plus complexe. Elle nécessite $n(n-1)/2$ interprétations comparatives à deux états, ce qui la rend très rapidement impraticable. Aucune solution ne semble applicable à ce problème.

Deuxièmement, le travail présenté est limité à des systèmes d'équations algébriques et donc à la comparaison d'états d'équilibre. La seconde extension suggérée est l'application de l'explication comparative à des systèmes d'équations différentielles. L'explication de résultats de simulations numériques deviendrait ainsi possible (entre deux instants). L'ordre causal est extensible à des systèmes d'équations différentielles ([KEVORKIAN 75], [IWASAKI 88]), mais le problème délicat est l'extension de la notion de responsabilité à de tels systèmes. Cette dernière est déjà le point faible de l'explication comparative dans des systèmes d'équations algébriques, son extension à des systèmes d'équations différentielles semble donc difficile.

Une troisième extension envisageable est l'application de l'explication comparative à des modèles qualitatifs plus riches que les réseaux d'influences. Nous avons vu au chapitre 2 que ces derniers constituent un outil simple pour raisonner qualitativement sur un système, mais parfois insuffisant à cause du peu d'information qu'ils contiennent. Les processus qualitatifs [FORBUS 84] (cf. chapitre 2) semblent la représentation la plus adaptée à cette extension. Les descriptions structurelles de B. Kuipers et J. De Kleer découlent en effet directement d'une traduction d'équations différentielles; il serait donc préférable d'appliquer l'explication comparative directement sur ces équations, plutôt que sur une représentation qualitative qui en dérive.

Il faut enfin remarquer qu'outre les modèles numériques et qualitatifs, l'ordre causal est également applicable à des modèles logiques: [SIMON 57b]. Il peut donc être envisagé de doter un agent logique de l'interprétation comparative afin que celui-ci puisse justifier un changement de décision ou d'état.

DEUXIEME PARTIE

INTRODUCTION

Cette partie présente les trois applications informatiques développées dans cette thèse pour le projet PIC.

L'interprétation comparative est à l'origine du développement de deux de ces applications: ANAFI, pour l'analyse financière et SURPLUS, pour l'analyse de la performance d'une entreprise. Dans le premier problème, il s'agit d'établir un bilan de santé d'une entreprise et d'expliquer les causes de ses éventuelles difficultés. Ceci est réalisé en comparant la situation de l'entreprise à celle de son secteur d'activité d'une part, et à une situation antérieure d'autre part. L'étude de la performance d'une entreprise par la méthode des surplus est abordée de manière identique. Le surplus de productivité global, un indicateur de l'évolution de la performance d'une entreprise entre deux exercices, est dégagé. La comparaison des quantités de facteurs consommés et de biens produits entre les deux exercices permet d'expliquer la formation du surplus, tandis que la comparaison des prix permet d'expliquer sa répartition entre les différentes parties prenantes de l'entreprise. Dans les deux problèmes, des équations comptables servent de modèle du fonctionnement de l'entreprise. Celui-ci est complété dans ANAFI par un réseau d'influences qui modélise le fonctionnement qualitatif d'une entreprise. ANAFI et SURPLUS ont été développés en CRESUS, un générateur de systèmes d'interprétation comparative dont la réalisation repose sur les idées présentées dans la première partie de ce mémoire.

L'imputation comptable est la troisième application abordée. Il s'agit de déduire les écritures associées aux différentes pièces comptables traitées par une entreprise: factures, avoirs, chèques, etc. SICA (Système pour l'Imputation Comptable Automatique), un système expert utilisant des connaissances profondes a été réalisé pour le traitement de cette application. Il analyse des phrases décrivant le contenu de pièces comptables et déduit l'écriture correspondante. Comme il n'existe pas

de modèle des pièces comptables, l'imputation n'est pas effectuée directement. L'opération économique sous-jacente est d'abord reconstituée sous forme d'un échange de flux entre l'entreprise et un tiers. L'imputation se déduit alors aisément de cet échange, qui constitue un véritable modèle du fonctionnement comptable de l'entreprise.

La seconde partie comprend deux chapitres. ANAFI et SURPLUS, les deux systèmes experts utilisant l'interprétation comparative sont présentés dans le chapitre 6 après avoir décrit CRESUS, le système qui a servi à leur réalisation. SICA, le système expert pour l'imputation comptable est ensuite présenté dans le chapitre 7.

CHAPITRE 6

DEUX APPLICATIONS DE L'INTERPRÉTATION COMPARATIVE DANS LE PROJET PIC

Ce chapitre présente les deux applications du projet PIC basées sur l'interprétation comparative: ANAFI, pour l'analyse financière et SURPLUS, pour l'analyse de la performance d'une entreprise par la méthode des surplus.

Ces deux applications semblent à première vue similaires: il s'agit d'analyser les documents comptables d'une entreprise et de produire un rapport, sur sa situation dans le premier cas, sur sa performance dans le second. Elles diffèrent cependant par bien des aspects. Par leur complexité tout d'abord: l'analyse de la performance est plus facile à traiter. Contrairement au diagnostic financier, elle n'exige pas d'évaluation comparative et le modèle utilisé pour l'explication comparative est plus simple: il ne comprend qu'un système d'équations comptables. L'objectif de la réalisation de ces deux systèmes est également différent. Il s'agissait avec ANAFI de valider l'interprétation comparative sur une application concrète. Une fois cette étape franchie, SURPLUS a été construit pour en tester l'efficacité, c'est-à-dire la possibilité de réaliser un système opérationnel. Des méthodes de travail différentes en ont résulté: ANAFI a été réalisé sans le concours d'un expert, ce qui n'est pas le cas de SURPLUS. Ceci se traduit bien entendu au niveau des performances: les résultats produits par SURPLUS sont bien plus proches du niveau d'expertise requis que ceux d'ANAFI.

CRESUS, le générateur de systèmes d'interprétation comparative est tout d'abord présenté en section 1. Les deux applications qu'il a permis de réaliser, ANAFI et SURPLUS font l'objet des deux sections suivantes.

1 - CRESUS: un générateur de systèmes d'interprétation comparative

CRESUS résulte de la même démarche qui permet de passer du système expert au générateur de systèmes experts. Le système ANAFI ayant d'abord validé la notion d'interprétation comparative, un générateur de systèmes a été réalisé pour cette classe d'applications.

Cette opération a un triple intérêt. Elle fournit un cadre méthodologique pour la réalisation d'un système: les aspects propres à chaque problème sont séparés de ceux qui sont généraux à toute la classe d'applications. Elle procure également un gain de rapidité pour le développement d'un système. Au lieu d'être introduites "manuellement" dans le système, les caractéristiques de chaque problème peuvent faire l'objet d'outils spécialisés: éditeur et analyseur de modèles, etc. Enfin, ces outils fournissent un gain de fiabilité: la réécriture du modèle d'ANAFI en CRESUS a ainsi permis de trouver des erreurs qui n'avaient pas été détectées lorsqu'il était codé directement en LISP.

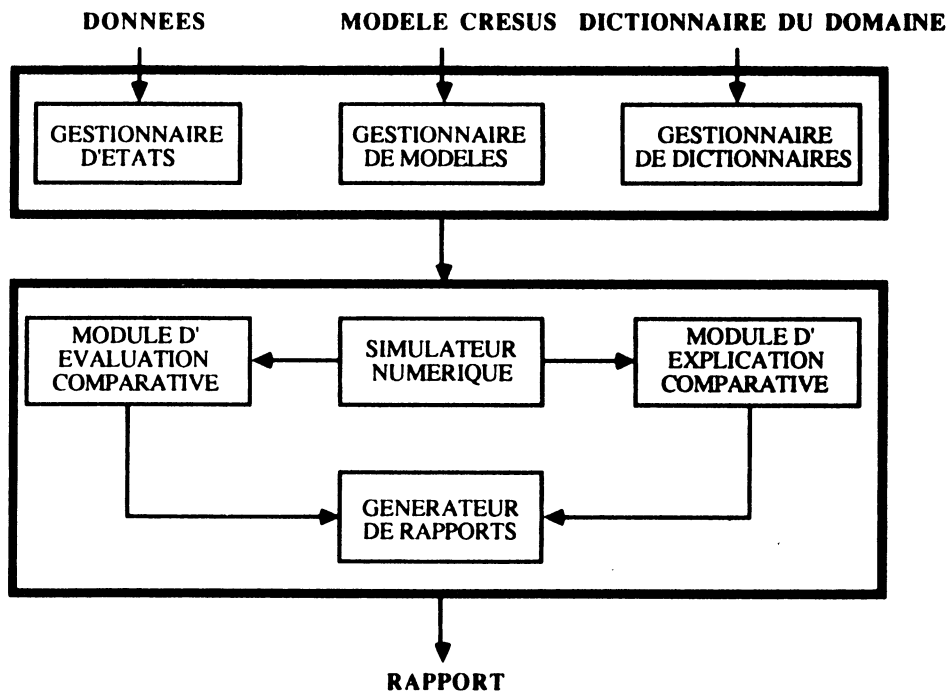


Figure 6.1: architecture de CRESUS

La réalisation d'un système d'interprétation comparative s'effectue en fournissant à CRESUS (figure 6.1) un modèle du système analysé, les données correspondant aux états à comparer et un dictionnaire des variables du domaine. CRESUS produit à partir de ces informations un rapport expliquant le passage du système de l'état de référence à l'état dérivé. Pour effectuer ces traitements, CRESUS est structuré en deux couches. La première joue le rôle d'interface entre le système et l'utilisateur, tandis que la seconde constitue le noyau du système.

1.1 - Les interfaces de CRESUS

Les interfaces de CRESUS permettent à l'utilisateur de spécifier les données du problème d'interprétation comparative à résoudre: modèle du système analysé, états à comparer et vocabulaire de l'application. A chacun de ces trois types de données correspond un gestionnaire.

1.1.1 - Le gestionnaire de modèles

Un modèle CRESUS est constitué de deux parties: un système d'équations algébriques et un réseau d'influences conditionnel (un ensemble d'expressions: [si <CONDITION> alors] $I+(x y)$ ou $I-(x y)$). Le gestionnaire de modèles comporte un éditeur pour la saisie et un analyseur pour vérifier la correction du modèle. Deux types de contrôles sont effectués: analyse syntaxique du modèle et, pour les équations, vérification de l'existence d'un couplage parfait (i.e. la possibilité d'associer une variable de sortie à chaque équation; cf. chapitre 5).

1.1.2 - Le gestionnaire d'états

Le gestionnaire d'états est essentiellement une interface entre CRESUS et un tableur. Ce dernier est utilisé pour la saisie des valeurs des variables dans les différents états comparés. Comme nous le verrons dans les sections 2 et 3, une application peut faire intervenir plusieurs problèmes d'interprétation comparative, donc plus de deux états, mais l'état de référence et dérivé de chacun de ces problèmes doit être spécifié par l'utilisateur. Pour les modèles numériques, seules les valeurs des variables exogènes sont saisies. Cela permet d'effectuer la distinction entre variables endogènes et exogènes nécessaire au calcul de l'ordre causal. Les variables endogènes sont calculées en résolvant le modèle.

1.1.3 - Le gestionnaire de dictionnaires

Bien que le principe de l'explication comparative soit indépendant du domaine d'application, le vocabulaire utilisé dans le rapport généré par CRESUS doit tenir compte des besoins de l'utilisateur. Pour cette raison, celui-ci fournit un dictionnaire des variables de l'application qui permet à CRESUS d'adapter son rapport au problème traité. Ce dictionnaire contient pour chaque variable:

- le nom complet de la variable et les informations nécessaires à l'élaboration de phrases: genre, nombre.
- le nombre de décimales et l'unité appropriés (quand il s'agit d'une variable numérique).
- les expressions associées à cette variable. Ainsi, à propos d'une augmentation, il faut selon les variables employer les termes: croissance, amélioration, hausse, alourdissement, etc.

1.2 - Le noyau de CRESUS

Modèle, états et dictionnaire de variables constituent les données du problème d'interprétation comparative résolu par le noyau de CRESUS. Ce dernier est constitué des modules d'évaluation et d'explication comparatives, d'un simulateur numérique et d'un générateur de rapports.

1.2.1 - Le module d'évaluation comparative

Comme nous l'avons signalé dans la première partie de la thèse, l'évaluation comparative dépend étroitement de l'application traitée. Aussi, ce module initialement vide doit-il être construit pour chaque application (cf. section 2).

1.2.2 - Le module d'explication comparative

Le module d'explication comparative n'est autre que le résultat de l'implantation des techniques présentées dans les chapitres 4 et 5.

1.2.3 - Le simulateur numérique

Le simulateur numérique remplit deux fonctions dans CRESUS. Il calcule la valeur des variables endogènes du modèle numérique. Il calcule également les responsabilités utilisées pour la présentation de la base d'explication d'une variable. Rappelons en effet que dans un modèle numérique, le commentaire de l'effet d'une variable sur une autre dépend du signe de sa responsabilité, c'est-à-dire la valeur qu'aurait la seconde si la première avait seule varié entre l'état de référence et l'état dérivé.

1.2.4 - Le générateur de rapports

Le générateur de rapports construit les phrases de l'explication comparative d'une variable. Le texte expliquant l'écart d'une variable entre l'état de référence et l'état dérivé est obtenu par un parcours de sa base d'explication. Ce parcours est effectué depuis les variables les plus reculées de la base d'explication vers la variable expliquée. La génération d'une phrase associée à une portion de base d'explication peut être illustrée par l'exemple simple de la figure 6.2. Le dictionnaire des variables est d'abord consulté pour obtenir le nom complet des variables; imaginons que ce soit: "variable *a*", "variable *b*" et "variable *c*". Supposons de plus que le terme approprié pour décrire un écart positif (resp. négatif) de ces variables soit "augmentation" (resp. "diminution"). Comme l'écart de la variable *c* est confirmé par celui de la variable *a* et infirmé par celui de *b*, la phrase suivante peut être générée: "l'augmentation de la variable *a* a entraîné la diminution de la variable *c*, malgré l'augmentation de la variable *b*".

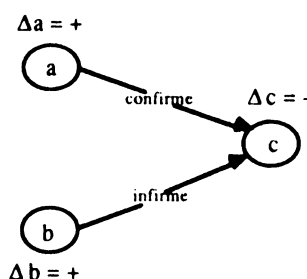


Figure 6.2: génération d'une phrase à partir d'une portion de base d'explication

Les phrases sont ensuite assemblées pour produire le rapport complet de l'interprétation comparative.

Le noyau de CRESUS est écrit en CEYX, un langage orienté objet développé à l'Inria autour de LE_LISP, tandis que les interfaces sont écrites en PROLOG. La totalité du système comporte environ 4000 lignes de code. CRESUS a été conçu pour être utilisé sur un matériel de type PC-compatibles. Néanmoins, à cause des limitations de mémoire dues au système d'exploitation MS-DOS, la version actuelle ne fonctionne que sur des modèles de taille assez réduite (moins d'une cinquantaine d'équations).

2 - ANAFI: une maquette pour le diagnostic financier

Le premier système ayant été réalisé en CRESUS est ANAFI, un système pour le diagnostic financier. Ce problème consiste à analyser la situation d'une entreprise à partir d'informations pour la plupart issues de documents comptables. C'est une application pour laquelle bon nombre de systèmes experts de première génération ont été réalisés. Ceci est paradoxal, car nous verrons que le traitement de ce problème en utilisant des connaissances de surface pose de sérieux problèmes. ANAFI, la maquette réalisée, se distingue de ces systèmes moins par ses performances que par l'originalité de sa conception: elle exploite des connaissances profondes sur le fonctionnement d'une entreprise. N'ayant pour but que de valider l'interprétation comparative sur une application concrète, elle a été réalisée sans aucun recours à un expert du domaine. Compte tenu de ce choix, les résultats produits par la maquette (cf. annexe A) reflètent assez bien l'intérêt de l'interprétation comparative. Ils n'atteignent cependant pas un niveau suffisant pour rendre ANAFI opérationnelle.

Le diagnostic financier est tout d'abord présenté. Les problèmes que posent son traitement par un système expert sont ensuite analysés. Puis sa formulation en terme d'interprétation comparative est explicitée; l'évaluation et l'explication comparatives sont alors examinées ainsi que leur coopération. Les résultats, limites et perspectives d'ANAFI sont enfin présentés.

2.1 - Le diagnostic financier

Le diagnostic financier a pour objectif l'analyse de la situation d'une entreprise. Ce problème joue un rôle critique dans maintes décisions. Il permet par exemple aux banquiers d'évaluer le risque d'un octroi de crédit ou aux actionnaires de juger la pertinence d'un investissement. Le diagnostic d'une entreprise est une tâche assez complexe, généralement confiée à un expert: l'analyste financier.

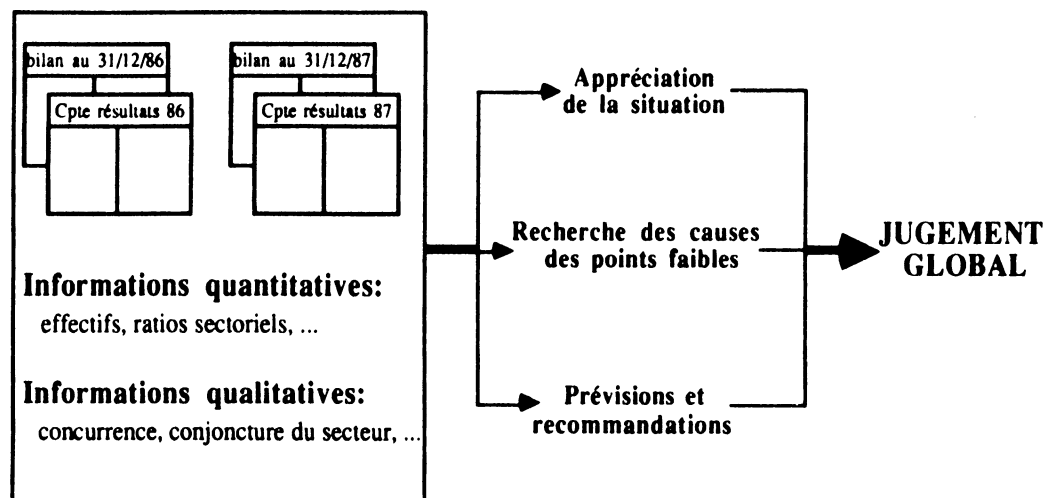


Figure 6.3: le diagnostic financier

Pour établir son diagnostic, l'analyste financier utilise plusieurs sources d'informations (figure 6.3). La principale est constituée par les documents comptables publiés chaque année par les entreprises, essentiellement le bilan et le compte de résultats. Il utilise aussi d'autres informations, qui sont de deux types: numériques et qualitatives. Les données numériques sont des variables concernant l'entreprise: effectifs, parts de marchés, etc ou son environnement: ratios du secteur d'activité, ... Les données qualitatives concernent également l'entreprise (efforts d'innovation, ...) et son environnement: agressivité de la concurrence, par exemple. L'évolution d'une entreprise représentant un aspect important de sa situation, l'expert collecte généralement ces données sur plusieurs années consécutives.

Une fois collectées, certaines informations font l'objet de redressements. Pour diverses raisons, une société peut en effet être amenée à présenter des résultats ne reflétant pas exactement sa situation économique. Afin de travailler sur des données significatives, l'analyste utilise son expérience pour corriger ces données. Ce point étant relativement marginal dans l'analyse, nous ne l'aborderons pas. A partir de cette étape, on peut considérer que le jugement final de l'expert résulte de trois étapes (cf. figure 6.3). Il apprécie d'abord les divers aspects du fonctionnement de l'entreprise: rentabilité, structure financière, etc. Ceci lui permet de mettre en évidence les points faibles (et bien sûr aussi les points forts) de l'entreprise. Puis, il analyse les faiblesses afin d'en rechercher les causes; il évalue ainsi la gravité des problèmes rencontrés par l'entreprise. Selon l'objectif de l'analyse, l'analyste est alors en mesure de formuler des prévisions et/ou des recommandations sur la situation de l'entreprise. Ces trois étapes sont maintenant développées.

2.1.1 - l'appréciation de la situation de l'entreprise

La méthode des ratios est l'une des techniques les plus fréquemment utilisée dans la pratique pour apprécier la situation d'une entreprise. L'idée en est simple: il s'agit de calculer un ensemble de ratios, c'est-à-dire de variables de la forme:

$$\frac{x}{y}$$

(où x et y sont des variables dont la valeur est connue), puis de comparer leur valeur à des normes. Un ratio judicieusement choisi évalue totalement ou partiellement un aspect de la situation de l'entreprise. Ainsi, le ratio:

$$\frac{\text{résultat_net}}{\text{capitaux}}$$

est un indicateur de la rentabilité des capitaux investis. La comparaison des ratios peut être effectuée par rapport à deux types de normes:

- des normes préétablies: un ratio de rentabilité des capitaux négatif constitue par exemple un indice défavorable sur la rentabilité de l'entreprise
- des normes sectorielles: la valeur d'un ratio de l'entreprise est comparée à celle de sociétés du même secteur d'activité

Etudier la situation de l'entreprise à un instant donné est insuffisant; l'appréciation de son évolution est également nécessaire. Pour cela, les variations des ratios doivent être analysées. Une augmentation du ratio de rentabilité des capitaux par exemple, constitue souvent un indice favorable de l'évolution d'entreprise.

Par le choix d'une batterie de ratios appropriée, l'analyste se forge donc une première opinion sur l'entreprise; en particulier, il cerne ses points faibles.

2.1.2 - La recherche des causes des points faibles

Détecter les points faibles de l'entreprise est insuffisant. Il convient de s'interroger sur leurs causes afin d'en évaluer la gravité. En effet, une baisse de rentabilité n'a pas les mêmes conséquences

sur l'avenir d'une entreprise si elle est due à des pertes exceptionnelles ou à des déséquilibres profonds dans la gestion.

L'étude de diagnostics réalisés par des experts montre que cette étape de l'analyse financière est effectuée grâce à un raisonnement sur un modèle du fonctionnement de l'entreprise ([BOUWMAN 83]). Ce raisonnement est mené qualitativement: la faiblesse de la rentabilité est analysée et non pourquoi elle a telle valeur précisément. Il apparaît également que deux types de modèles sont utilisés au cours de ce processus: des équations et des relations qualitatives. Les premières sont utilisées dans des raisonnements tels que: "les difficultés de trésorerie s'expliquent par le fait que les besoins en fonds de roulement ont augmenté alors que dans le même temps le fonds de roulement a diminué"; l'équation: $\text{trésorerie} = \text{fonds_de_roulement} - \text{besoin_en_fonds_de_roulement}$ est sous-jacente à ce raisonnement. Les relations qualitatives permettent quant à elles d'effectuer des raisonnements tels que: "le vieillissement de la gamme de produits explique la réduction des parts de marché".

2.1.3 - Prédications et recommandations

En général, l'intérêt porté à une société ne concerne pas tant sa situation actuelle que future. Aussi, un diagnostic financier se termine-t-il fréquemment par un pronostic sur l'avenir de l'entreprise analysée. De plus, si l'analyste est conseiller dans l'entreprise, ce pronostic est assorti de recommandations lui permettant d'améliorer sa situation. Cette étape fait appel aux résultats de l'étape précédente, mais surtout à l'expérience de l'analyste. C'est en effet en fonction des cas semblables auxquels a déjà été confronté ce dernier, qu'il sera en mesure de formuler des prédictions et recommandations pertinentes. Le traitement de cette étape par un système expert repose sur la manipulation de connaissances heuristiques traduisant l'expérience de l'analyste et non sur des connaissances profondes modélisant le fonctionnement d'une entreprise.

Le diagnostic financier exige donc deux types de compétences. D'une part, la manipulation de connaissances heuristiques pour la mise en oeuvre de la méthode des ratios et pour le pronostic sur l'avenir de l'entreprise; d'autre part, un raisonnement systématique sur le fonctionnement de l'entreprise afin de déterminer la cause de ses problèmes. C'est la conjonction de ces deux compétences qui fait du diagnostic financier une application complexe pour les systèmes experts.

2.2 - Les systèmes experts et l'analyse financière

De nombreux prototypes pour le diagnostic financier ont vu le jour depuis le développement des systèmes experts. La multiplication des maquettes masque néanmoins les difficultés de la conception d'un système complet traitant les trois étapes qui viennent d'être présentées. L'utilisation conjointe de connaissances heuristiques et profondes pour résoudre ce problème est, nous l'avons vu, la principale source de difficultés.

Des solutions basées sur les deux approches de conception de systèmes experts ont été proposées, mais aucune n'est à elle seule suffisante.

2.2.1 - Les systèmes experts de diagnostic financier de première génération

La plupart des systèmes experts réalisés pour le diagnostic financier exploitent les connaissances de surface de l'analyste: [MICHA 86], [RECHENMANN 87] par exemple. Cette approche conduit cependant à deux difficultés importantes.

La première difficulté, soulignée dans [PINSON 89], se manifeste lorsque l'on tente de coder la méthode des ratios par des règles. Deux types de règles sont nécessaires pour mener à bien ce travail: des règles jugeant individuellement la valeur d'un ratio et des règles agrégeant ces différents jugements. La difficulté provient de ces dernières; en supposant n ratios jugés selon 3 modalités: "faible", "moyen", et "élevé", l'agrégation des jugements requiert 3^n règles pour couvrir tous les cas pouvant se présenter. Ceci conduit rapidement à une base de connaissances de taille prohibitive, d'autant qu'il est nécessaire d'écrire un nombre de règles identique pour analyser l'évolution de l'entreprise (en supposant l'évolution de chaque variable également jugée selon trois modalités: "diminution", "stabilité" et "augmentation").

La seconde difficulté rencontrée par les systèmes experts utilisant des connaissances de surface concerne la recherche des causes des problèmes de l'entreprise. Cette opération nécessite le raisonnement sur des équations (cf. §2.1.2); il a été démontré dans le chapitre 1 combien ceci est délicat avec des règles.

Il est facile de comprendre que les systèmes utilisant exclusivement des connaissances de surface se heurtent au problème de l'écriture d'une base de connaissances de taille réaliste.

2.2.2 - Les systèmes experts de diagnostic financier utilisant des connaissances profondes

L'idée d'utiliser des connaissances sur le fonctionnement d'une entreprise pour réaliser un système expert de diagnostic financier n'est pas nouvelle. L'application du raisonnement qualitatif à un modèle où l'entreprise est vue comme un système hydraulique ([HELFERT 82]) a suscité plusieurs tentatives en ce sens: [BOUWMAN 83], [HART 86], [APTE 86]. Cette idée constitue un point de départ intéressant pour analyser la cause des difficultés de l'entreprise. Elle n'apporte cependant aucune solution à la coopération de cette tâche avec les connaissances heuristiques nécessaires au cours des autres étapes du diagnostic. Seule, elle est donc insuffisante pour réaliser un système expert complet de diagnostic financier. Nous allons examiner comment l'interprétation comparative permet de faire coopérer des connaissances de différentes natures pour résoudre ce problème.

2.3 - Diagnostic financier et interprétation comparative: ANAFI

Le diagnostic financier peut être formulé comme un double problème d'interprétation comparative: figure 6.4. La situation de l'entreprise est comparée à des normes sectorielles, ce qui permet de la situer par rapport à des entreprises similaires. Elle est également comparée à une situation antérieure afin d'en analyser l'évolution. Il faut remarquer que la restriction à la comparaison de deux états interdit l'analyse de l'évolution d'une entreprise sur plusieurs années consécutives. La comparaison sera donc effectuée entre la première et la dernière année où les informations sont disponibles sur l'entreprise.

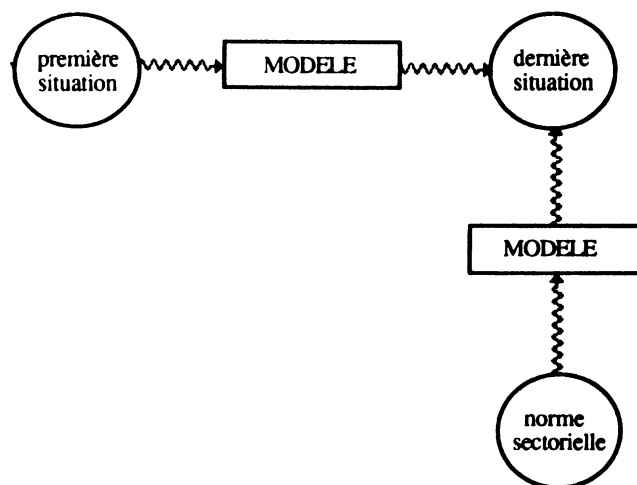


Figure 6.4: Diagnostic financier et interprétation comparative

L'évaluation comparative est l'appréciation des divers aspects du fonctionnement de l'entreprise et de son évolution. Elle a pour but de mettre en évidence les points faibles de l'entreprise (des écarts) soit par rapport au secteur, soit par rapport à la situation antérieure. La recherche des causes de ces faiblesses forme l'explication comparative du problème. La dernière étape du diagnostic financier, la composante prédictions/recommandations ne s'insère pas dans le schéma de l'interprétation comparative. Faisant appel à des connaissances heuristiques, elle est plus aisément traitée par un système à base de règles par exemple. Nous reviendrons sur ce point dans le paragraphe 2.3.3 afin de voir comment un tel système peut coopérer avec l'interprétation comparative.

2.3.1 - L'évaluation comparative

L'appréciation de la situation d'une entreprise, tant par rapport à son passé qu'à son secteur d'activité, peut être décrite de manière arborescente, ainsi que l'illustre la figure 6.5. La situation dépend de trois facteurs: la rentabilité, la structure financière et l'activité. Chacun de ces facteurs dépend à son tour d'autres aspects du fonctionnement de l'entreprise, etc. Ce processus prend fin avec les ratios qui sont les feuilles de l'arborescence.

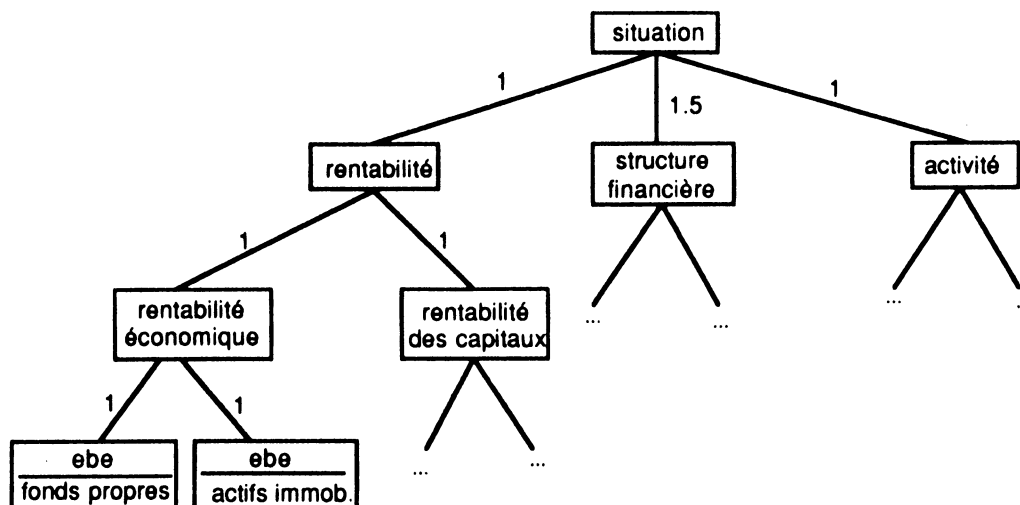


Figure 6.5: Un exemple d'arbrescence de variables pour apprécier la situation d'une entreprise

Cette arbrescence mêle deux types de variables: des variables intermédiaires qualitatives (situation, activité, ...) et des variables terminales numériques (ratios). L'évaluation comparative consiste à juger ces différentes variables par un mécanisme numérique hiérarchique appelé notation. A chaque variable sont associés deux **notes** qui sont des nombres réels dans l'intervalle $[-5,+5]$. La première (appelée note sectorielle) traduit l'écart de la variable par rapport au secteur d'activité, la seconde (appelée note d'évolution) traduit l'écart de la variable par rapport à la situation antérieure.

La notation des variables débute avec les ratios (feuilles de l'arbrescence). La note sectorielle d'un ratio est obtenue par comparaison de la valeur du ratio de l'entreprise par rapport à celle du secteur d'activité. Une note sectorielle proche de 0 traduit une valeur "normale"; une note proche de -5 (resp. +5), une valeur très inférieure (resp. supérieure) à la normale. La note d'évolution d'un ratio est calculée en fonction de son taux de variation. Une note d'évolution proche de 0 indique une valeur stable; une note proche de -5 (resp. +5), une forte diminution (resp. augmentation). Les notes sont ensuite remontées dans l'arbrescence en considérant pour une variable la moyenne des notes des variables subordonnées, pondérées par des coefficients d'importance (les nombres étiquetant les arcs de l'arbrescence sur la figure 6.5). Il faut noter que certaines informations qualitatives constituant des données de l'entreprise: efforts d'innovation, ... échappent à ce processus arborescent; leurs deux notes sont directement fournies au système.

Le mécanisme de notation est implanté dans le module d'évaluation comparative d'ANAFI sous forme d'un langage comportant deux primitives: les échelles et les fonctions de notation.

Les échelles de notation attribuent aux ratios leurs deux notes. Une échelle de notation est une expression: $[V_1 \rightarrow N_1 \dots V_n \rightarrow N_n]$ telles que $V_1 < V_2 < \dots < V_n$ et $N_1 = -5 < N_2 < \dots < N_n = 5$, où les V_i sont des valeurs et les N_i les notes qui leur sont associées. Deux échelles sont associées à chaque variable: la première pour le calcul de la note sectorielle, la seconde pour le calcul de la note d'évolution. La notation d'une variable X au moyen d'une échelle s'écrit dans ANAFI:

$$X \text{ vaut } [V_1 \rightarrow N_1 \dots V_n \rightarrow N_n] [V_1' \rightarrow N_1' \dots V_n' \rightarrow N_n']$$

Le calcul de la note sectorielle (resp. d'évolution) N d'une variable ayant une valeur V (resp. un taux de variation V), illustré sur la figure 6.6, est effectué de la manière suivante:

- si $V \leq V_1$ alors $N = -5$
- si $V \in [V_i, V_{i+1}[$ alors $N = N_i + \frac{(V - V_i)}{(V_{i+1} - V_i)} (N_{i+1} - N_i)$
- si $V \geq V_n$ alors $N = 5$

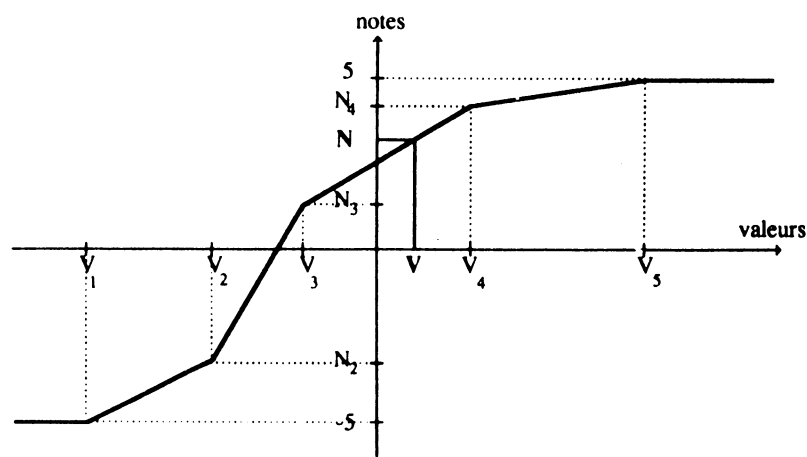


Figure 6.6: calcul d'une note sur une échelle de notation

Supposons ainsi le ratio de rentabilité des capitaux d'une entreprise égal à 0,02 avec un taux de variation de 5%, jugé sur les deux échelles: $[-0,02 \rightarrow -5 \quad 0,03 \rightarrow 0 \quad 0,10 \rightarrow 5]$, pour la note sectorielle et $[-0,2 \rightarrow -5 \quad 0 \rightarrow 0 \quad 0,2 \rightarrow 5]$ pour la note d'évolution. Sa note sectorielle est -1 et sa note d'évolution est 1,25.

Les fonctions de notation forment la deuxième primitive du langage d'évaluation comparative d'ANAFI. Destinées à la notation des variables intermédiaires de l'arborescence, elles ont pour forme:

$$X \text{ vaut } p_1 \cdot X_1 + p_2 \cdot X_2 + \dots + p_n \cdot X_n$$

où les X_i sont les variables intervenant dans la notation de la variable X et les p_i les coefficients de pondération (éventuellement négatifs) qui leur sont associés.

La note sectorielle (resp. d'évolution) N de la variable qualitative X est calculée par l'expression:

$$N = \frac{\sum_{i=1}^n p_i N_i}{\sum_{i=1}^n |p_i|} \quad \text{où } N_i \text{ est la note sectorielle (resp. d'évolution) de la variable } X_i$$

La variable situation, racine de l'arborescence de la figure 6.5, est ainsi décrite par la fonction de notation:

$$\text{situation vaut } 1 \cdot \text{rentabilité} + 1.5 \cdot \text{structure_financière} + 1 \cdot \text{activité}$$

L'agrégation numérique des variables autorise un codage de la méthode des ratios beaucoup plus condensé qu'une agrégation symbolique par des règles (cf. §2.2.1): une fonction de notation suffit où nombre de règles serait nécessaire. Elle permet également la jonction entre l'évaluation et l'explication comparatives pour les variables qualitatives.

2.3.2 - L'explication comparative

L'explication comparative utilise le même modèle pour les deux aspects de la comparaison de la situation de l'entreprise. Ce modèle est composé d'équations comptables et d'un réseau d'influences.

Les équations comptables sont définitionnelles et forment un système exempt de boucle de rétroaction. Elles sont de deux types: des identités (ex: valeur_ajoutée = production_globale - consommations_intermédiaires) et des définitions de ratios (ex: ratio_de_liquidité_générale = $\frac{\text{capitaux_circulants}}{\text{dettes_a_court_terme}}$). Mis à part les ratios, les variables numériques ne reçoivent pas de note; les équations ne permettent donc aucune comparaison sectorielle et sont exclusivement utilisées pour expliquer l'évolution des ratios.

Les équations ne décrivent que la partie numérique du fonctionnement de l'entreprise. Ce dernier a également une composante qualitative. Celle-ci est représentée dans ANAFI par un réseau d'influences conditionnel qui lie les variables qualitatives de l'entreprise. L'explication comparative est effectuée sur les variables qualitatives à partir des notes qui leur sont associées grâce au processus suivant:

- 1 - Les conditions éventuelles de toutes les relations d'influences sont d'abord évaluées. Lorsque la condition d'une relation d'influence n'est pas vérifiée, cette relation est retirée du réseau. On se ramène ainsi à un réseau d'influences simple.
- 2 - Les notes des variables sont ensuite traduites sous forme d'un écart afin que puisse être appliquée l'explication comparative. Cette opération est basée sur l'idée suivante: seules les notes s'écartant significativement de 0 méritent d'être expliquées. L'écart sectoriel (resp. d'évolution) ΔX d'une variable X est défini à partir de la note sectorielle (resp. d'évolution) $N(X)$ de la manière suivante:

- si $N(X) \in [-5, -1[$ alors $\Delta X = -$
- si $N(X) \in [-1, 1]$ alors $\Delta X = 0$
- si $N(X) \in]1, 5]$ alors $\Delta X = +$

On se ramène ainsi à un problème d'explication comparative sur un réseau d'influences avec des variables qualitatives tel qu'il a été vu au chapitre 4. Lorsque l'écart sectoriel et/ou d'évolution d'une variable est différent de 0, l'explication comparative est déclenchée afin d'en déterminer les causes.

2.3.3 - Coopération entre évaluation et interprétation comparatives

ANAFI produit un rapport qui constitue le dossier d'analyse financière en combinant l'évaluation et l'explication comparatives. Ce rapport est généré par un parcours (en largeur d'abord) de l'arborescence des variables. A chaque variable traversée est associé un commentaire sur sa note sectorielle et sa note d'évolution par un mécanisme analogue à la génération de phrases pour l'explication comparative (cf. §1.2.4). Lorsqu'une variable présente un écart sectoriel ou d'évolution différent de 0, l'explication comparative est déclenchée sur cet (ou ses) écart(s), puis propagée aux variables explicatives, jusqu'à ce que les variables exogènes soient atteintes. La propagation de l'explication

comparative aux variables explicatives souffre toutefois deux exceptions: lorsque l'écart d'une variable a déjà été expliqué ou lorsqu'il doit être étudié ultérieurement, c'est-à-dire lorsque la variable explicative est un noeud de l'arborescence de notation non encore exploré.

Le mécanisme de notation constitue l'élément central de la coopération entre connaissances profondes et heuristiques dans ANAFI. Notons cependant qu'il n'est pas indispensable que l'évaluation comparative s'effectue numériquement pour que l'explication comparative soit applicable. Il est seulement nécessaire que chaque variable qui apparaît dans le modèle servant à l'explication comparative prenne ses valeurs dans un totalement ordonné (cf. chapitre 3 - §2.2.3) afin que son écart soit défini. Il est donc possible de faire coopérer l'explication comparative avec un module d'évaluation comparative basé sur des règles de production respectant cette contrainte. A l'inverse, on peut envisager un système à base de règles manipulant les écarts des variables, en aval de l'explication comparative. Ceci semble d'ailleurs la solution adéquate pour réaliser le module de prédictions/recommandations qui constitue la dernière étape du diagnostic financier, non traitée dans ANAFI.

2.4 - Résultats, limitations et perspectives

ANAFI met en évidence la puissance du cadre conceptuel fourni par l'interprétation comparative en démontrant la possibilité d'intégrer dans un même système des connaissances heuristiques (évaluation comparative pour la méthode des ratios) et des connaissances profondes de différentes natures (équations et réseau d'influences). Les résultats n'atteignent certes pas un niveau suffisant pour être utilisables dans la pratique, mais sont très encourageants compte tenu du peu de connaissances contenues dans le système. La version actuelle d'ANAFI ne comporte en effet qu'une trentaine de relations d'influences et une quarantaine d'équations. Ceci est très largement insuffisant pour décrire finement les divers aspects du fonctionnement d'une entreprise.

La réalisation d'ANAFI a également mis en évidence une limite importante de l'explication comparative sur des réseaux d'influences. Ces derniers constituent un outil simple pour modéliser le fonctionnement qualitatif d'une entreprise, mais de trop bas niveau pour le diagnostic financier. Dans un système aussi complexe qu'une entreprise, un événement est généralement interprétable à différents niveaux, qu'un réseau d'influences ne permet pas de distinguer. Considérons ainsi l'exemple de l'annexe A: les variables du cycle d'exploitation influant sur les liquidités constituent un premier niveau d'analyse. D'autres niveaux pourraient toutefois s'avérer nécessaires pour fournir des explications pertinentes. Pour exprimer, par exemple, qu'un vol dans une entreprise peut affecter ses liquidités, il faudrait ajouter la

relation d'influence: vol l- liquidité. A cause de cet ajout cependant, les variables vol et durée_phase_distribution risquent de se trouver mêlées dans une même explication alors qu'elles concernent des niveaux d'analyse totalement différents. Il faudrait pouvoir considérer les différentes relations d'influences ne s'appliquant qu'à un niveau donné. La notion de processus qualitatif ([FORBUS 84], cf. chapitre 2) semble tout à fait adéquate pour résoudre ce problème. En regroupant les relations d'influences d'un niveau donné dans un processus qualitatif, il deviendrait possible de proposer différents niveaux d'explication. Ceci nous amène à suggérer une extension importante du travail entrepris: l'application de l'explication comparative aux processus qualitatifs.

3 - SURPLUS: un système expert pour l'analyse de la performance d'une entreprise

L'étude de la performance de l'entreprise par la méthode des surplus ([VINCENT 68], [MORVAN 85]) fait partie des développements assez récents de l'analyse financière. Elle permet de déterminer si une entreprise améliore la qualité de sa gestion et d'analyser les rapports de force qu'elle entretient avec son environnement.

Comme le diagnostic financier, cette tâche incombe à un spécialiste: l'expert comptable ou l'analyste financier. Elle nécessite plusieurs heures de travail, et bien que la méthode des surplus comporte une étape de calculs pouvant être traitée de manière algorithmique, le travail essentiel est l'interprétation des résultats numériques qu'elle fournit. L'interprétation étant menée à partir d'équations, l'utilisation de connaissances de surface est peu adaptée à ce problème (cf. chapitre 1). D'ailleurs, aucun système expert de première génération n'a été réalisé pour ce problème (à notre connaissance).

SURPLUS, un système expert pour l'analyse de la performance d'une entreprise par la méthode des surplus a été réalisé. Comme ANAFI, il utilise l'interprétation comparative, mais s'insère plus simplement dans ce schéma. La composante évaluation comparative de l'interprétation est en effet inutile et le modèle pour l'explication est réduit à un système d'équations définitionnelles. Par conséquent, les résultats obtenus (cf. annexe B) sont de meilleure qualité et permettent d'envisager une application opérationnelle de ce système.

Cette section présente le système expert SURPLUS. Le problème traité, l'analyse de la performance d'une entreprise par la méthode des surplus, est tout d'abord présenté. Sa formulation en

terme d'interprétation comparative est ensuite examinée, avant que son implantation en CRESUS ne soit décrite. Les résultats obtenus et les extensions possibles sont enfin analysés.

3.1 - La méthode des surplus¹

L'objectif de la méthode des surplus est à la fois plus limité et plus large que celui du diagnostic financier. Il est plus limité dans la mesure où il vise seulement à l'analyse des performances de l'entreprise. Il est aussi plus large car il est applicable au plan macro-économique et n'est donc pas réduit à l'analyse d'une entreprise (ou d'une branche d'entreprises). Dans SURPLUS, la méthode est appliquée à l'analyse de la performance d'une entreprise.

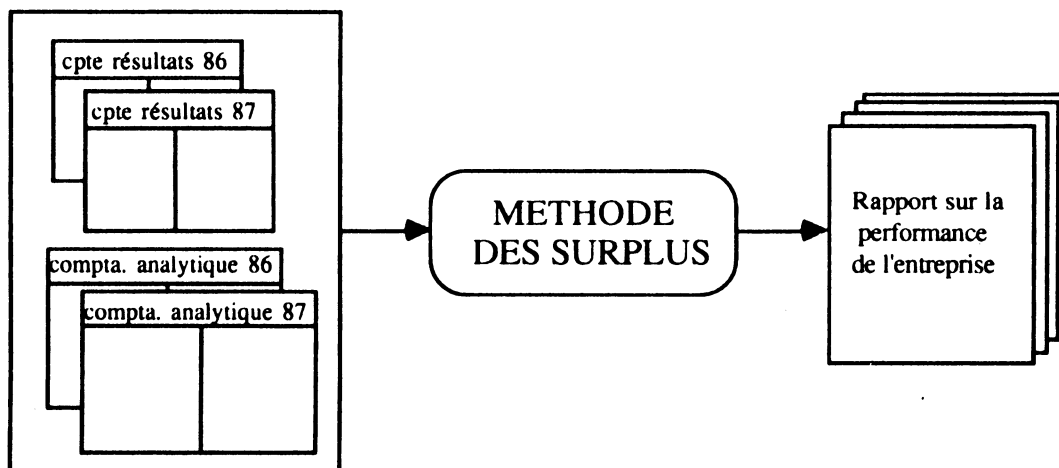


Figure 6.7: l'analyse de la performance d'une entreprise par la méthode des surplus

Un rapport sur la performance de l'entreprise est produit à partir de l'étude de ses documents comptables (figure 6.7). Le compte de résultats qui regroupe les chiffres caractéristiques de l'exploitation de l'entreprise (produits et charges) constitue le point de départ de la méthode des surplus. Les comptes de deux exercices consécutifs sont nécessaires à l'analyse. Comme les informations y figurant ont un niveau de détail insuffisant pour effectuer une analyse fine, la comptabilité analytique de l'entreprise qui comprend des données beaucoup plus détaillées est également requise sur les deux exercices.

¹ Cette présentation de la méthode des surplus est adaptée de [MORVAN 85]

L'excédent brut d'exploitation (EBE), l'un des postes du compte de résultats, fournit une mesure précise de l'enrichissement de l'entreprise dû aux facteurs de production. Il ne constitue par pour autant un indicateur suffisant de la performance de l'entreprise. Il exprime en effet la différence en valeur entre les produits, et les charges utilisées pour obtenir ces produits. Or, pour analyser la performance de l'entreprise, il faut distinguer les gains d'**efficacité** et les gains de **rentabilité**. Les premiers résultent de variations de **quantité** de biens produits et de facteurs de production utilisés. Les seconds résultent de variations de **prix**.

Pour étudier ces deux aspects, un indicateur appelé "surplus de productivité global" (SPG), ou "surplus" simplement, est utilisé. La méthode, centrée sur le SPG, se déroule en deux étapes. La formation du surplus qui traduit les variations de quantités entre les deux exercices est tout d'abord analysée afin d'évaluer l'efficacité de l'entreprise. Sa répartition, se traduisant par des variations de prix, est ensuite examinée pour déterminer qui (fournisseurs, clients, salariés, ...) bénéficie du surplus formé.

3.1.1 - La formation du surplus

La comparaison en quantité des variations de biens ou services produits P_1, P_2, \dots, P_n entre deux exercices 1 et 2 avec les facteurs de productions utilisés (travail, matières premières, ...) F_1, F_2, \dots, F_m a pour but de mettre en évidence le gain (ou la perte) d'efficacité de l'entreprise. Un indicateur, le surplus (SPG), est utilisé à cette fin; il est défini par la relation:

$$\text{SPG} = \Sigma \text{variations de quantités des produits } P_i - \Sigma \text{variations de quantités des facteurs } F_i$$

Si, la variation entre les deux exercices de quantité de produits dépasse celle de facteurs utilisés, le SPG est positif. Cela traduit une amélioration de la productivité de l'entreprise. A l'inverse, un SPG négatif indique une dégradation de sa productivité.

Les éléments intervenant dans l'expression du SPG doivent être exprimés dans la même unité pour être comparables. Aussi, le SPG est calculé, non pas à partir de variations de quantité, mais de volume; c'est-à-dire des produits (quantité \times prix). L'unité commune est alors le franc. Pour ne pas introduire de biais à cause des variations de prix, le calcul est mené en supposant le prix de chaque produit et facteur constant, exprimé en francs constants de l'année 1. Le SPG s'écrit donc:

$$SPG = \sum p_i^1 \Delta P_i - \sum f_i^1 \Delta F_i$$

où p_i^1 et f_i^1 sont le prix des produits et des facteurs à l'année 1, ΔP_i et ΔF_i leur variation en quantité entre les exercices 1 et 2.

L'explication de la formation du surplus constitue la partie "intelligente" de cette première étape. Elle rend compte des éléments ayant favorisé ou empêché l'amélioration de la productivité de l'entreprise. Elle examine la contribution de chacun des produits et des facteurs de production à la formation du surplus. Considérant par exemple le facteur de production travail, sa contribution au surplus est: $\text{-salaire_horaire}^1 \times \Delta \text{nb_heures_travaillées}$ (l'exposant 1 indique qu'il s'agit de la valeur dans l'exercice 1). Cette analyse peut être poursuivie jusqu'à un niveau très fin pour chaque produit et facteur. Ainsi, pour le travail, il est possible d'étudier séparément la productivité d'une catégorie d'employés, etc.

3.1.2 - La répartition du surplus

Le surplus formé se trouve distribué à travers des variations de prix entre les diverses parties prenantes de l'entreprise (clients, salariés, fournisseurs, ...). Ces variations de prix traduisent les rapports de force entre l'entreprise et ses partenaires. L'étude de la répartition du surplus permet donc d'analyser comment l'entreprise maîtrise son environnement.

A partir des identités comptables, il est aisé de démontrer la relation suivante:

$$SPG = -\sum P_i^2 \Delta p_i + \sum F_i^2 \Delta f_i + \Delta EBE$$

Cette égalité peut être interprétée de la manière suivante: le surplus sert, outre la formation du bénéfice (ΔEBE), à récompenser les partenaires de l'entreprise pour le gain de productivité qu'ils ont permis d'effectuer. Chaque terme $F_i^2 \Delta f_i$ représente en effet un avantage procuré au facteur F_i , sous forme d'une augmentation de prix; dans le cas du travail, par exemple, l'avantage accordé aux salariés est égal à l'augmentation de salaire horaire multiplié par le nombre d'heure dépensé au cours de l'exercice 2. De même, chaque terme $-P_i^2 \Delta p_i$ représente un avantage procuré aux clients sous forme d'une baisse de prix du produit P_i . Le SPG peut donc être reformulé de la manière suivante:

$$\text{SPG} = \Sigma \text{avantages accordés à la clientèle} - \Sigma \text{avantages accordés aux facteurs} + \Delta \text{EBE}$$

Si un facteur de production contribue à la formation du surplus, il est logique qu'il s'en trouve récompensé par un avantage. Cependant, l'entreprise peut utiliser son pouvoir pour modifier cet ordre des choses. Ainsi, si l'un des facteurs f est tel que Δf est négatif, cela signifie qu'il a subi une perte (et donc l'entreprise un gain).

Une fois calculée la part accordée à (ou cédée par) chaque partenaire de l'entreprise, l'interprétation de ces rapports de force constitue la partie "intelligente" de l'analyse de la répartition. De manière duale à la formation du surplus, elle consiste à analyser les variations de prix expliquant la part de chacun.

3.2 - La méthode des surplus et l'interprétation comparative

Les deux étapes de la méthode des surplus peuvent être mises sous forme d'un problème d'interprétation comparative. L'analyse de la formation du surplus est en effet l'explication des variations de quantités de produits et de facteurs entre les deux exercices, tandis que l'analyse de sa répartition est une explication en terme de prix. Les données numériques des deux exercices étant directement comparables, l'évaluation comparative est inutile.

3.2.1 - Le modèle

Le modèle utilisé pour les deux aspects de la méthode des surplus est identique. Il comprend deux parties: l'une générale à toutes les entreprises, l'autre spécifique à l'entreprise analysée. Le modèle général est constitué d'équations définitionnelles liant les postes du compte de résultats. C'est donc un sous-ensemble du modèle utilisé pour le diagnostic financier. Le modèle spécifique détaille les postes du compte de résultats qu'il est nécessaire d'étudier séparément pour une analyse fine de la performance de l'entreprise. Ces postes se décomposent de manière hiérarchique. La figure 6.8 présente un exemple d'une telle décomposition.

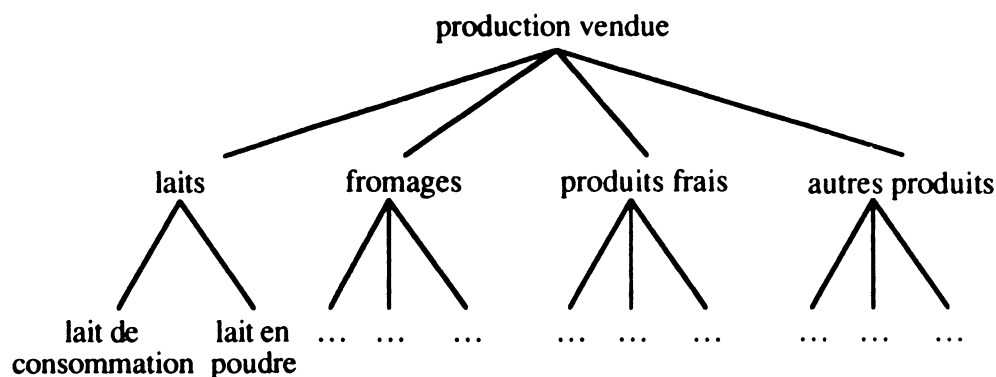


Figure 6.8: hiérarchie du poste production vendue d'une entreprise de produit laitier

Comme les ventes d'un élément de cette arborescence sont égales en valeur à la somme de tous les composants qu'il regroupe, il est possible d'écrire une équation pour chaque élément (non terminal) de l'arborescence. Ainsi:

$$\begin{aligned}
 \text{production_vendue} &= \text{vente_laits} + \text{vente_fromages} + \text{vente_produits_frais} + \text{vente_autres_produits} \\
 \text{vente_laits} &= \text{vente_lait_consommation} + \text{vente_lait_poudre} \\
 &\dots
 \end{aligned}$$

Ces équations constituent le modèle spécifique de l'entreprise traitée. L'explication comparative pour la formation et la répartition du surplus utilise l'ensemble des équations des modèles général et spécifique.

3.2.2 - La formation du surplus

La formation du surplus est formulée comme un problème d'explication comparative à partir de la remarque suivante. Imaginons, à la place du compte de résultats de l'exercice 2, un compte fictif où les divers éléments au lieu d'être valorisé au prix de l'année 2 le sont au prix de l'année 1. La variation de l'EBE entre l'exercice 1 et l'exercice fictif s'exprimerait alors:

$$\begin{aligned}
 \Delta \text{EBE} &= \text{EBE}^{\text{fictif}} - \text{EBE}^1 \\
 &= \left(\sum p_i^1 P_i^2 - \sum f_i^1 F_i^2 \right) - \left(\sum p_i^1 P_i^1 - \sum f_i^1 F_i^1 \right) \\
 &= \sum p_i^1 \Delta P_i - \sum f_i^1 \Delta F_i \\
 &= \text{SPG}
 \end{aligned}$$

L'analyse de la formation du surplus est donc l'explication comparative de la variable EBE entre deux états (figure 6.9). L'état de référence est réel; il est constitué par les résultats de l'entreprise à l'année 1. L'état dérivé est fictif; il est construit à partir des quantités de l'année 2 valorisés aux prix de l'année 1.

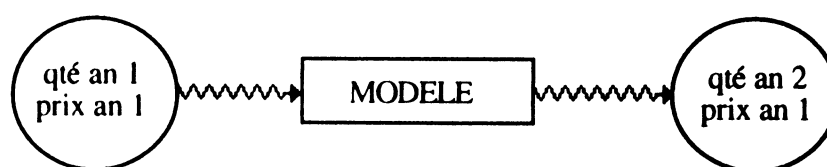


Figure 6.9: formation du surplus et interprétation comparative

3.2.3 - La répartition du surplus

La répartition du surplus est formulée comme un problème d'explication comparative de manière similaire. Au lieu du compte de résultats de l'exercice 1, considérons le compte de résultats fictif précédemment défini. La variation de l'EBE entre les deux exercices serait:

$$\begin{aligned}
 \Delta EBE &= EBE^2 - EBE^{\text{fictif}} \\
 &= (\sum p_i^2 P_i^2 - \sum f_i^2 F_i^2) - (\sum p_i^1 P_i^2 - \sum f_i^1 F_i^2) \\
 &= \sum P_i^2 \Delta p_i - \sum F_i^2 \Delta f_i \\
 &= -\sum \text{avantages distribués}
 \end{aligned}$$

La répartition du surplus est donc également l'explication comparative de la variable EBE entre deux états (figure 6.10). L'état de référence est l'état dérivé utilisé pour l'explication de la formation du surplus, construit à partir des quantités de l'année 2 valorisées aux prix de l'année 1. L'état dérivé est constitué des résultats de l'année 2.

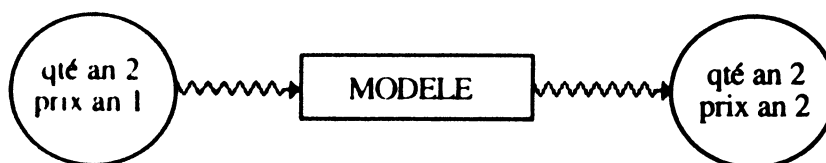


Figure 6.10: répartition du surplus et interprétation comparative

3.3 - Implantation

Une fois mis sous forme de problème d'interprétation comparative, la méthode des surplus est directement implantable en CRESUS.

Le modèle général comporte une dizaine d'équations. Les valeurs des variables intervenant dans ces équations sont récupérées automatiquement d'un programme de saisie de liasses fiscales. Le modèle spécifique à l'entreprise n'est pas introduit sous forme d'équations. Un outil de description de nomenclatures a été écrit afin de rendre sa saisie plus rapide. Les équations sont générées automatiquement à partir de ces nomenclatures (cf. §3.1). Il permet également de décrire pour chaque élément les informations nécessaires à la génération du rapport (expression française, nombre d'unités, etc), ainsi que la décomposition en prix et quantité, pour appliquer la méthode. Par ailleurs, des modules de tracé de tableaux et de diagrammes ont été ajoutés, afin que le rapport généré soit comparable au travail effectué par l'expert comptable.

3.4 - Résultats et extensions

Les résultats fournis par le système SURPLUS laissent entrevoir une utilisation réelle par des experts comptables. Le système fournissant l'essentiel de l'analyse, le comptable n'aurait qu'à corriger le style trop frustré du rapport et à ajouter ses propres remarques.

Deux extensions importantes sont possibles à partir du système actuel. La première consiste à doter SURPLUS d'un modèle qualitatif lui permettant de fournir des analyses plus fines. Ceci est utile pour expliquer de brusques variations de certaines variables exogènes, qui ne sont appréciables que qualitativement. Il deviendrait ainsi possible d'analyser si une diminution soudaine des stocks est due à une cause accidentelle ou résulte d'une décision de l'entreprise, par exemple. Cette extension est réalisable en plaçant en amont des variables exogènes un réseau d'influences représentant leurs causes potentielles de variation; on retombe cependant ainsi dans le problème mentionné à propos d'ANAFI: l'impossibilité de fournir différents niveaux d'explications. Le second prolongement vise à introduire un facteur supplémentaire dans l'analyse de la performance: la qualité. Cette extension est cependant délicate à introduire car des problèmes se posent déjà pour la décomposition en prix et quantité de certains éléments; la difficulté pour mesurer la qualité de chaque produit et facteur utilisé aurait donc pour effet

d'amplifier ces problèmes. Les nombreux travaux actuellement menés en Gestion sur ce sujet laissent cependant espérer une solution.

Notons enfin que SURPLUS permet d'envisager une extension de l'application de la méthode des surplus. Etant assez longue à mettre en oeuvre manuellement, elle est généralement utilisée à des fins d'analyse rétrospective. S'il devient possible de simuler l'effet d'une décision sur la performance de l'entreprise, elle peut également devenir un outil d'aide à la décision important.

CHAPITRE 7

SICA: UN SYSTEME EXPERT POUR L'IMPUTATION COMPTABLE

La tenue des comptes d'une entreprise apparaît souvent aux personnes étrangères à la comptabilité comme un problème totalement résolu par l'informatique. Ce n'est pourtant pas le cas; la comptabilité comporte en effet deux aspects différents, comme le montre la figure 7.1. Le premier, l'imputation, consiste à déterminer les écritures comptables associées aux pièces émises et reçues par l'entreprise: factures, quittances, avoirs, ... Les écritures font ensuite l'objet de traitements afin de constituer les documents comptables de l'entreprise: balance, grand-livre, etc. Ce dernier problème est totalement structuré. C'est à lui seul que fait référence l'informatisation de la comptabilité. L'imputation, en revanche n'est pas structurée. Elle exige toujours l'intervention d'un expert, le comptable.

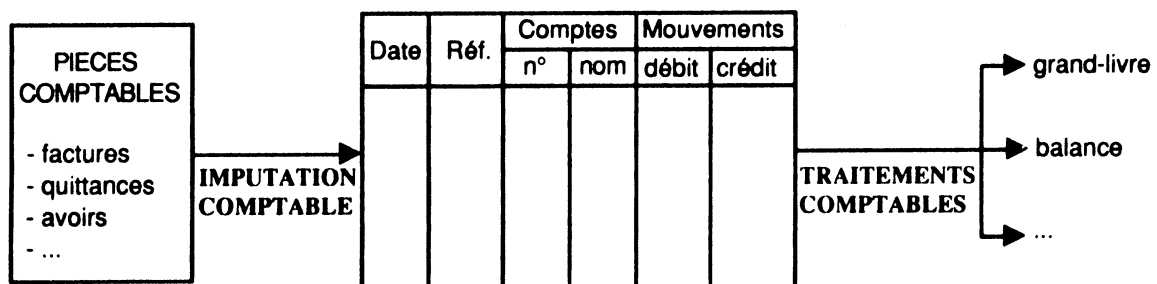


Figure 7.1: la tenue des comptes d'une entreprise

Nous avons développé un système expert pour ce problème: SICA (Système d'Imputation Comptable Automatique). Il analyse des phrases décrivant des pièces comptables et détermine l'écriture associée en reconstituant l'opération économique sous-jacente.

Après avoir décrit l'imputation comptable en section 1, les réalisations de systèmes experts dans ce domaine sont présentées en section 2. Le système SICA fait ensuite l'objet de la section 3.

1 - L'imputation comptable

Les documents commerciaux émis et reçus par l'entreprise constituent l'information de base pour l'imputation. Ils sont traités par le comptable qui détermine pour chacun d'eux l'écriture associée et l'inscrit sur une pièce spéciale appelée bordereau d'imputation (d'autres techniques existent également), ainsi que l'illustre la figure 7.2.

SOCIETE POUETTE 12 rue de Bonne 38000 Grenoble				
Doit: Ets DUGLOUX rue de la République 38000 GRENOBLE		Facture n°565 Grenoble, le 3/6/1988		
REF.	DESIGNATION	Qte	P. U.	Montant H. T.
R56	câbles de transmission x908	10	56,00	560,00
Q21	connecteurs K/80-C9	10	224,00	2240,00
TOTAL				2800,00
DEBOURS				0,00
T.V.A 18,6%				520,80
NET A PAYER				3320,80

Date	Réf.	Comptes		Mouvements	
		n°	nom	débit	crédit
...
...
3/6/1988	565	607	achat mar	2800,00	
3/6/1988	565	445	TVA	520,80	
3/6/1988	565	400	fournisseurs		3320,80
...
...

Figure 7.2: l'imputation d'une facture d'achat sur bordereau chez Dugloux

Une écriture se compose de plusieurs mouvements (chaque ligne du bordereau d'imputation). Un mouvement est essentiellement un triplet:

(n° de compte, sens: débit ou crédit, montant)

Les autres informations se déduisent en effet directement: la date et la référence d'après la pièce traitée et le nom du compte d'après son numéro.

La déduction des mouvements associés à une opération est l'objet de la technique comptable. L'imputation de la facture d'achat ci-dessous par exemple, donne lieu chez Dugloux à une écriture comportant trois mouvements: le premier enregistre en débit le montant hors-taxe de l'achat de marchandises dans le compte prévu à cet effet: "607 - achats de marchandises"; le second enregistre en

débit le montant de la TVA. Enfin, le troisième constate par un crédit la dette envers le fournisseur pour le montant total de l'opération.

Il faut noter que le numéro de compte d'un mouvement dépend parfois du plan comptable adopté par l'entreprise. Les comptes forment en effet une nomenclature dont les entreprises peuvent choisir, dans une certaine mesure, le niveau de détail. Les imputations peuvent donc varier légèrement d'une entreprise à l'autre. Signalons enfin, qu'une fois un plan comptable retenu, une pièce comptable n'a qu'une imputation possible (il existe malgré tout quelques exceptions).

2 - Systèmes experts et imputation comptable

La réalisation d'un système expert déduisant les écritures vise à faire traiter l'imputation par un personnel moins qualifié que le comptable. Même s'il ne sait pas imputer, un employé chargé de la saisie pour les traitements informatiques comptables par exemple, sait généralement reconnaître les différentes pièces. Il suffit donc de lui en faire décrire le contenu au système expert qui se charge ensuite de l'imputation; l'automatisation de la tenue des comptes serait alors totale.

Deux prototypes ont déjà été réalisés autour de cette idée. Le premier, présenté dans [GERVAIS 86] et [STEPNIEWSKI 87] est un système expert à base de règles de productions. Le second dénommé KIWI est décrit dans [CLARKE 75]; il utilise un modèle logique. Comme nous allons le voir cependant, aucun de ces systèmes n'est en mesure d'atteindre un stade opérationnel.

2.1 - Le prototype de Stepniewski et Gervais

Ce système détermine les écritures associées à des phrases décrivant des opérations de l'entreprise. Les connaissances en comptabilité nécessaires à l'imputation sont codées par des règles de production dans une base écrite en SNARK [LAURIERE 82b]. Une règle de ce système tirée de [GERVAIS 86] est présentée ci-dessous.

Règle 1

si le document est une facture
et cette facture a été émise par nous
et la mention «doit» existe
alors l'opération est une vente et le tiers qui apparaît après le terme «doit» est: un client.

Le moteur d'inférence détermine les imputations en trois étapes. Il déduit d'abord la nature de l'opération et le tiers qu'elle concerne. Il déduit ensuite les numéros des comptes à mouvementer. Le passage de l'écriture correspondant à l'opération est enfin effectué.

Cette démarche semble difficilement utilisable pour réaliser un système expert d'imputation comptable complet. Elle conduit en effet à une base de connaissances très complexe. Considérons par exemple la règle 35 ci-dessous (d'après [GERVAIS 86]) traitant la vente d'un produit fini. Des dizaines de règles similaires sont nécessaires pour traiter les différentes combinaisons des variantes de cette opération: l'escompte peut ne pas être accordé, le port non facturé, les emballages non consignés, etc. Chacun de ces cas conduit à une écriture différente et donc requiert une règle différente. De plus, des règles symétriques doivent être écrites pour le cas dual: l'achat de matières premières. La représentation des connaissances de surface du comptable est donc très peu adaptée à la réalisation d'un système expert pour l'imputation.

Règle 35

si l'opération est une vente

et le nom du client est mentionné

et l'objet de la vente est le produit fini

et l'escompte est accordé

et le port est facturé

et les emballages sont consignés

et la mention «payée» n'apparaît pas

alors débiter le compte du client mentionné sur la facture pour le montant «net à payer» et débiter le compte 665 pour le montant de l'escompte et créditer le compte de vente du produit fini pour le montant «net commercial» hors-taxes et créditer le compte 4457 pour le montant de la TVA et créditer le compte 4196 pour le montant des emballages et créditer le compte 7085 pour le montant du port

2.2 - Le système KIWI de Clarke

B. Clarke est un précurseur, tant par son système KIWI que par sa démarche. Il est le premier à proposer un système pour l'imputation de phrases décrivant des opérations économiques élémentaires. Son approche est également originale: il rend opérationnelle l'axiomatique comptable de Y. Ijiri [IJIRI 64] grâce à un démonstrateur de théorèmes.

L'axiomatique de Y. Ijiri, bien que concernant la théorie de la mesure comptable, s'applique aussi à l'imputation. Les opérations effectuées par une entreprise y sont décrites en termes d'entités économiques gagnant ou perdant le contrôle de ressources. Par exemple, l'axiome de l'échange dans cette théorie s'écrit en logique des prédicats:

$$C(z,w) \wedge N(w,n) \wedge V(w,x) \wedge C(z,u) \wedge N(u,p) \wedge V(u,v) \Rightarrow J(z,u,v,w,x)$$

où:

- $C(x,y)$ signifie que l'agent économique x contrôle la ressource y .
- $N(x,p)$ (resp. $N(x,n)$) signifie que le contrôle sur la ressource x est positif (resp. négatif).
- $V(x,y)$ signifie que la valeur de la ressource x est y .
- $J(z,u,v,w,x)$ signifie que le compte u débité du montant v et le compte w crédité du montant x est une entrée au journal de l'agent économique z .

B. Clarke étend la théorie de Y. Ijiri, manquant en certains points de rigueur, afin de la rendre utilisable par un démonstrateur de théorèmes. L'imputation revient avec cette méthode à faire la preuve d'un théorème. Au préalable, une analyse linguistique de la phrase décrivant l'opération économique est effectuée pour en fournir le contenu sous forme prédicative au démonstrateur. Pour cela, la syntaxe de la phrase est analysée par des ATN [WOODS 70], suivie d'une analyse sémantique reposant sur une grammaire de cas [FILLMORE 68]. Enfin, utilisant l'étude des verbes commerciaux de [HENDRIX 73], la description de la phrase est mise sous une forme canonique centrée sur la notion d'échange. L'obtention de la forme prédicative est alors immédiate. Malheureusement, l'axiomatique de Y. Ijiri ne permet de modéliser que des opérations très simples dont l'écriture comptable ne comprend que deux mouvements. Elle est donc insuffisante pour imputer la plupart des opérations d'une entreprise.

Les deux systèmes qui viennent d'être étudiés, s'ils n'atteignent pas un niveau de performance leur permettant d'être utilisables en pratique, n'en fournissent pas moins de précieux enseignements. Ils montrent qu'il est vain d'essayer de déduire les écritures directement du contenu des pièces comptables comme le font J. Stepniewski et M. Gervais; un autre niveau de description est nécessaire pour représenter les opérations de l'entreprise, comme dans KIWI. C'est précisément la démarche adoptée pour la réalisation de SICA.

3 - Présentation de SICA

Poursuivant l'idée de B. Clarke, un modèle des opérations économiques est utilisé dans SICA. Il est cependant différent: les opérations économiques sont représentées comme des échanges de flux entre l'entreprise et un tiers. Grâce à ce modèle, SICA peut imputer n'importe quelle opération d'une entreprise.

Ce point constitue la principale différence entre SICA et KIWI car, par ailleurs, SICA suit une voie assez similaire à celle de KIWI. Il analyse des phrases décrivant les opérations d'une entreprise et déduit l'écriture associée (figure 7.3). Un analyseur linguistique reconstitue d'abord l'échange qui est ensuite imputé afin d'obtenir l'écriture.

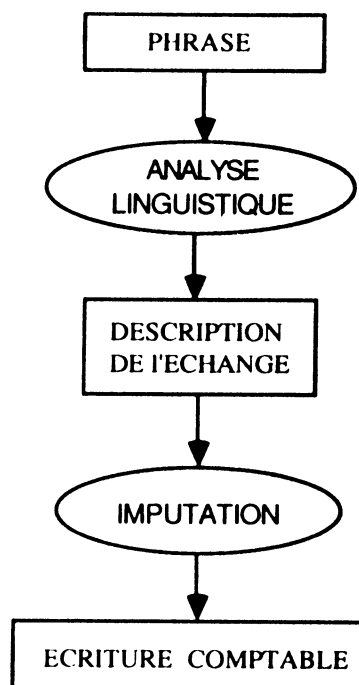


Figure 7.3: le traitement d'une phrase dans SICA

Puisque la description des opérations économiques sous forme d'échanges de flux constitue l'apport principal de SICA, il nous faut la décrire en détail.

3.1 - La description des opérations comme échanges de flux

Les pièces comptables servent de justificatifs des opérations économiques effectuées. Elles mémorisent les flux de diverses natures entrant et sortant de l'entreprise, afin de pouvoir les comptabiliser. Pour cette raison, une opération de l'entreprise est vue dans SICA comme un échange de flux avec un tiers (figure 7.4).

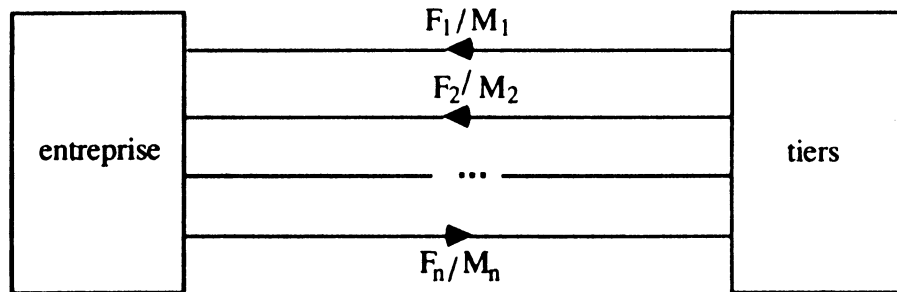


Figure 7.4: la représentation d'une opération économique entre une entreprise et un tiers

Chaque flux possède les trois caractéristiques suivantes:

- nature : chaque F_j spécifie l'objet échangé.
- sens: indiqué par le sens de la flèche entre l'émetteur et le récepteur du flux.
- intensité: M_j est le montant (en francs) du flux F_j .

Le nombre des flux d'une opération peut être quelconque, mais la somme des intensités des flux entrant dans l'entreprise doit être égale à celle des flux sortant. La figure 7.5 représente par exemple l'échange (vu de chez Dugloux) sous-jacent à la facture de la figure 7.2. L'entreprise reçoit un flux de marchandises et un flux de TVA. Le sens de ce dernier flux peut paraître surprenant; pour en comprendre la signification, il faut se souvenir que c'est Dugloux qui paie la TVA au cours de l'achat. Ce paiement correspond donc bien à un "service" accordé par le fournisseur, d'où le sens du flux. En contrepartie de ces deux flux, le flux de nature "dette" matérialise la dette contractée par Dugloux envers son fournisseur au cours de l'opération.

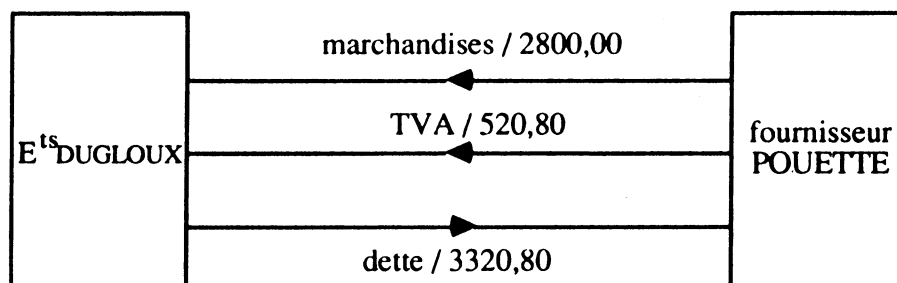


Figure 7.5: représentation de l'échange associé à la facture d'achat de la figure 7.2

Cette représentation permet également de décrire des opérations internes: l'entreprise est alors considérée par un artifice comme le tiers. La figure 7.6 schématisant un virement de 5000 francs de la banque à la caisse, illustre cette possibilité.

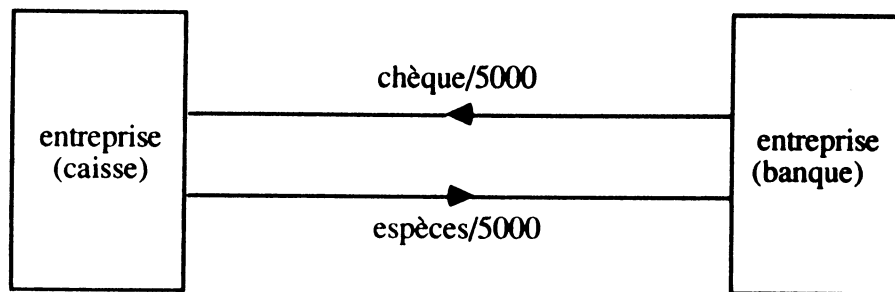


Figure 7.6 : représentation d'un virement de 5000 francs de la caisse à la banque

Comme on le voit, la représentation d'une opération comme un échange de flux est beaucoup plus puissante que celle de B. Clarke. Elle permet de représenter des opérations complexes dont le nombre de flux n'est pas limité. Par ailleurs, la déduction de l'écriture comptable est à partir de cette représentation est très simple.

3.2 - De la description de l'échange à l'écriture comptable

Le passage de la description des flux à l'écriture comptable est immédiat, car à chaque flux est associé un mouvement de l'écriture comptable. Les trois éléments d'un mouvement: numéro de compte, sens et montant sont déterminés de la manière suivante:

- le numéro du compte est identifié de manière unique par le triplet (nature du flux, émetteur, récepteur), comme le montre la table de la figure 7.7. Cette table représente donc le plan comptable de l'entreprise.
- le sens du mouvement est déterminé par celui du flux: c'est un DEBIT si le flux entre dans l'entreprise et un CREDIT sinon.
- le montant du mouvement est l'intensité du flux.

NATURE DU FLUX	EMETTEUR DU FLUX	RECEPTEUR DU FLUX	NUMERO DU COMPTE
marchandises	entreprise	client	707
marchandises	fournisseur	entreprise	607
dette	entreprise	fournisseur	400
tva	fournisseur	entreprise	445
...

Figure 7.7: un extrait du plan comptable de SICA

Les mouvements de l'écriture sont totalement déterminés par les trois points ci-dessus et le plan comptable. La figure 7.8 présente ainsi l'écriture comptable associée à l'opération de la figure 7.5.

Date	Réf.	Comptes		Mouvements	
		n°	nom	débit	crédit
3/6/1988	565	607	achat mse	2800,00	
3/6/1988	565	445	TVA	520,80	
3/6/1988	565	400	fournisseurs		3320,80

Figure 7.8: l'écriture comptable de l'opération de la figure 7.5

Comme on le voit, l'imputation comptable est très simple lorsque les opérations sont mises sous forme d'un échange de flux. L'essentiel du travail de SICA consiste donc à obtenir cette représentation.

3.3 - De la phrase à la description de l'échange

Ce travail est effectué par un analyseur linguistique. Plusieurs raisons nous ont poussé à développer un analyseur pour cette application plutôt qu'en adapter un existant déjà. Tout d'abord, le vocabulaire comptable ne forme qu'un petit sous-ensemble du français. Ensuite, contrairement à la plupart des applications de traitement du langage naturel, la syntaxe des phrases admises dans SICA ne dérive pas d'une étude de corpus: c'est nous qui l'avons imposée de sorte que toute opération puisse être décrite. De plus, les phrases étant restreintes à la description d'une opération de l'entreprise, leur

sémantique est très limitée. Enfin, le langage comptable présente ses propres caractéristiques: la présence de dates, de montants, de sigles dans les phrases en sont des exemples.

Dans SICA, l'analyse linguistique (figure 7.9) comporte quatre niveaux: morphologie, syntaxe, sémantique et logique; nous allons successivement les détailler.

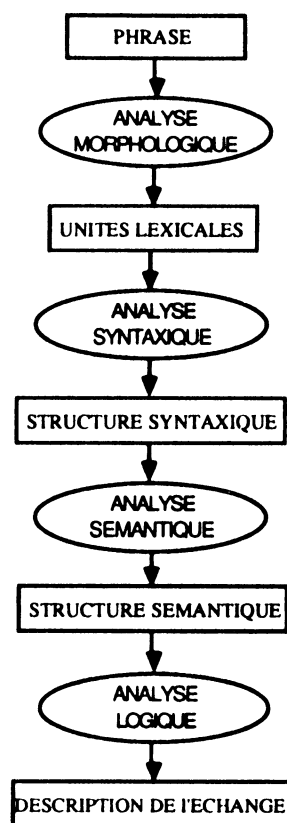


Figure 7.9: l'analyse linguistique dans SICA

3.3.1 - L'analyse morphologique

L'analyseur morphologique découpe en unités lexicales la chaîne de caractères composant la phrase entrée par l'utilisateur. La taille des dictionnaires étant réduite, 500 racines environ, il n'est pas nécessaire d'utiliser de techniques d'accès et d'analyse sophistiquées.

3.1.2 - L'analyse syntaxique

Le choix d'une grammaire pour la description des opérations d'une entreprise est guidé par deux contraintes. Il doit d'une part permettre à toute opération d'être décrit. Il est d'autre part souhaitable d'éviter les problèmes du traitement du langage naturel appliqué à des textes généraux. Ces considérations nous ont amené à fixer pour les phrases la syntaxe suivante:

DATE [SUJET] VERBE CD* CI* (MONTANT)*

où:

{ } indique un élément dont la place n'est pas définie dans la phrase

[] indique un élément optionnel

* indique un nombre quelconque (≥ 0) d'occurrences

Une phrase commence par la date de l'opération, suivie du verbe. Une phrase ne peut comporter qu'un verbe principal (les compléments peuvent cependant contenir des participes passés); ceci facilite le rattachement des compléments. Un verbe peut se présenter sous forme d'un nom verbal comme dans (P1), d'un participe passé comme dans (P2), ou conjuguée comme dans (P3); dans ce dernier cas, il est précédé d'un sujet. Une caractéristique importante du langage comptable est que du point de vue de l'imputation, on peut ramener les compléments suivant le verbe à deux types seulement: direct (CD) ou indirect (CI); nous verrons les raisons de cette particularité. Les premiers occupent la fonction de complément d'objet direct (pour un verbe transitif); les seconds regroupent tous les autres compléments. Ainsi, (P1) comporte un CD: "d'un chèque", et un CI: "au fournisseur". Les compléments peuvent être de simples groupes nominaux comme le CD "d'un chèque" dans (P1) ou complexes comme "pour les marchandises achetées le 29/6" dans (P2). Enfin, la phrase comporte un ou plusieurs montants se rapportant aux compléments; leur place n'est pas déterminée: ils peuvent précéder le complément auquel ils se rapportent comme dans (P3) ou le suivre comme dans (P1) et (P2).

(P1) le 12/7/88: envoi d'un chèque au fournisseur de 5000 F

(P2) le 13/7, reçu la facture n°7565 pour les marchandises achetées le 29/6: 6000 F HT

(P3) 18/7: le client Fourroux retourne 760 F HT de marchandises détériorées

L'analyse syntaxique construit la structure de la phrase suivant cette syntaxe. Pour la phrase (P1), cette structure est présentée sur la figure 7.10.

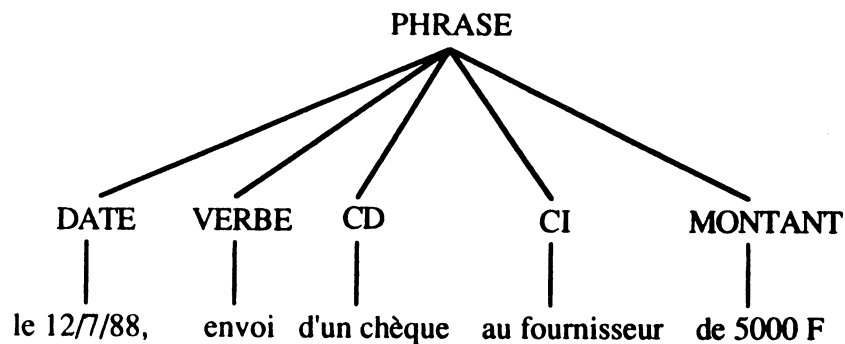


Figure 7.10: structure syntaxique de (P1)

Une particularité de SICA est d'effectuer les étapes suivant l'analyse syntaxique par des règles logiques. Aussi, une fois la structure syntaxique de la phrase obtenue, celle-ci est transformée en une proposition exprimée en logique des prédicats. Pour cela, à chacun des éléments de la phrase est associé un prédicat. La description syntaxique de la phrase est donc assurée au moyen des six prédicats suivants:

date(d) : d est la date de l'opération

verbe(v) : v est le verbe ou le nom verbal ayant valeur prédicative

sujet(x) : x est le sujet du verbe

cd(x) : x est un complément d'objet direct du verbe

ci(p,x) : x est un complément indirect introduit par la préposition p

montant(x, m) : m est le montant du complément x

Les arguments de ces prédicats sont des constantes représentant les éléments de la phrase. Ainsi, aux expressions "chèque bancaire", "chèque",... sont associées la constante CHEQUE. La description syntaxique de la phrase (P1) qui illustrera tout au long de ce chapitre le fonctionnement de SICA s'exprime donc:

date(12/7/88) \wedge verbe(ENVOYER) \wedge cd(CHEQUE)
 \wedge ci(A, FOURNISSEUR) \wedge montant(CHEQUE, 5000)

3.3.3 - L'analyse sémantique

Une fois la syntaxe de la phrase analysée, sa "signification" est recherchée. Ceci revient à déterminer pour chaque élément de la description syntaxique, la partie de l'échange de flux qu'il représente. Cette tâche est effectuée au cours de l'analyse sémantique, qui repose sur une **grammaire de cas**.

Introduites en linguistique générale par C. Fillmore [FILLMORE 68], les grammaires de cas définissent autour du verbe, élément central de la phrase, un ensemble de rôles sémantiques. Au nombre de six dans le système casuel initial de C. Fillmore, ces cas sont:

- l'agentif: l'instigateur (une entité animée) de l'action décrite par le verbe.
- l'instrumental: l'objet impliqué dans l'action.
- le datif: l'entité animée affectée par l'action.
- le factitif: l'objet résultant de l'action.
- le locatif: le lieu ou l'orientation spatiale de l'action.
- l'objectif: l'objet affecté par l'action.

Ce système casuel est illustré par l'exemple de la phrase (P4):

(P4) Jean (AGENTIF) ouvre la porte (OBJECTIF) avec un couteau (INSTRUMENTAL)

L'univers sémantique de l'imputation comptable est beaucoup plus restreint que celui du langage général, car il est limité à la description d'échanges de flux. Aussi, est-il possible de le décrire en utilisant seulement quatre rôles sémantiques, qui sont:

- l'AGENT: l'instigateur de l'opération économique décrite par le verbe.
- le CONTRE-AGENT: le partenaire de l'AGENT dans l'opération.
- l'OBJET: l'objet sur lequel s'applique l'opération.
- le CONTRE-OBJET: la contrepartie de l'OBJET au cours de l'opération.

Les deux premiers cas décrivent les partenaires de l'échange, tandis que les deux derniers indiquent les flux échangés. Ce système casuel est illustré sur la phrase (P5) ci-dessous.

(P5) le 12 juin 1988, nous (AGENT) achetons des marchandises (OBJET) au fournisseur (CONTRE-AGENT) par chèque bancaire (CONTRE-OBJET): 3500 F TTC

Il faut remarquer que tous les cas ne sont pas nécessairement présents dans une phrase. (P1) par exemple, ne comporte ni AGENT, ni CONTRE-OBJET. Par ailleurs, si l'AGENT et le CONTRE-AGENT sont uniques dans la phrase, il peut en revanche exister plusieurs OBJETs et CONTRE-OBJETs, comme dans la phrase (P6)

(P6) le 5/4: règlement de marchandises (OBJET) au fournisseur (CONTRE-AGENT) : 5000 F par chèque (CONTRE-OBJET) et 3000 F en espèces (CONTRE-OBJET).

Le but de l'analyse sémantique est de déterminer le cas de chaque complément ainsi que du sujet de la phrase, lorsqu'il existe. Deux éléments sont déterminants pour effectuer cette opération: les groupes nominaux et les prépositions.

Les groupes nominaux du vocabulaire de l'imputation comptable se décomposent essentiellement en deux classes: les acteurs et les objets d'échange (figure 7.11).

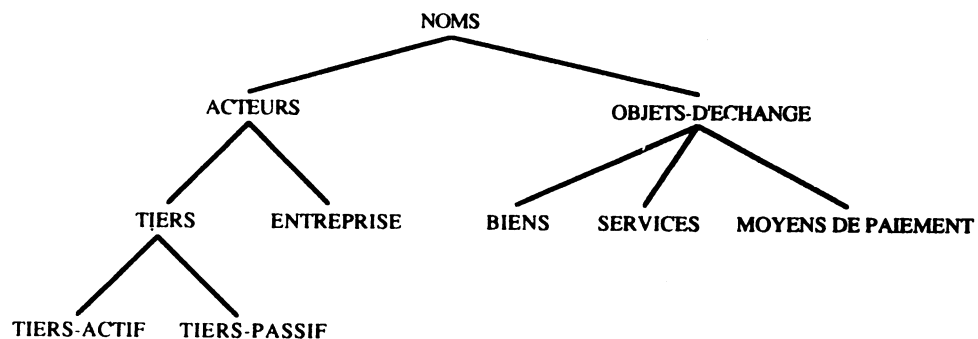


Figure 7.11: le vocabulaire comptable

Les acteurs sont les entités économiques pouvant apparaître dans une phrase. Ils sont de deux types:

- des tiers: ce sont les partenaires de l'entreprise au cours de l'échange. Ils se subdivisent eux-mêmes en tiers de l'actif et tiers du passif. Les premiers sont, en situation normale, des débiteurs de l'entreprise: clients, locataires,... Les seconds sont des créiteurs de l'entreprise: fournisseurs, Sécurité Sociale, ...

- l'entreprise: étant toujours partie prenante dans les échanges, elle fait également partie de la classe des acteurs.

Les objets d'échange désignent tout ce qui peut être échangé au cours d'une opération économique. Ils se subdivisent en trois classes:

- les biens représentent des entités matérielles pouvant être échangées: marchandise, matériel, matière première, ...
- les moyens de paiement: espèce, chèque, effet de commerce,...
- les services représentent des entités d'échange ni matérielles, ni monétaires : salaire, loyer, ...

La classification des groupes nominaux a une fonction importante au cours de l'analyse sémantique, car elle a un rôle équivalent aux traits sémantiques dans les grammaires casuelles générales. Les rôles d'AGENT et de CONTRE-AGENT ne peuvent en effet être attribués qu'à des éléments de la classe des acteurs tandis qu'OBJET et CONTRE-OBJET sont réservés aux objets-d'échange.

Les prépositions sont également importantes pour l'analyse sémantique. Elles fonctionnent en français comme marqueurs de cas. Ceci se vérifie aussi pour l'imputation comptable, mais comme pour les groupes nominaux, leur classification est plus simple que dans le langage général. Trois types de prépositions sont considérés dans SICA:

- les préposition-origine (par, ...) suivies d'un groupe nominal de type acteur introduisent le sujet logique de la phrase et donc l'AGENT de l'opération.
- les préposition-destination (à, au bénéfice de, ...) suivies d'un groupe nominal de type acteur précèdent le partenaire de l'AGENT, donc le CONTRE-AGENT.
- les préposition-instrument (par, en, ...) suivies d'un objet-d'échange introduisent le CONTRE-OBJET.

Il faut ici noter l'ambiguïté de la préposition "de". Considérons les deux phrases (P7) et (P8), dont la structure de surface est identique. Dans la phrase (P7), "de" est une préposition-origine: c'est le client

qui paie, car c'est un tiers de l'actif; c'est donc l'AGENT. Dans la phrase (P8) au contraire, elle est préposition-destination: c'est l'entreprise qui paie le fournisseur, le CONTRE-AGENT. Dans certains cas comme celui des phrases (P7) et (P8), l'ambiguïté peut être levée grâce à la distinction entre tiers d'actif et tiers de passif, mais ce n'est pas toujours le cas; la phrase est alors rejetée.

(P7) le 12 mai, paiement du client par chèque des marchandises achetées le 5 mai: 5000 F

(P8) le 13 mai, paiement du fournisseur par chèque des marchandises achetées le 5 mai: 5000 F

L'attribution des rôles sémantiques est effectuée à partir des classifications des groupes nominaux et des prépositions de la manière suivante:

- l'AGENT est le sujet logique de la phrase (le SUJET ou un CI introduit par une préposition-origine).
- le CONTRE-AGENT est un CD de type acteur ou un CI de même type introduit par une préposition-destination.
- l'OBJET est un CD de type objet-d'échange.
- le CONTRE-OBJET est un CI de type objet-d'échange introduit par une préposition-instrument.

Ces quatre points sont représentés dans SICA par des règles logiques. Pour cela, à chacune des classes de groupes nominaux et de prépositions définies est associée un prédicat. De même, un prédicat est défini pour chaque cas de la grammaire casuelle. Les instances des prédicats de chacune des classes élémentaires (tiers-actif, tiers-passif, bien, moyen-de-paiement, service, préposition-origine, préposition-destination, préposition-instrument) constituent le dictionnaire sémantique de SICA (cf. annexe C.2). La structuration de ces classes est représentée par les règles RU₁ à RU₃ de l'annexe C.1. Les règles RS₁ à RS₅ formalisent les quatre points présentés ci-dessus pour l'attribution des rôles sémantiques.

Illustrons le fonctionnement de l'analyse sémantique en appliquant ces règles sur la phrase (P1), dont on rappelle que la description syntaxique est:

date(12/7/88) \wedge verbe(ENVOYER) \wedge cd(CHEQUE)
 \wedge ci(A, FOURNISSEUR) \wedge montant(CHEQUE, 5000)

Sa description sémantique est obtenue de la manière suivante:

FAIT INFERE	REGLE UTILISEE
cd(CHEQUE)	description syntaxique
mp(CHEQUE)	dictionnaire sémantique
objet-d'échange(CHEQUE)	RU3
objet(CHEQUE)	RS4
ci(A, FOURNISSEUR)	description syntaxique
prep-dest(A)	dictionnaire sémantique
tiers-passif(FOURNISSEUR)	dictionnaire sémantique
tiers(FOURNISSEUR)	RU1
acteur(FOURNISSEUR)	RU2
contre-agent(FOURNISSEUR)	RS3

d'où la description sémantique de (P1):

date(12/7/88) \wedge verbe(ENVOYER) \wedge objet(CHEQUE)
 \wedge contre-agent(FOURNISSEUR) \wedge montant(CHEQUE, 5000)

3.3.4 - L'analyse logique

Au cours de cette dernière étape de l'analyse linguistique, l'échange est totalement reconstitué. Pour cela, les éléments restés implicites dans la phrase sont déduits: AGENT et CONTRE-AGENT sont inférés s'ils sont absents de la description sémantique; de même les flux absents de la phrase sont déduits, ainsi que les caractéristiques de chacun d'eux: nature, sens et intensité.

3.3.4.1 - Déduction des partenaires de l'échange

Lorsque l'un et/ou l'autre des partenaires (AGENT ou CONTRE-AGENT) de l'échange est absent de la description sémantique, il existe plusieurs moyens de le déduire. Tout d'abord, un échange concerne toujours l'entreprise; cela signifie que si l'un des partenaires est connu et appartient à la classe des tiers, l'autre est nécessairement l'entreprise. C'est cet argument qui permet de déterminer l'AGENT

de l'opération dans la phrase (P1); celle-ci ne comporte pas d'AGENT et le CONTRE-AGENT (fournisseur) est un tiers: l'AGENT est donc l'entreprise. Il est également possible dans certains cas d'inférer l'un des partenaires en analysant les flux échangés. Il existe en effet des tiers spécialisés dans la réception ou l'envoi de flux de nature particulière. Dans la phrase (P9), le CONTRE-AGENT n'est pas précisé mais on peut déduire qu'il s'agit de l'EDF, car c'est le seul tiers susceptible d'envoyer un flux de nature électricité à l'entreprise.

(P9) le 13, nous payons la quittance d'électricité par chèque: 1200 F

Lorsque l'AGENT ou le CONTRE-AGENT ne peut être déduit, la phrase est ambiguë; elle n'est donc pas imputée.

3.3.4.2 - Déduction des flux échangés

Chaque OBJET et CONTRE-OBJET de la phrase représente un flux. Cependant, il arrive que certains flux soient absents de la phrase; il faut dans ce cas les déduire. Pour cela, on attribue aux flux une propriété particulière appelée **induction**. Cette propriété représente le fait que certains flux ne peuvent exister isolément: un flux de marchandises sujettes à TVA par exemple s'accompagne toujours d'un flux de TVA. Ce mécanisme permet de déduire l'existence du flux induit lorsque le flux inducteur est présent dans l'échange. On appelle coefficient d'induction le rapport:

$$\frac{\text{intensité du flux induit}}{\text{intensité du flux inducteur}}$$

Ce coefficient est positif si le flux induit est de même sens que le flux inducteur et négatif sinon. Pour les marchandises taxées à 18,6%, le flux de marchandises induit donc un flux de TVA avec un coefficient d'induction égal à +0,186.

Le mécanisme d'induction présente un cas particulier important. Il arrive fréquemment qu'une phrase décrivant une opération ne comporte qu'un flux (ou plusieurs flux de même sens). Or, il existe nécessairement un flux de sens inverse pour que l'échange soit équilibré. Ce dernier correspond en fait à la création ou l'extinction d'une dette, implicite dans la phrase. Ceci revient à poser que tout flux induit par défaut (c'est-à-dire en l'absence d'un flux de sens contraire) un flux de dette avec un coefficient d'induction égal à -1. C'est par ce moyen que le CONTRE-OBJET de la phrase (P1) est déterminé: le flux de chèque n'ayant pas de contrepartie, un flux de dette de sens opposé et d'intensité égale à celui de chèque est ajouté à la description de l'échange.

3.3.4.3 - Caractérisation des flux

Déterminer la totalité des flux de l'échange n'est pas suffisant; chacun d'eux doit être totalement explicite pour que l'imputation soit possible. Il faut donc en rechercher la nature, le sens et l'intensité.

La nature d'un flux est simplement déterminée par l'OBJET ou le CONTRE-OBJET associé. Dans (P1) le flux associé à l'OBJET chèque est donc de nature chèque.

Le sens des flux dépend du verbe de la phrase. Cependant, il est inutile de décrire la sémantique de chaque verbe comme le fait B. Clarke dans son système KIWI. Il est en effet possible de classer tous les verbes du vocabulaire comptable en deux catégories dénommées A et V. Les verbes de type A (comme acheter) sont tels que l'OBJET est reçu par l'AGENT et le CONTRE-OBJET donné. Des verbes de ce type sont: acheter, payer, recevoir, etc. Les verbes de type V (comme vendre) ont un comportement inverse: l'OBJET est donné par l'AGENT et le CONTRE-OBJET reçu; vendre, envoyer, ..., sont des verbes de ce type. Ce point est très important car il simplifie considérablement la construction des dictionnaires et les traitements à effectuer. Dans la phrase (P1), le verbe "envoyer" étant de type V, l'OBJET chèque est donné par l'AGENT, le flux associé circule donc de l'entreprise (l'AGENT) vers le fournisseur (le CONTRE-AGENT). Le flux de dette étant induit par ce flux avec un coefficient négatif, il circule en sens inverse.

L'intensité d'un flux est caractérisée par le montant de l'OBJET ou du CONTRE-OBJET associé (certains flux nécessitent cependant des traitements particuliers; par exemple, dans une facture d'achat de marchandises comportant une remise, le montant de cette remise est soustrait de l'intensité du flux de marchandises). Lorsque le montant est absent de la phrase, deux règles permettent de déterminer l'intensité d'un flux. La première est liée au mécanisme d'induction. Si un flux induit un autre flux avec un coefficient d'induction C, et que l'intensité du flux induit est inconnue, elle peut être calculée de la manière suivante: $I_{\text{induit}} = |C| \cdot I_{\text{inducteur}}$. C'est ainsi qu'est déterminée l'intensité du flux de dettes dans la phrase (P1): $I_{\text{dette}} = |-1| \cdot I_{\text{chèque}} = 5000$. La deuxième règle est liée à ce qu'on pourrait appeler la loi de conservation des flux; un échange est tel que pour les deux partenaires, on a l'égalité: $\sum I_{\text{flux-entrant}} = \sum I_{\text{flux-sortant}}$. Si tous les flux de l'échange sont connus excepté un, son intensité est telle que la contrainte ci-dessus est respectée. Lorsque les intensités de plusieurs flux sont inconnues, la phrase n'est pas imputable; elle est rejetée.

L'analyse logique de la phrase est réalisée, comme l'analyse sémantique, au moyen de règles logiques. Les règles RL₁ à RL₄ de l'annexe C.1 permettent de déduire les partenaires manquants de l'échange. Les règles RL₅ à RL₈ déterminent le sens des flux. Enfin, les règles RL₉ et RL₁₀ construisent les flux de l'échange. Remarquons cependant que la règle de conservation de flux ne peut être exprimée en logique; elle fait l'objet d'un traitement indépendant.

Illustrons l'utilisation de ces règles sur la phrase (P1) dont nous rappelons que la description sémantique est:

date(12/7/88) \wedge verbe(ENVOYER) \wedge objet(CHEQUE)
 \wedge contre-agent(FOURNISSEUR) \wedge montant(CHEQUE, 5000)

Le passage à la description de l'échange s'effectue de la manière suivante:

FAIT INFERE	REGLE UTILISEE
verbe(ENVOYER)	description syntaxique
tv(ENVOYER,V)	dictionnaire sémantique
objet(CHEQUE)	description sémantique
mp(CHEQUE)	dictionnaire sémantique
sens(CHEQUE,DON)	RL7
contre-agent(FOURNISSEUR)	description sémantique
tiers-passif(FOURNISSEUR)	dictionnaire sémantique
tiers(FOURNISSEUR)	RU1
agent(ENTREPRISE)	RL2
montant(CHEQUE,5000)	description syntaxique
flux(CHEQUE,ENTREPRISE,FOURNISSEUR,5000)	RL10
flux(DETTE,FOURNISSEUR,ENTREPRISE,5000)	flux induit

d'où la description de l'échange associée à (P1):

flux(CHEQUE, ENTREPRISE, FOURNISSEUR, 5000)
 \wedge flux(DETTE, FOURNISSEUR, ENTREPRISE, 5000)

Nous pouvons maintenant revenir sur l'imputation de l'échange pour obtenir l'écriture comptable associée à la phrase (P1). La déduction de l'écriture s'effectue par les deux règles RI₁ et RI₂ (cf. annexe C.1) de la manière suivante:

FAIT INFERE	REGLE UTILISEE
flux(CHEQUE,ENTREPRISE,FOURNISSEUR,5000)	description échange
compte(CHEQUE,ENTREPRISE,FOURNISSEUR,510)	plan comptable
mouvement(510,CREDIT,5000)	RI2
flux(DETTE,FOURNISSEUR,ENTREPRISE,5000)	description échange
compte(DETTE,FOURNISSEUR,ENTREPRISE,400)	plan comptable
mouvement(400,DEBIT,5000)	RI1

d'où l'écriture associée à (P1):

mouvement(400,DEBIT,5000) A mouvement(510,CREDIT,5000)

Voici donc présentées toutes les étapes menant d'une phrase à l'écriture comptable qui lui est associée. Cette présentation est toutefois légèrement simplifiée car elle ne présente que le noyau du système réalisé. En particulier, il est nécessaire d'introduire différents niveaux de contrôles au cours de l'imputation pour des raisons de fiabilité. Ainsi, lorsque l'utilisateur signale par une phrase le paiement d'une facture, il faut vérifier que cette facture a été au préalable effectivement reçue, que les flux et les montants correspondent, etc. Toutes ces vérifications sont effectuées par des traitements particuliers, en dehors des règles logiques présentées.

3.4 - Résultats et perspectives

SICA est totalement écrit en PROLOG. Il comporte environ 2000 lignes de code (hormis les prédicats utilisés pour les dictionnaires), dont 1/5 seulement pour les règles logiques. Le temps de réponse du système est inférieur à 4 secondes sur un PC-AT pour traiter des phrases comme celle de la

figure 7.12. Il traite la plupart des opérations courantes; les phrases n'ayant pu être imputées sont stockées sur un fichier et doivent être traitées manuellement par un comptable.

SAISIE DES OPERATIONS

Phrase

le 12 juin 1988, reçu la facture n° 3432 du fournisseur Fernand:
 marchandises type 1: 1200,00 F HT
 remise: 100,00 F
 TVA: 200,88 F

Imputation

J'ai passé l'écriture suivante:

12/06/88	facture 3432	débit	crédit
607	achat de marchandises	1100,00	
445	TVA	200,88	
	400 fournisseurs		1300,88

Figure 7.12: exemple d'imputation d'une phrase par SICA

Sur le plan de l'application, la démarche utilisée dans SICA est tout à fait concluante. L'installation de SICA en entreprise pose néanmoins des problèmes. Au niveau de la sécurité d'abord, la tenue des comptes ne tolère pas (ou très peu) les erreurs. On comprend donc les craintes suscitées par la mise en place d'un système comme SICA. Le système fait actuellement preuve de fiabilité pour l'imputation de phrases correctes. Mais la possibilité d'interprétation inattendue d'une phrase mal formulée est inquiétante. Au niveau de l'ergonomie enfin, la description des pièces comptables par des phrases s'avère à l'usage extrêmement lourde. Le nombre d'écritures passées quotidiennement dans une PME est généralement de l'ordre de quelques dizaines, voire quelques centaines. De plus, elles concernent pour la plupart deux types d'opérations: les achats et les ventes. Il est donc très pénible de devoir taper pour chacune d'elles une phrase. Pour cette raison, une autre solution semble devoir être adoptée pour la saisie. La proposition à l'utilisateur de "squelettes" de phrases qu'il compléterait par exemple, semble une solution envisageable; elle permettrait par ailleurs d'améliorer la fiabilité du système.

Sur le plan méthodologique, la réalisation de SICA montre une nouvelle fois l'intérêt d'exploiter les connaissances profondes acquises dans un domaine, lorsqu'elles existent. Cela autorise une

représentation très concise des connaissances: le noyau de SICA ne comprend que 20 règles logiques; plusieurs centaines de règles seraient nécessaires à un système expert utilisant les connaissances de surface du comptable pour résoudre le même problème (cf. paragraphe 2.1). Corollairement, le temps de développement se trouve raccourci: moins de quatre mois pour la conception du noyau de SICA (ce qui représente environ un tiers du temps total). L'utilisation des connaissances profondes procure également un gain en flexibilité du système: à cause des différences entre les plans comptables adoptés par les entreprises, une partie des règles d'un système expert de première génération devrait être réécrite pour porter l'application d'une entreprise à une autre. Dans SICA, il suffit de modifier la table qui représente le plan comptable (cf. paragraphe 3.2). Enfin, la réalisation de SICA a mis en évidence un point important: lorsque la formulation initiale d'un problème n'est pas modélisable, il peut être très profitable de changer le niveau de description du problème afin de pouvoir tirer profit des connaissances profondes du domaine. C'est ainsi le passage de la description des pièces comptables à celui des opérations économiques sous forme d'échanges de flux qui a permis la réalisation de SICA.

Notons pour conclure ce chapitre que SICA ouvre la voie à la réalisation d'un outil puissant d'EAO en comptabilité. Une partie de la formation des comptables s'effectue en effet par des exercices d'imputation de phrases qui décrivent les opérations d'une entreprise. Nous avons pu constater qu'il est très facile de doter SICA d'un module d'explication des écritures générées; ces explications sont d'autant mieux adaptées aux étudiants que le modèle d'échange de flux est très répandu dans les manuels de comptabilité. Actuellement, l'analyseur de SICA n'est capable de traiter qu'une faible proportion des exercices présentés dans les livres d'exercices. La construction d'un analyseur linguistique à cette fin constitue donc un défi très intéressant à la fois pour le traitement du langage naturel et la comptabilité.

CONCLUSION

Résumons tout d'abord les lignes essentielles de cette thèse. Les lacunes des systèmes experts de première génération ont d'abord été analysées: difficulté de construction et de maintenance d'une base de connaissances, faible réutilisabilité de cette dernière pour résoudre un problème légèrement différent du problème initial, insuffisance de l'explication du raisonnement. Une alternative reposant sur l'exploitation des connaissances théoriques acquises dans le domaine d'application a été exposée. L'utilisation de cette dernière approche dans le projet PIC a conduit au développement de la notion d'interprétation comparative avec pour objectif l'analyse et l'explication des différences entre deux situations grâce à un modèle du fonctionnement du système étudié. Des techniques applicables à des modèles qualitatifs et numériques ont été proposées et implantées dans CRESUS, un générateur de systèmes d'interprétation comparative. Grâce à CRESUS, deux applications du projet PIC ont été réalisées: ANAFI, pour le diagnostic financier et SURPLUS, pour l'analyse de la performance d'une entreprise par la méthode des surplus. Enfin, SICA un système expert pour l'imputation comptable utilisant également des connaissances profondes a été présenté. Diverses conclusions peuvent être tirées de ce travail tant sur le plan des applications que méthodologique.

Au niveau des applications, cette thèse montre que la réalisation d'un poste de travail "intelligent" pour le comptable est possible avec les techniques actuelles de l'Intelligence Artificielle. En amont de la tenue des comptes, SICA est en effet suffisamment près de la réalité pour qu'une application opérationnelle soit envisageable. En aval, les résultats de SURPLUS laissent également espérer un stade opérationnel. Enfin, si ANAFI n'atteint pas ce niveau de performances, il convient de rappeler qu'il a été réalisé sans le concours d'un expert, indispensable pour construire un modèle réaliste du fonctionnement qualitatif d'une entreprise. Outre les deux derniers systèmes, c'est toute une classe d'applications qui s'ouvre aux systèmes experts en Gestion avec l'interprétation comparative. En contrôle de Gestion notamment, l'analyse des coûts et le contrôle budgétaire semblent pouvoir bénéficier de cette démarche.

Au niveau méthodologique, trois conclusions peuvent être formulées. La première concerne l'interprétation comparative. CRESUS pose un premier jalon vers la résolution de cette classe d'applications. Néanmoins, ce système ne doit pas masquer la difficulté du chemin restant à parcourir. D'une part, des hypothèses simplificatrices ont été introduites: comparaison de deux états seulement, limitation à des modèles qualitatifs et numériques simples. La relaxation de chacune de ces hypothèses constitue autant de prolongements intéressants des travaux entrepris. D'autre part, certains problèmes n'ont pas reçu de solution satisfaisante, en particulier la validation de la notion de responsabilité pour mesurer l'impact d'une variable explicative sur la variable expliquée, dans un système d'équations. Le second enseignement pouvant être dégagé provient du système SICA. Il met en évidence l'intérêt de trouver une formulation d'un problème capable d'exploiter les connaissances profondes du domaine d'application, même si cette opération requiert un travail supplémentaire. Le passage de la description des pièces comptables à celle des opérations économiques sous-jacentes introduit ainsi une étape supplémentaire dans la déduction des écritures. Cet effort est cependant largement récompensé par la simplification de l'imputation qui en résulte. Enfin, le troisième enseignement, que nous allons développer, concerne la comparaison des deux approches de conception de systèmes experts.

Tout d'abord, l'utilisation de connaissances profondes autorise l'écriture de bases de connaissances extrêmement concises. Ce point, déjà noté dans [KOTON 85], est remarquable dans ANAFI et plus encore dans SICA, dont le noyau ne comporte qu'une vingtaine de règles logiques. Il se comprend aisément dans la mesure où, pour ces deux problèmes, la représentation des connaissances de surface de l'expert conduit à l'écriture d'une base analysant tous les cas pouvant se présenter. Le problème traité devient ainsi combinatoire et donc la base de connaissances de grande taille. Par ailleurs, l'utilisation de connaissances profondes facilite notablement l'acquisition des connaissances: les équations comptables utilisées dans ANAFI et les différents flux représentés dans SICA sont accessibles dans des ouvrages et non seulement par interview d'un expert. Il ne faut pour autant conclure que le concours d'un expert est inutile pour réaliser un système expert à base de connaissances profondes. Celui-ci est indispensable, ne serait-ce que pour introduire le concepteur du système expert au domaine d'application. Il est également indispensable lorsque des connaissances profondes existent dans le domaine d'application, mais ne sont pas systématiquement regroupées dans un ouvrage. C'est par exemple le cas pour la modélisation qualitative du fonctionnement d'une entreprise dans ANAFI. C'est également le cas de la mise sous forme d'échange de flux de certaines opérations comptables complexes.

Le problème de la maintenance de la base de connaissances dans les systèmes experts de première génération se pose aussi avec une acuité moindre lorsque des connaissances théoriques sont utilisées. Ces dernières ne sont pas sujettes à erreur comme peuvent l'être les connaissances de surface d'un expert. Ceci est d'autant plus vrai que la concision des bases de connaissances profondes diminue encore le risque d'erreurs.

Les connaissances profondes se caractérisent également par leur réutilisabilité pour résoudre un problème différent du problème initial. Cette caractéristique est évidente dans SURPLUS et ANAFI, puisque l'explication comparative utilise directement les équations qui permettent le calcul des soldes intermédiaires de gestion, donc celles-là même qui seraient utilisées pour une simulation qualitative, par exemple. Ceci contraste singulièrement avec les différents jeux de règles nécessaires pour effectuer ces raisonnements (cf. chapitre 1). Cette caractéristique s'illustre également dans SICA par la faible quantité de modifications nécessaires pour adapter le système d'une entreprise à une autre (le vocabulaire et le plan comptable de l'entreprise, donc les données de l'application), opération qui nécessite la réécriture d'une partie de la base de connaissances avec un système expert de première génération.

Avec de tels avantages, on peut se demander pourquoi la plupart des systèmes experts actuels continuent à utiliser des connaissances de surface. La réponse à cette question est très simple: peu de domaines ont atteint un stade de formalisation suffisant pour rendre possible l'utilisation de connaissances profondes. Les applications abordées dans cette thèse constituent un terrain favorable pour l'utilisation de telles connaissances car elles appartiennent à des domaines d'applications où existent d'importants efforts de formalisation. De plus, si l'utilisation de connaissances profondes est suffisante pour deux des applications traitées (SICA et SURPLUS), il faut bien se garder de généraliser ce point. La plupart des problèmes requièrent la combinaison d'un savoir-faire et de connaissances profondes, ainsi que l'illustre ANAFI. Les deux approches de conception de systèmes experts ne sont donc pas concurrentes, mais complémentaires. Aussi, la prochaine étape consistera à les faire coopérer à l'intérieur d'un même système, lorsque les recherches sur les systèmes experts à base de connaissances profondes seront suffisamment avancées. C'est là toute la problématique, entrevue dans ANAFI, des systèmes experts de seconde génération.

BIBLIOGRAPHIE

- [ALTY 87] J. Alty: *the limitations of rule-based expert systems*. In Jiri Kriz ed.: *knowledge-based expert systems in industry*, Ellis Horwood, Chichester, 1987, pp. 17-22.
- [APTE 86] C. Apté, S. Hong: *using qualitative reasoning to understand financial arithmetic*, AAAI, 1986, pp. 942-948.
- [AYEL 86] M. Ayel: *l'incohérence dans une base de connaissances*. Rapport interne, Université de Chambéry, 1986.
- [BERGE 83] C. Berge: *graphes*, Gauthier-Villard, Paris, 1983.
- [BLUM 83] R. Blum: *representation of empirically derived causal relationships*, IJCAI, 1983, pp. 268-271.
- [BOBROW 84] D. Bobrow: *qualitative reasoning about physical systems*, North-Holland, 1984. (également publié sous: *Artificial Intelligence*, vol. 24, 1984).
- [BOUWMAN 83] M. Bouwman: *human diagnostic reasoning by computer: an illustration from financial analysis*. *Management Science*, vol. 29, n° 6, 1983, pp. 653-672.
- [CALOUD 87] P. Caloud: *towards continuous process supervision*, IJCAI, 1987, pp. 1086-1089.
- [CALOUD 88] P. Caloud: *raisonnement qualitatif - application à l'aide à la supervision des procédés continus*. Thèse de Doctorat de l'Institut National Polytechnique de Grenoble, 1988.
- [CHANDRASEKARAN 83] B. Chandrasekaran: *Towards a taxonomy of problem solving types*. *AI magazine*, vol. 4, pp. 9-17, 1983.

- [CLANCEY 83] W. Clancey: *The epistemology of a rule-based expert system - a framework for explanation*. Artificial Intelligence 20, 1983, pp. 215-251.
- [CLANCEY 85] W. Clancey: *heuristic classification*. Artificial Intelligence 27, 1985, pp. 289-350.
- [CLANCEY 87a] W. Clancey, T. Thompson: *applying a qualitative modeling shell to process diagnosis: the Caster system*. Report STAN-CS-87-1169, Stanford University, 1987.
- [CLANCEY 87b] W. Clancey: *the science and engineering of qualitative models*. Report STAN-CS-87-1170, Stanford University, 1987.
- [CLARKE 75] B.J. Clarke: *an intelligent management information system as a tool for formal accounting theory*. Ph.D. Thesis, University of British Columbia, 1975.
- [CLOCKSIN 86] W. Clocksin, A. Morgan: *qualitative control*, ECAI, 1986, pp. 350-356.
- [D'AMBROSIO 87] B. D'Ambrosio: *extending the mathematics in qualitative process theory*, AAI, 1987, pp. 595-599.
- [DAGUE 86] P. Dague, P. Devès, O. Raiman: *raisonnement qualitatif dans le diagnostic de pannes*, 5^e journées sur les systèmes experts et leurs applications, Avignon, 1986, pp. 613-632.
- [DAVIS 82] R. Davis: *expert systems: where are we ? and where do we go from here ?* MIT Artificial Intelligence memo. 665, 1982.
- [DAVIS 84] R. Davis: *Diagnostic reasoning based on structure and behavior*. Artificial Intelligence 24, 1984, pp. 347-410.
- [DE KLEER 84a] J. De Kleer, J.S. Brown: *qualitative physics based on confluences*. Artificial Intelligence 24, 1984, pp. 7-83.
- [DE KLEER 84b] J. De Kleer, D. Bobrow: *qualitative reasoning with higher-order derivatives*, AAI, 1984, pp. 86-91.

- [DE KLEER 86] J. De Kleer, J. S. Brown: *theories of causal ordering*. Artificial Intelligence 29, 1986, pp. 33-61.
- [DE KLEER 87] J. De Kleer, B. Williams: *diagnosing multiple faults*. Artificial Intelligence, vol. 32, 1987, pp. 97-130.
- [FEIGENBAUM 77] E. Feigenbaum: *the art of Artificial Intelligence: themes and case studies of knowledge engineering*, IJCAI, 1977, pp. 1014-1029.
- [FILLMORE 68] C. Fillmore: *the case for case*. In Bach and Harms eds.: *Universals in linguistic theory*, Holt, Rinehart and Winston Inc., Chicago, 1968, pp. 1-88.
- [FORBUS 84] K. Forbus: *qualitative process theory*. Artificial Intelligence 24, 1984, pp. 85-168.
- [FORBUS 86a] K. Forbus: *interpreting observations of physical systems*. Report R-86-1248, University of Illinois, 1986.
- [FORBUS 86b] K. Forbus: *the qualitative process engine*. Report R-86-1288, University of Illinois, 1986.
- [FORBUS 88] K. Forbus: *interpreting observations of physical systems*. IEEE transactions on systems, man and cybernetics, vol. 17, n°3, 1987, pp. 350-359.
- [GERVAIS 83] M. Gervais: *contrôle de gestion*, Economica, Paris, 1983.
- [GERVAIS 86] M. Gervais, J. Stepniewski: *un système expert pour effectuer l'imputation comptable - principe de fonctionnement d'un tel système*. Revue Française de Comptabilité 172, octobre 1986, pp. 47-57.
- [HART 82] P. Hart: *directions for AI in the eighties*. SIGART Newsletters, n° 79, 1982, pp. 11-15.
- [HART 86] P. Hart, A. Barzilay, R. Duda: *qualitative reasoning for financial assessments: a prospectus*, AI magazine, vol. 7, n°1, 1986, pp. 62-68.

- [HAYES 79] P. Hayes: *the naive physics manifesto*. In D. Michie ed.: *Expert systems in the micro-electronics age*, Edinburgh University Press, 1979, pp. 242-270.
- [HAYES-ROTH 83] F. Hayes-Roth, D. Waterman, D. Lenat: *building expert systems*, Addison-Wesley, Reading, MA, 1983.
- [HEISE 75] D. Heise: *causal analysis*, John Wiley & Sons, New-York, 1975.
- [HELPERT 82] E. Helfert: *techniques of financial analysis*, Homewood, 1982.
- [HENDRIX 73] G. Hendrix, C. Thompson, J. Slocum: *language processing via canonical verbs and semantic models*, IJCAI, 1973, pp. 262-269.
- [HOPCROFT 73] J. Hopcroft, R. Karp: *an $o(n^{5/2})$ algorithm for maximum matchings in bipartite graphs*. SIAM Journal of computing, vol. 2, 1973, pp. 225-231.
- [IJIRI 64] Y. Ijiri: *axioms and structures of conventional accounting measurement*. Technical Report n°128, Stanford, 1964.
- [IVANESCU 66] P. Ivanescu, S. Rudeanu: *a pseudo-boolean approach to matching problems*, International Symposium on Theory of Graphs, 1966, pp. 161-175.
- [IWASAKI 86a] Y. Iwasaki, H. Simon: *causality in device behavior*. Artificial Intelligence 29, 1986, pp. 3-32.
- [IWASAKI 86b] Y. Iwasaki, H. Simon: *theories of causal ordering: reply to De Kleer and Brown*. Artificial Intelligence 29, 1986, pp. 63-72.
- [IWASAKI 87] Y. Iwasaki: *generating behavior equations from explicit representation of mechanisms*. Report CMU-CS-87-131, Carnegie Mellon University, 1987.
- [IWASAKI 88] Y. Iwasaki: *causal ordering in a mixed structure*, AAAI, 1988, pp. 313-318.
- [KEVORKIAN 75] A. Kevorkian: *structural aspects of large dynamic systems*, 6^e congrès IFAC, 1975, pp. 19.3, 1-10.

- [KING 86] D. King: *ERGO : an explanation facility for decision support systems*, 6^e Journées Systèmes Experts et leurs applications, Avignon, 1986, pp. 991-1013.
- [KOSY 84] D.W. Kosy, B.P. Wise: *self-explanatory financial planning models*, AAAI, 1984, pp. 176-181.
- [KOSY 86] D.W. Kosy, B.P. Wise: *overview of ROME: a reason-oriented modeling environment*, Workshop on Artificial Intelligence in Economics and Management, 1986, pp. 21-30.
- [KOTON 85] P. Koton: *empirical and model-based reasoning in expert systems*, IJCAI, 1985, pp. 297-299.
- [KUIPERS 84a] B. Kuipers: *commonsense reasoning about causality: deriving behavior from structure*. Artificial Intelligence 24, 1984, pp. 169-203.
- [KUIPERS 84b] B. Kuipers, J. Kassirer: *causal reasoning in medicine: analysis of a protocol*. Cognitive Science, vol. 8, 1984, pp. 363-385.
- [KUIPERS 86] B. Kuipers: *qualitative simulation*. Artificial Intelligence, vol. 29, 1986, pp. 289-338.
- [KUIPERS 87] B. Kuipers, C. Chiu: *taming intractible branching in qualitative simulation*, IJCAI, 1987, pp. 1079-1085.
- [KUIPERS 88] B. Kuipers, D. Berleant: *using incomplete quantitative knowledge in qualitative reasoning*, AAAI, 1988, pp. 324-329.
- [KUNZ 83] J. Kunz: *analysis of physiological behavior using a causal model based on first principles*, AAAI, 1983, pp. 225-228.
- [LAURIERE 82a] J. L. Laurière: *représentation et utilisation des connaissances*. Technique et Science Informatiques, vol. 1, n° 1 et 2, 1982.

- [LAURIERE 82b] J.L. Laurière: *le système SNARK: Symbolic Normalized Acquisition and Representation of Knowledge*. Rapport de l'Institut de Programmation n°427, 1982.
- [LESCA 86] H. Lesca: *Structure et système d'information - facteurs de compétitivité de l'entreprise*, Masson, Paris, 1986.
- [LMI 88] *Vous avez dit «opérationnel» ?*. Le Monde Informatique, 30 mai 1988.
- [LONG 83] W. Long: *reasoning about state from causation and time in medical domain*, AAAI, 1983, pp. 251-254.
- [MALLOUM 80] A. Malloum: *analyse structurale et partition des systèmes macro-économiques dynamiques complexes. Application à Mini-D.M.S.* Thèse de troisième cycle, Université Lyon I, 1980.
- [MICHA 86] B. Micha: *AIDE: un système expert de diagnostic d'entreprise*. Rapport de la Direction Générale des Etudes - Banque de France, nov. 1986.
- [MICHIE 82] D. Michie: *high-road and low-road programs*. AI magazine, vol.3, 1982, pp. 21-22.
- [MORGAN 87] A. Morgan: *predicting the behaviour of dynamic systems with qualitative vectors*. AISB, pp. 81-95, 1987.
- [MORVAN 85] Y. Morvan: *fondements d'Economie Industrielle*, Economica, Paris, 1985, pp. 73-85.
- [NEPOMIASTCHY 78] P. Népomiashtchy, A. Ravelli, F. Rechenmann: *an automatic method to get an econometric model in a quasi-triangular form*. Rapport de recherche IRIA n° 313, 1978.
- [NGUYEN 85] T. Nguyen, W. Perkins, T. Laffrey, D. Pecora: *checking an expert systems knowledge base for consistency and completeness*, IJCAI, 1985, pp. 375-378.
- [PAGE 87] M. Page, B. Oudet, C. Lamrani: *Systèmes experts et étude de la causalité en analyse financière*, Colloque AFCET: Aide à la Décision dans l'Organisation, Paris, 1987.

- [PAGE 89a] M. Page, B. Lamrani, B. Oudet: *systèmes experts et modélisation en Comptabilité*, 9^e Journées Systèmes Experts et leurs applications, Avignon, 1989.
- [PAGE 89b] M. Page: *l'interprétation comparative ou l'explication des différences entre deux situations*, Congrès RFLA, Paris, 1989.
- [PATIL 81] R. Patil, P. Szolovits, W. Schwartz: *causal understanding of patient illness in medical diagnosis*, IJCAI, 1981, pp. 893-899.
- [PEARL 85] J. Pearl: *distributed diagnosis in causal models with continuous variables*. Technical Report CSD-860051, University of California, Los Angeles, 1985.
- [PENG 87] Y. Peng, J. Reggia: *a probabilistic causal model for diagnostic problem solving*. IEEE transactions on systems, man and cybernetics, vol. 17 n° 2, 1987, pp. 146-162.
- [PINSON 89] S. Pinson: *une évaluation multi-expert du risque d'entreprise: le système CREDEX*. Technique et Science Informatiques, vol. 8, n° 2, 1989.
- [PUCCIA 85] C. Puccia, R. Levins: *qualitative modeling of complex systems*, Harvard University Press, Cambridge MA, 1985.
- [RAIMAN 86] O. Raiman: *order of magnitude reasoning*, AAAI, 1986, pp. 100-104.
- [RAIMAN 87] O. Raiman, P. Devès: *troubleshooting: when modeling is the trouble*, AAAI, 1987, pp. 600-605.
- [RAJAGOPALAN 84] R. Rajagopalan: *qualitative modeling in the turbojet engine domain*, AAAI, 1984, pp. 283-287.
- [RECHENMANN 87] F. Rechenmann, M. Doize: *SAFIR-SHIRKA: un système à base de connaissance centrée objet pour l'analyse financière*, 7^e Journées Systèmes Experts et leurs applications, Avignon, 1987, pp. 947-969.

- [RIEGER 77] C. Rieger, M. Grinberg: *the declarative representation and procedural simulation of causality in physical mechanisms*, IJCAI, 1977, pp. 250-256.
- [ROUSSET 87] M.C. Rousset: *sur la validité des bases de connaissances: le système COVADIS*, 6^e journées systèmes experts et leurs applications, Avignon, 1987, pp. 269-282.
- [SACKS 88] E. Sacks: *qualitative analysis by piecewise linear approximation*. International Journal of Artificial Intelligence Engineering, vol.3, n°3, 1988, pp.151-155.
- [SHIBAHARA 83] T. Shibahara, J. Tsotsos, J. Mylopoulos, D. Covvey: *CAA: a knowledge-based system using causal knowledge to diagnose cardiac rhythm disorders*, IJCAI, 1983, pp. 242-245.
- [SHIBAHARA 85] T. Shibahara: *on using causal knowledge to recognize vital signals: knowledge-based interpretation of arrhythmias*, IJCAI, 1985, pp. 307-314.
- [SHORTLIFFE 76] E. Shortliffe: *computer-based medical consultations: MYCIN*, Elsevier, New-York, 1976.
- [SIMMONS 83] R. Simmons: *the use of qualitative and quantitative simulations*, AAAI, 1983, pp. 364-368.
- [SIMMONS 86] R. Simmons: *"commonsense" arithmetic reasoning*, AAAI, 1986, pp. 118-124.
- [SIMON 57a] H. Simon: *causal ordering and identifiability*. In H. Simon: *Models of man*, John Wiley & Sons, New-York, 1957, pp. 10-36.
- [SIMON 57b] H. Simon: *on the definition of the causal relation*. In H. Simon: *Models of man*, John Wiley & Sons, New-York, 1957, pp. 50-61.
- [STANSFIELD 80] J. Stansfield: *Conclusion from the commodity expert project*, MIT Artificial Intelligence memo. 601, 1980.
- [STEELS 85] L. Steels: *second generation expert systems*. Future Generation Computer Systems, vol. 1, n°4, 1985, pp. 213-221.

- [STEPNIEWSKI 87] J. Stepniewski: *principes de la comptabilité événementielle*, Masson, Paris, 1987.
- [STEWART 62] D. Steward: *on an approach to techniques for the analysis of the structure of large systems of equations*. SIAM Review, vol. 4, n° 4, 1962, pp. 321-342.
- [STRUSS 88] P. Struss: *mathematical aspects of qualitative reasoning*. International Journal of Artificial Intelligence Engineering, vol.3, n°3, 1988, pp.156-169.
- [SWARTOUT 83] W. Swartout: *XPLAIN: a system for creating and explaining expert consulting programs*. Artificial Intelligence 21, 1983, pp. 285-325.
- [TARJAN 72] R. Tarjan: *depth-first search and linear graph algorithms*. SIAM Journal of computing, vol. 1, 1972, pp. 146-160.
- [VINCENT 68] L. Vincent: *mesure de la productivité globale au niveau de l'entreprise*. Gestion, mai 1968, pp. 373-387.
- [WATERMAN 86] D. Waterman: *a guide to expert systems*, Addison-Wesley, Reading, MA, 1986.
- [WEISS 78] S. Weiss, C. Kulikowski, S. Amarel, A. Safir: *a model-based method for computer-aided medical reasoning*. Artificial Intelligence 11, 1978, pp. 145-172.
- [WELD 86] D. Weld: *the use of causal aggregation in causal reasoning*. Artificial Intelligence, vol. 30, 1986, pp. 1-34.
- [WELD 88a] D. Weld: *choice for comparative analysis: DQ analysis or exaggeration ?* International Journal of Artificial Intelligence Engineering, vol.3, n°3, 1988, pp.174-180.
- [WELD 88b] D. Weld: *comparative analysis*. Artificial Intelligence, vol. 36, 1988, pp. 333-373.
- [WELLMAN 87] M. Wellman: *probabilistic semantics for qualitative influences*, AAAI, 1987, pp. 660-664.

- [WIELINGA 86] B. Wielinga, J. Breuker: *models of expertise*. Proc. ECAI, 1986, pp.306-318.
- [WILLIAMS 84a] B. Williams: *the use of continuity in a qualitative physics*, AAI, 1984, pp. 350-354.
- [WILLIAMS 84b] B. Williams: *qualitative analysis of MOS circuits*. Artificial Intelligence 24, 1984, pp. 281-346.
- [WOODS 70] W. Woods: *transition network grammars for natural language analysis*. Communications of the ACM, vol. 13, n°10, 1970, pp. 591-606.
- [YIP 87] K. Yip: *extracting qualitative dynamics from numerical experiments*, AAI, 1987, pp. 665-670.

ANNEXE A

UN EXEMPLE DE RÉSULTAT FOURNI PAR LE SYSTÈME ANAFI

A.1 - L'exemple

Un extrait de la trace d'une exécution du système ANAFI sur une entreprise fictive simplifiée est présentée ci-dessous. Les équations, relations d'influences et expressions de jugement relatifs à cet extrait figurent en sections A.2 et A.3.

2.1 - Les liquidités sont faibles et diminuent nettement.

La faiblesse des liquidités s'explique par la longueur de la phase de fabrication et la courte durée de la phase d'approvisionnement.

La diminution des liquidités est causé par l'allongement de la phase de fabrication et par l'allongement de la phase de distribution, non compensé par l'allongement de la phase d'approvisionnement.

2.1.1 - Le ratio trésorerie/CA est faible et diminue de 78,9 % : 28 jours de CA en 86, 7 jours de CA en 87.

La diminution du ratio trésorerie/CA est dû à la diminution de la trésorerie de 80,0 %.

La diminution de la trésorerie s'explique par l'augmentation du besoin en fonds de roulement de 20 kF en 86 à 65 kF en 87, non compensé par l'augmentation du fonds de roulement de 7,1 %.

L'augmentation du besoin en fonds de roulement est causée par l'augmentation des emplois cycliques de 11,4 % et à un degré moindre par la diminution des ressources cycliques de 1,5 %.

2.1.2 - Les ratios du cycle d'exploitation sont assez mauvais et se dégradent nettement.

2.1.2.1 - La durée d'immobilisation des stocks est longue et augmente de 48,0 % : 28 jours de CA en 86, 41 jours de CA en 87.

L'allongement de la durée d'immobilisation des stocks est dû à l'augmentation du stock moyen de 25,0 % et à la diminution du CA de 5,4 %.

2.1.2.2 - La durée du crédit clients est normale et augmente de 8,6 % : 14 jours de CA en 86, 15 jours de CA en 87.

L'allongement de la durée du crédit clients s'explique par la diminution du CA de 5,4 % et par l'augmentation des créances clients de 2,2 % .

2.1.2.3 - la durée du crédit fournisseurs est courte et augmente de 7,1 % : 11 jours d'achats en 86, 12 jours d'achats en 87.

L'allongement de la durée du crédit fournisseurs est causé par l'augmentation des dettes aux fournisseurs de 10,6 %, non compensé par la progression des achats de 3,2 %.

A.2 - Le système d'équations et les relations d'influences relatifs à l'extrait présenté

$duree_credit_client = ((creance_client + effet_recevoir) / chiffre_affaire) * 360;$
 $duree_immob_stock = (stock_moyen / chiffre_affaire) * 360;$
 $duree_credit_fournisseur = ((dette_fournisseur + effet_a_payer) / achat) * 360;$
 $tresorerie_sur_chiffre_affaire = (tresorerie / chiffre_affaire) * 360;$
 $tresorerie = fonds_de_roulement - besoin_fonds_de_roulement;$
 $fonds_de_roulement = ressource_stable - emploi_stable;$
 $besoin_fonds_de_roulement = emploi_cyclique - ressource_cyclique;$

$duree_phase_approvisionnement$ l+ liquidite;
 $duree_phase_fabrication$ l- liquidite;
 $duree_phase_distribution$ l- liquidite;

A.3 - Les expressions de jugement relatives à l'extrait présenté

liquidite vaut $3 * tresorerie_sur_chiffre_affaire + 1 * ratios_cycle_exploitation$;
 $r ratios_cycle_exploitation$ vaut $-1 * duree_immob_stock + -1 * duree_credit_client$
 $+ 1 * duree_credit_fournisseur$;
 $duree_phase_approvisionnement$ vaut $1 * duree_credit_fournisseur$;
 $duree_phase_fabrication$ vaut $1 * duree_immob_stock$;
 $duree_phase_distribution$ vaut $1 * duree_credit_client$;

$duree_credit_client$ vaut [0 -> -5 15 -> 0 40 -> 5] [-0.3 -> -5 0 -> 0 0.3 -> 5];
 $duree_credit_fournisseur$ vaut [0 -> -5 15 -> 0 40 -> 5] [-0.3 -> -5 0 -> 0 0.3 -> 5];
 $duree_immob_stock$ vaut [0 -> -5 15 -> 0 40 -> 5] [-0.3 -> -5 0 -> 0 0.3 -> 5];
 $tresorerie_sur_chiffre_affaire$ vaut [0 -> -5 10 -> 0 30 -> 5] [-0.3 -> -5 0 -> 0 0.3 -> 5];

ANNEXE B

UN EXEMPLE DE RÉSULTAT FOURNI PAR LE SYSTEME SURPLUS

B.1 - L'exemple

La trace d'une exécution de SURPLUS ci-dessous est issue de l'analyse de surplus d'une entreprise réelle. Le nom de cette entreprise a été supprimée pour des raisons de confidentialité. Les systèmes d'équations général et spécifique associés à cette entreprise sont présentés dans les sections B.2 et B.3 de cette annexe. Le tableau résumant les principales variations des soldes intermédiaires de gestion et le diagramme de répartition du surplus n'ayant pu être insérés directement dans ce texte, ont été reproduits à partir du résultat réel.

Analyse du surplus de l'entreprise xxxxxxxxx de 1986 à 1987

1 - Analyse de la rentabilité de l'entreprise

L'excédent brut d'exploitation augmente de 27,5 % : 63 975,00 kF en 1986, 81 560,00 kF en 1987. Ceci est dû à l'augmentation du produit global de 9,3 %, non compensé par l'accroissement des consommations intermédiaires de 9,0 %.

La progression du produit global s'explique par l'augmentation de la production de l'exercice de 9,2 %.

L'accroissement de la production de l'exercice est causé par la progression de la production vendue de 9,0 %.

L'augmentation de la production vendue est dû à l'accroissement des ventes de lait de consommation de 15,5 %, à la progression des ventes de crème d'expédition de 60,6 % et à l'augmentation des ventes de produits frais de 12,5 %, non compensé par la diminution des ventes de beurres et poudres de 29,6 %.

Le tableau suivant résume les principales variations des soldes intermédiaires de gestion:

Annexe B - un exemple de résultat fourni par le système SURPLUS

postes	valeur en 1986 kF	valeur en 1987 kF	var. valeur %	var. volume %	var. prix %
PRODUIT GLOBAL	1 508 421,00	1 648 313,00	9,27	12,97	-5,81
ventes de lait de consommation	453 000,00	523 000,00	15,45	28,36	-12,42
ventes de lait d'expédition	94 000,00	151 000,00	60,64	55,41	0,65
ventes de produits frais	360 000,00	405 000,00	12,50	9,42	0,11
ventes de lait d'expédition	430 000,00	431 000,00	0,23	4,93	-6,99
ventes de beurres et de poudres	108 000,00	76 000,00	-29,63	-33,65	3,28
CONSOMMATIONS INTERMEDIAIRES	1 342 278,00	1 462 434,00	8,95	12,60	-5,78
achats de lait	801 134,00	924 969,00	15,46	11,13	1,16
emballages	115 499,00	134 282,00	16,26	12,60	0,54
autres charges extérieures	92 598,00	104 192,00	12,52	9,56	0,00
sous-traitance	32 267,00	39 225,00	21,56	18,37	0,00
achats d'autres produits laitiers	236 526,00	192 509,00	-18,61	20,21	-34,07
CHARGES DE PERSONNEL	104 550,00	106 876,00	2,22	1,19	-1,64
IMPOTS ET TAXES	6 206,00	7 383,00	18,97	15,84	0,00

2 - Analyse de la formation du surplus de productivité global

L'entreprise xxxxxxxx a enregistré un surplus de productivité globale de 25 371,07 kF entre 1986 et 1987. Ceci s'explique par l'accroissement en volume du produit global de 13,0 %, non compensé par la progression en volume des consommations intermédiaires de 12,6 %.

L'augmentation en volume du produit global est causé par l'accroissement en volume de la production de l'exercice de 13,0 %.

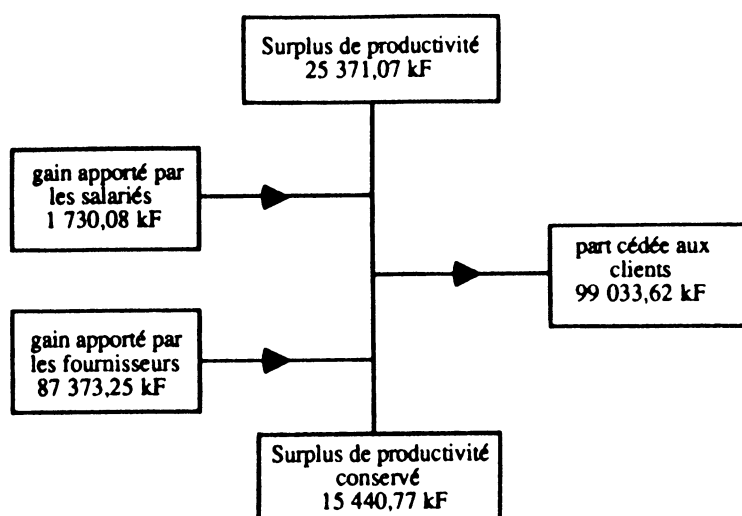
La progression en volume de la production de l'exercice est dû à l'augmentation en volume de la production vendue de 12,7 %.

L'accroissement en volume de la production vendue s'explique par la progression en volume des ventes de lait de consommation de 28,4 % et à un degré moindre par l'augmentation en volume des ventes de crème d'expédition de 55,4 %, par l'accroissement en volume des ventes de produits frais de 9,4 % et par la progression en volume des ventes de lait d'expédition de 4,9 %, non compensé par la chute en volume des ventes de beurres et poudres de 33,7 %.

3 - Analyse de la répartition du surplus de productivité global

Le surplus conservé par l'entreprise s'élève à 15 440,77 kF. Le diagramme suivant montre la répartition du surplus de productivité global.

Annexe B - un exemple de résultat fourni par le système SURPLUS



L'interprétation de ce diagramme pour les différents partenaires de l'entreprise est la suivante:

a - les salariés

Le gain apporté par les salariés de 1 730,08 kF représente une évolution du prix des charges de personnel favorable à l'entreprise.

b - les fournisseurs

Le gain apporté par les fournisseurs de 87 373,25 kF représente une évolution du prix des consommations intermédiaires favorable à l'entreprise. Ceci est causé par la décroissance du prix des approvisionnements consommés de 6,4 %.

La diminution du prix des approvisionnements consommés est dû à la chute du prix des matières premières de 6,4 %.

La décroissance du prix des matières premières s'explique par la diminution du prix des autres produits laitiers de 34,1 %, non compensé par l'augmentation du prix du lait de 1,2 %.

c - les clients

La part cédée aux clients de 99 033,62 kF représente une évolution du prix du produit global défavorable à l'entreprise. Ceci est causé par la chute du prix de la production de l'exercice de 5,8 %.

La décroissance du prix de la production de l'exercice est dû à la diminution du prix de la production vendue de 5,8 %.

La chute du prix de la production vendue s'explique par la décroissance du prix du lait de consommation de 12,4 % et à un degré moindre par la diminution du prix du lait d'expédition de 7,0 %.

B.2 - Les équations comptables communes à toutes les entreprises utilisées pour l'analyse de surplus

$ebe = \text{produit_global} - \text{conso_intermediaire} - \text{impot_taxe} + \text{subvention_exploitation} - \text{charge_personnel} + \text{transfert_charge_personnel};$

$\text{produit_global} = \text{produit_exercice} + \text{marge_commerciale};$
 $\text{produit_exercice} = \text{production_vendue} + \text{production_stockee} + \text{production_immobilisee} + \text{transfert_charge_production};$

$\text{charge_personnel} = \text{effectif} * \text{salaire};$

$\text{marge_commerciale} = \text{vente_marchandise} - \text{cout_achat_marchandise};$
 $\text{cout_achat_marchandise} = \text{achat_marchandise} + \text{variation_stock_marchandise};$

$\text{conso_intermediaire} = \text{appro_consomme} + \text{transfert_charge_conso_intermediaire} + \text{autre_achat};$
 $\text{appro_consomme} = \text{achat_matiere_premiere} + \text{variation_stock_matiere_premiere};$

$\text{autre_achat} = \text{sous_traitance} + \text{personnel_exterieur} + \text{autre_charge_exterieure};$

B.3 - Les équations spécifiques à l'entreprise analysée

$\text{production_vendue} = \text{lait_consommation} + \text{lait_expedition} + \text{creme_expedition} + \text{beurre_poudre} + \text{produit_frais} + \text{fromage};$

$\text{achat_matiere_premiere} = \text{lait} + \text{autre_produit_laitier};$

$\text{lait_consommation} = \text{prix_lait_consommation} * \text{qte_lait_consommation};$
 $\text{lait_expedition} = \text{prix_lait_expedition} * \text{qte_lait_expedition};$
 $\text{creme_expedition} = \text{prix_creme_expedition} * \text{qte_creme_expedition};$
 $\text{beurre_poudre} = \text{prix_beurre_poudre} * \text{qte_beurre_poudre};$
 $\text{produit_frais} = \text{prix_produit_frais} * \text{qte_produit_frais};$
 $\text{fromage} = \text{prix_fromage} * \text{qte_fromage};$

$\text{lait} = \text{prix_lait} * \text{qte_lait};$
 $\text{autre_produit_laitier} = \text{prix_autre_produit_laitier} * \text{qte_autre_produit_laitier};$

Remarque:

les huit dernières équations sont générées automatiquement à partir de la nomenclature des produits vendus et des matières premières achetées par un outil intégré à SURPLUS.

ANNEXE C

LE MODELE LOGIQUE DE SICA

C.1 - Les règles logiques pour l'imputation comptable

C.1.1 - Les règles de description de l'univers comptable

RU₁ $\forall x$ [tiers-actif(x) \vee tiers-passif(x)] \Rightarrow tiers(x)

RU₂ $\forall x$ [tiers(x) \vee x = ENTREPRISE] \Rightarrow acteur(x)

RU₃ $\forall x$ [bien(x) \vee service(x) \vee mp(x)] \Rightarrow objet-d'échange(x)

C.1.2 - Les règles de l'analyse sémantique

RS₁ $\forall x$ sujet(x) \wedge acteur(x) \Rightarrow agent(x)

RS₂ $\forall x \forall p$ ci(p,x) \wedge prep-origine(p) \wedge acteur(x) \Rightarrow agent(x)

RS₃ $\forall x \forall p$ ci(p,x) \wedge prep-dest(p) \wedge acteur(x) \Rightarrow contre-agent(x)

RS₄ $\forall x$ cd(x) \wedge objet-d'échange(x) \Rightarrow objet(x)

RS₅ $\forall x \forall p$ ci(p,x) \wedge prep-inst(p) \wedge objet-d'échange(x) \Rightarrow contre-objet(x)

C.1.3 - Les règles de l'analyse logique

RL₁ $\forall a \text{ agent}(a) \wedge \text{tiers}(a) \Rightarrow \text{contre-agent}(\text{ENTREPRISE})$

RL₂ $\forall a \text{ contre-agent}(a) \wedge \text{tiers}(a) \Rightarrow \text{agent}(\text{ENTREPRISE})$

RL₃ $\forall o \forall s \forall t \text{ agent}(\text{ENTREPRISE}) \wedge [\text{objet}(o) \vee \text{contre-objet}(o)] \wedge \text{sens}(o,s) \wedge \text{spécialité}(t,o,s) \Rightarrow \text{contre-agent}(t)$

RL₄ $\forall o \forall s \forall t \text{ contre-agent}(\text{ENTREPRISE}) \wedge [\text{objet}(o) \vee \text{contre-objet}(o)] \wedge \text{sens}(o,s) \wedge \text{spécialité}(t,o,s) \Rightarrow \text{agent}(t)$

RL₅ $\forall v \forall o \text{ verbe}(v) \wedge \text{objet}(o) \wedge \text{tv}(v,A) \Rightarrow \text{sens}(o,\text{REC})$

RL₆ $\forall v \forall o \text{ verbe}(v) \wedge \text{contre-objet}(o) \wedge \text{tv}(v,A) \Rightarrow \text{sens}(o,\text{DON})$

RL₇ $\forall v \forall o \text{ verbe}(v) \wedge \text{objet}(o) \wedge \text{tv}(v,V) \Rightarrow \text{sens}(o,\text{DON})$

RL₈ $\forall v \forall o \text{ verbe}(v) \wedge \text{contre-objet}(o) \wedge \text{tv}(v,V) \Rightarrow \text{sens}(o,\text{REC})$

RL₉ $\forall a \forall c \forall o \forall m [\text{objet}(o) \vee \text{contre-objet}(o)] \wedge \text{agent}(a) \wedge \text{contre-agent}(c) \wedge \text{sens}(o,\text{REC}) \wedge \text{montant}(o,m) \Rightarrow \text{flux}(o,c,a,m)$

RL₁₀ $\forall a \forall c \forall o \forall m [\text{objet}(o) \vee \text{contre-objet}(o)] \wedge \text{agent}(a) \wedge \text{contre-agent}(c) \wedge \text{sens}(o,\text{DON}) \wedge \text{montant}(o,m) \Rightarrow \text{flux}(o,a,c,m)$

C.1.4 - Les règles d'imputation

RI₁ $\forall e \forall o \forall m \forall c \text{ flux}(o,e,\text{ENTREPRISE},m) \wedge \text{compte}(o,e,\text{ENTREPRISE},c) \Rightarrow \text{mouvement}(c,\text{DEBIT},m)$

RI₂ $\forall r \forall o \forall m \forall c \text{ flux}(o,\text{ENTREPRISE},r,m) \wedge \text{compte}(o,\text{ENTREPRISE},r,c) \Rightarrow \text{mouvement}(c,\text{CREDIT},m)$

C.2 - Un extrait de la base des termes primitifs

bien(MARCHANDISE)

bien(MATIERE-PREMIERE)

bien(PRODUIT-FINI)

...

mp(ESPECE)

mp(CHEQUE)

mp(CCP)

...

service(LOYER)

service(IMPOT)

service(COTISATION-SOCIALE)

...

tiers-actif(LOCATAIRE)

tiers-actif(CLIENT)

tiers-actif(DEBITEUR-DIVERS)

...

tiers-passif(PROPRIETAIRE)

tiers-passif(FOURNISSEUR)

tiers-passif(CREDITEUR-DIVERS)

...

prep-origine(PAR)

prep-origine(DE)

...

prep-dest(A)

prep-dest(DE)

...

prep-inst(EN)

prep-inst(PAR)

...

tv(RECEVOIR,A)

tv(ACHETER,A)

tv(PAYER,A)

tv(VENDRE,V)

tv(ENVOYER,V)

...

compte(MARCHANDISE, ENTREPRISE, CLIENT, 707)

compte(MARCHANDISE, FOURNISSEUR, ENTREPRISE, 607)

compte(DETTE, ENTREPRISE, FOURNISSEUR, 400)

compte(DETTE, CLIENT, ENTREPRISE, 410)

compte(TVA, ENTREPRISE, FOURNISSEUR, 445)

compte(TVA, FOURNISSEUR, ENTREPRISE, 445)

$\forall x$ compte(ESPECE, x, ENTREPRISE, 530)

$\forall x$ compte(ESPECE, ENTREPRISE, x, 530)

...

spécialité(FOURNISSEUR, MARCHANDISE, REC)

spécialité(CLIENT, MARCHANDISE, DON)

spécialité(BANQUE, CHEQUE, REC)

spécialité(BANQUE, CHEQUE, DON)

spécialité(CAISSE, ESPECE, REC)

spécialité(CAISSE, ESPECE, DON)

spécialité(PROPRIETAIRE, LOYER, REC)

spécialité(LOCATAIRE, LOYER, DON)

...

$\forall e \forall r \forall i$ flux(MARCHANDISE,e,r,i) \Rightarrow flux(TVA,e,r,0.186*i)

...