



HAL
open science

Modèles mathématiques de l'imprécis et de l'incertain en vue d'applications aux techniques d'aide à la décision

Didier Dubois

► **To cite this version:**

Didier Dubois. Modèles mathématiques de l'imprécis et de l'incertain en vue d'applications aux techniques d'aide à la décision. Modélisation et simulation. Université Joseph-Fourier - Grenoble I; Institut National Polytechnique de Grenoble - INPG, 1983. tel-00277223

HAL Id: tel-00277223

<https://theses.hal.science/tel-00277223>

Submitted on 5 May 2008

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

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

THESE

présentée à

l'Université Scientifique et Médicale de Grenoble

et à

l'Institut National Polytechnique de Grenoble

pour obtenir le grade de
DOCTEUR ES SCIENCES
«Mathématiques appliquées»

par

Didier DUBOIS



**MODELES MATHEMATIQUES DE L'IMPRECIS
ET DE L'INCERTAIN EN VUE D'APPLICATIONS
AUX TECHNIQUES D'AIDE A LA DECISION.**



Thèse soutenue le 19 novembre 1983 devant la commission d'examen.

M. METIVIER	}	Président
A. BENSOUSSAN		Examineurs
A. KAUFMANN		
A. LE BRETON		
D. PONASSE		
B. ROY		
M. SAKAROVITCH		

UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

Année universitaire 1982-1983

Président de l'Université : M. TANCHE

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

(RANG A)

SAUF ENSEIGNANTS EN MEDECINE ET PHARMACIE

PROFESSEURS DE 1ère CLASSE

ARNAUD Paul	Chimie organique
ARVIEU Robert	Physique nucléaire I.S.N.
AUBERT Guy	Physique C.N.R.S.
AYANT Yves	Physique approfondie
BARBIER Marie-Jeanne	Electrochimie
BARBIER Jean-Claude	Physique expérimentale C.N.R.S. (labo de magnétisme)
BARJON Robert	Physique nucléaire I.S.N.
BARNOUD Fernand	Biosynthèse de la cellulose-Biologie
BARRA Jean-René	Statistiques - Mathématiques appliquées
BELORISKY Elie	Physique
BENZAKEN Claude (M.)	Mathématiques pures
BERNARD Alain	Mathématiques pures
BERTRANDIAS Françoise	Mathématiques pures
BERTRANDIAS Jean-Paul	Mathématiques pures
BILLET Jean	Géographie
BONNIER Jean-Marie	Chimie générale
BOUCHEZ Robert	Physique nucléaire I.S.N.
BRAVARD Yves	Géographie
CARLIER Georges	Biologie végétale
CAUQUIS Georges	Chimie organique
CHIBON Pierre	Biologie animale
COLIN DE VERDIERE Yves	Mathématiques pures
CRABBE Pierre (détaché)	C.E.R.M.O.
CYROT Michel	Physique du solide
DAUMAS Max	Géographie
DEBELMAS Jacques	Géologie générale
DEGRANGE Charles	Zoologie
DELOBEL Claude (M.)	M.I.A.G. Mathématiques appliquées
DEPORTES Charles	Chimie minérale
DESRE Pierre	Electrochimie
DOLIQUE Jean-Michel	Physique des plasmas
DUCROS Pierre	Cristallographie
FONTAINE Jean-Marc	Mathématiques pures
GAGNAIRE Didier	Chimie physique

.../...

GASTINEL Noël	Analyse numérique - Mathématiques appliquées
GERBER Robert	Mathématiques pures
GERMAIN Jean-Pierre	Mécanique
GIRAUD Pierre	Géologie
IDELMAN Simon	Physiologie animale
JANIN Bernard	Géographie
JOLY Jean-René	Mathématiques pures
JULLIEN Pierre	Mathématiques appliquées
KAHANE André (détaché DAFCO)	Physique
KAHANE Josette	Physique
KOSZUL Jean-Louis	Mathématiques pures
KRAKOWIAK Sacha	Mathématiques appliquées
KUPTA Yvon	Mathématiques pures
LACAZE Albert	Thermodynamique
LAJZEROWICZ Jeannine	Physique
LAJZEROWICZ Joseph	Physique
LAURENT Pierre	Mathématiques appliquées
DE LEIRIS Joël	Biologie
LLIBOUTRY Louis	Géophysique
LOISEAUX Jean-Marie	Sciences nucléaires I.S.N.
LOUP Jean	Géographie
MACHE Régis	Physiologie végétale
MAYNARD Roger	Physique du solide
MICHEL Robert	Minéralogie et pétrographie (géologie)
MOZIERES Philippe	Spectrométrie - Physique
OMONT Alain	Astrophysique
OZENDA Paul	Botanique (biologie végétale)
PAYAN Jean-Jacques (détaché)	Mathématiques pures
PEBAY PEYROULA Jean-Claude	Physique
PERRIAUX Jacques	Géologie
PERRIER Guy	Géophysique
PIERRARD Jean-Marie	Mécanique
RASSAT André	Chimie systématique
RENARD Michel	Thermodynamique
RICHARD Lucien	Biologie végétale
RINAUDO Marguerite	Chimie CERMAV
SENGEL Philippe	Biologie animale
SERGERAERT Francis	Mathématiques pures
SOUTIF Michel	Physique
VAILLANT François	Zoologie
VALENTIN Jacques	Physique nucléaire I.S.N.
VAN CUTSEN Bernard	Mathématiques appliquées
VAUQUOIS Bernard	Mathématiques appliquées
VIALON Pierre	Géologie

PROFESSEURS DE 2ème CLASSE

ADIBA Michel	Mathématiques pures
ARMAND Gilbert	Géographie

AURIAULT Jean-Louis	Mécanique
BEGUIN Claude (M.)	Chimie organique
BOEHLER Jean-Paul	Mécanique
BOITET Christian	Mathématiques appliquées
BORNAREL Jean	Physique
BRUN Gilbert	Biologie
CASTAING Bernard	Physique
CHARDON Michel	Géographie
COHENADDAD Jean-Pierre	Physique
DENEUVILLE Alain	Physique
DEPASSEL Roger	Mécanique des fluides
DOUCE Roland	Physiologie végétale
DUFRESNOY Alain	Mathématiques pures
GASPARD François	Physique
GAUTRON René	Chimie
GIDON Maurice	Géologie
GIGNOUX Claude (M.)	Sciences nucléaires I.S.N.
GUITTON Jacques	Chimie
HACQUES Gérard	Mathématiques appliquées
HERBIN Jacky	Géographie
HICTER Pierre	Chimie
JOSELEAU Jean-Paul	Biochimie
KERCKOVE Claude (M.)	Géologie
LE BRETON Alain	Mathématiques appliquées
LONGEQUEUE Nicole	Sciences nucléaires I.S.N.
LUCAS Robert	Physiques
LUNA Domingo	Mathématiques pures
MASCLE Georges	Géologie
NEMOZ Alain	Thermodynamique (CNRS - CRTBT)
OUDET Bruno	Mathématiques appliquées
PELMONT Jean	Biochimie
PERRIN Claude (M.)	Sciences nucléaires I.S.N.
PFISTER Jean-Claude (détaché)	Physique du solide
PIBOULE Michel	Géologie
PIERRE Jean-Louis	Chimie organique
RAYNAUD Hervé	Mathématiques appliquées
ROBERT Gilles	Mathématiques pures
ROBERT Jean-Bernard	Chimie physique
ROSSI André	Physiologie végétale
SAKAROVITCH Michel	Mathématiques appliquées
SARROT REYNAUD Jean	Géologie
SAXOD Raymond	Biologie animale
SOUTIF Joanne	Physique
SCHOOL Pierre-Claude	Mathématiques appliquées
STUTZ Pierre	Mécanique
SUBRA Robert	Chimie
VIDAL Michel	Chimie organique
VIVIAN Robert	Géographie

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

Année universitaire 1982-1983

Président de l'Université : D. BLOCH

**Vice-Président : René CARRE
Hervé CHERADAME
Marcel IVANES**

PROFESSEURS DES UNIVERSITES :

ANCEAU François	E.N.S.I.M.A.G.
BARRAUD Alain	E.N.S.I.E.G.
BAUDELET Bernard	E.N.S.I.E.G.
BESSON Jean	E.N.S.E.E.G.
BLIMAN Samuel	E.N.S.E.R.G.
BLOCH Daniel	E.N.S.I.E.G.
BOIS Philippe	E.N.S.H.G.
BONNETAIN Lucien	E.N.S.E.E.G.
BONNIER Etienne	E.N.S.E.E.G.
BOUVARD Maurice	E.N.S.H.G.
BRISSONNEAU Pierre	E.N.S.I.E.G.
BUYLE BODIN Maurice	E.N.S.E.R.G.
CAVAIGNAC Jean-François	E.N.S.I.E.G.
CHARTIER Germain	E.N.S.I.E.G.
CHENEVIER Pierre	E.N.S.E.R.G.
CHERADAME Hervé	U.E.R.M.C.P.P.
CHERUY Arlette	E.N.S.I.E.G.
CHIAVERINA Jean	U.E.R.M.C.P.P.
COHEN Joseph	E.N.S.E.R.G.
COUMES André	E.N.S.E.R.G.
DURAND Francis	E.N.S.E.E.G.
DURAND Jean-Louis	E.N.S.I.E.G.
FELICI Noël	E.N.S.I.E.G.
FOULARD Claude	E.N.S.I.E.G.
GENTIL Pierre	E.N.S.E.R.G.
GUERIN Bernard	E.N.S.E.R.G.
GUYOT Pierre	E.N.S.E.E.G.
IVANES Marcel	E.N.S.I.E.G.
JAUSSAUD Pierre	E.N.S.I.E.G.
JOUBERT Jean-Claude	E.N.S.I.E.G.
JOURDAIN Geneviève	E.N.S.I.E.G.
LACOUME Jean-Louis	E.N.S.I.E.G.
LATOMBE Jean-Claude	E.N.S.I.M.A.G.

.../...

LESSIEUR Marcel	E.N.S.H.G.
LESPINARD Georges	E.N.S.H.G.
LONGEQUEUE Jean-Pierre	E.N.S.I.E.G.
MAZARE Guy	E.N.S.I.M.A.G.
MOREAU René	E.N.S.H.G.
MORET Roger	E.N.S.I.E.G.
MOSSIERE Jacques	E.N.S.I.M.A.G.
PARIAUD Jean-Charles	E.N.S.E.E.G.
PAUTHENET René	E.N.S.I.E.G.
PERRET René	E.N.S.I.E.G.
PERRET Robert	E.N.S.I.E.G.
PIAU Jean-Michel	E.N.S.H.G.
POLOUJADOFF Michel	E.N.S.I.E.G.
POUPOT Christian	E.N.S.E.R.G.
RAMEAU Jean-Jacques	E.N.S.E.E.G.
RENAUD Maurice	U.E.R.M.C.P.P.
ROBERT André	U.E.R.M.C.P.P.
ROBERT François	E.N.S.I.M.A.G.
SABONNADIERE Jean-Claude	E.N.S.I.E.G.
SAUCIER Gabrielle	E.N.S.I.M.A.G.
SCHLENKER Claire	E.N.S.I.E.G.
SCHLENKER Michel	E.N.S.I.E.G.
SERMET Pierre	E.N.S.E.R.G.
SILVY Jacques	U.E.R.M.C.P.P.
SOHM Jean-Claude	E.N.S.E.E.G.
SOUQUET Jean-Louis	E.N.S.E.E.G.
VEILLON Gérard	E.N.S.I.M.A.G.
ZADWORNY François	E.N.S.E.R.G.

PROFESSEURS ASSOCIES

BASTIN Georges	E.N.S.H.G.
BERRIL John	E.N.S.H.G.
CARREAU Pierre	E.N.S.H.G.
GANDINI Alessandro	U.E.R.M.C.P.P.
HAYASHI Hirashi	E.N.S.I.E.G.

PROFESSEURS UNIVERSITE DES SCIENCES SOCIALES (Grenoble II)

BOLLIET Louis
Chatelin Françoise

PROFESSEURS E.N.S. Mines de Saint-Etienne

RIEU Jean
SOUSTELLE Michel

CHERCHEURS DU C.N.R.S.

FRUCHART Robert
VACHAUD Georges

Directeur de Recherche
Directeur de Recherche

.../...

ALLIBERT Michel	Maître de Recherche
ANSARA Ibrahim	Maître de Recherche
ARMAND Michel	Maître de Recherche
BINDER Gilbert	
CARRE René	Maître de Recherche
DAVID René	Maître de Recherche
DEPORTES Jacques	
DRIOLE Jean	Maître de Recherche
GIGNOUX Damien	
GIVORD Dominique	
GUELIN Pierre	
HOPFINGER Emil	Maître de Recherche
JOUD Jean-Charles	Maître de Recherche
KAMARINOS Georges	Maître de Recherche
KLEITZ Michel	Maître de Recherche
LANDAU Ioan-Dore	Maître de Recherche
LASJAUNIAS J.C.	
MERMET Jean	Maître de Recherche
MUNIER Jacques	Maître de Recherche
PIAU Monique	
PORTESEIL Jean-Louis	
THOLENCE Jean-Louis	
VERDILLON André	

CHERCHEURS du MINISTERE de la RECHERCHE et de la TECHNOLOGIE (Directeurs et Maîtres de Recherches, ENS Mines de St. Etienne)

LESBATS Pierre	Directeur de Recherche
BISCONDI Michel	Maître de Recherche
KOBYLANSKI André	Maître de Recherche
LE COZE Jean	Maître de Recherche
LALAUZE René	Maître de Recherche
LANCELOT Francis	Maître de Recherche
THEVENOT François	Maître de Recherche
TRAN MINH Canh	Maître de Recherche

PERSONNALITES HABILITEES à DIRIGER des TRAVAUX de RECHERCHE (Décision du Conseil Scientifique)

ALLIBERT Colette	E.N.S.E.E.G.
BERNARD Claude	E.N.S.E.E.G.
BONNET Rolland	E.N.S.E.E.G.
CAILLET Marcel	E.N.S.E.E.G.
CHATILLON Catherine	E.N.S.E.E.G.
CHATILLON Christian	E.N.S.E.E.G.
COULON Michel	E.N.S.E.E.G.
DIARD Jean-Paul	E.N.S.E.E.G.
EUSTAPOPOULOS Nicolas	E.N.S.E.E.G.
FOSTER Panayotis	E.N.S.E.E.G.

.../...

GALERIE Alain	E.N.S.E.E.G.
HAMMOU Abdelkader	E.N.S.E.E.G.
MALMEJAC Yves	E.N.S.E.E.G. (CENG)
MARTIN GARIN Régina	E.N.S.E.E.G.
NGUYEN TRUONG Bernadette	E.N.S.E.E.G.
RAVAINE Denis	E.N.S.E.E.G.
SAINFORT	E.N.S.E.E.G. (CENG)
SARRAZIN Pierre	E.N.S.E.E.G.
SIMON Jean-Paul	E.N.S.E.E.G.
TOUZAIN Philippe	E.N.S.E.E.G.
URBAIN Georges	E.N.S.E.E.G. (Laboratoire des ultra-réfractaires ODEILLON)
GUILHOT Bernard	E.N.S. Mines Saint Etienne
THOMAS Gérard	E.N.S. Mines Saint Etienne
DRIVER Julien	E.N.S. Mines Saint Etienne
BARIBAUD Michel	E.N.S.E.R.G.
BOREL Joseph	E.N.S.E.R.G.
CHOVET Alain	E.N.S.E.R.G.
CHEHIKIAN Alain	E.N.S.E.R.G.
DOLMAZON Jean-Marc	E.N.S.E.R.G.
HERAULT Jeanny	E.N.S.E.R.G.
MONLLOR Christian	E.N.S.E.R.G.
BORNARD Guy	E.N.S.I.E.G.
DESCHIZEAU Pierre	E.N.S.I.E.G.
GLANGEAUD François	E.N.S.I.E.G.
KOFMAN Walter	E.N.S.I.E.G.
LEJEUNE Gérard	E.N.S.I.E.G.
MAZUER Jean	E.N.S.I.E.G.
PERARD Jacques	E.N.S.I.E.G.
REINISCH Raymond	E.N.S.I.E.G.
ALEMANY Antoine	E.N.S.H.G.
BOIS Daniel	E.N.S.H.G.
DARVE Félix	E.N.S.H.G.
MICHEL Jean-Marie	E.N.S.H.G.
OBLED Charles	E.N.S.H.G.
ROWE Alain	E.N.S.H.G.
VAUCLIN Michel	E.N.S.H.G.
WACK Bernard	E.N.S.H.G.
BERT Didier	E.N.S.I.M.A.G.
CALMET Jacques	E.N.S.I.M.A.G.
COURTIN Jacques	E.N.S.I.M.A.G.
COURTOIS Bernard	E.N.S.I.M.A.G.
DELLA DORA Jean	E.N.S.I.M.A.G.
FONLUPT Jean	E.N.S.I.M.A.G.
SIFAKIS Joseph	E.N.S.I.M.A.G.
CHARUEL Robert	U.E.R.M.C.P.P.
CADET Jean	C.E.N.G.
COEURE Philippe	C.E.N.G. (LETI)

.../...

DELHAYE Jean-Marc
DUPUY Michel
JOUBE Hubert
NICOLAU Yvan
NIFENECKER Hervé
PERROUD Paul
PEUZIN Jean-Claude
TAIEB Maurice
VINCENDON Marc

C.E.N.G. (STT)
C.E.N.G. (LETI)
C.E.N.G. (LETI)
C.E.N.G. (LETI)
C.E.N.G.
C.E.N.G.
C.E.N.G. (LETI)
C.E.N.G.
C.E.N.G.

LABORATOIRES EXTERIEURS

DEMOULIN Eric
DEVINE
GERBER Roland
MERCKEL Gérard
PAULEAU Yves
GAUBERT C.

C.N.E.T.
C.N.E.T. (R.A.B.)
C.N.E.T.
C.N.E.T.
C.N.E.T.
I.N.S.A. Lyon

Je remercie tous ceux qui ont bien voulu s'intéresser à mon travail, l'encourager, ou l'examiner.

- Monsieur Michel Métivier me fait l'honneur d'assurer la présidence du jury. Je suis très touché de l'attention qu'a prêté à cette thèse un éminent spécialiste de la théorie des probabilités.

- Monsieur Michel Sakarovitch, mon directeur de recherches, n'a cessé de me prodiguer des conseils utiles. La correspondance fournie que nous avons échangée m'a fortement aidé à clarifier mes idées à propos de notions que leur caractère récent rend délicates à saisir selon toutes leurs facettes.

- La caution que m'apporte ici Monsieur Bensoussan justifie à mes yeux mes efforts pour sortir la théorie des ensembles flous de ses chapelles.

- Monsieur Kaufmann m'a très tôt encouragé dans la voie que j'ai suivie. Son enthousiasme a eu une influence réelle sur mon gout pour la recherche.

- Monsieur Le Breton, en apportant l'opinion d'un statisticien, confère un surcroît de crédibilité à ce travail.

- Monsieur Ponasse, en tant que logicien, m'honore beaucoup en se penchant sur un travail d'ingénieur

- Monsieur Roy m'a plusieurs fois invité à présenter mon travail dans le cadre des séminaires qu'il anime au LAMSADE, et m'a ainsi fourni l'occasion de fructueuses discussions.

- Monsieur Zadeh m'accorde son soutien depuis longtemps déjà et son appui m'a beaucoup aidé à surmonter mes inquiétudes quand l'avenir restait pour moi quelque peu aléatoire.

Cette thèse n'aurait jamais pu être menée à bien sans l'accueil des différents laboratoires où j'ai passé les six dernières années.

Je suis très reconnaissant au Professeur K.S. Fu pour une année passée à l'Université de Purdue, à Michel Sakarovitch et l'équipe de Jean Mermet

pour une année passée à l'IMAG, Grenoble. J'ai trouvé au Département d'Automatique du Centre d'Etudes et de Recherches de Toulouse un environnement de travail très enrichissant, tant sur le plan scientifique qu'humain. Je sais gré à l'Equipe Production Automatisée de m'avoir permis de continuer des recherches sur l'imprécis et l'incertain de rédiger ma thèse, tout en participant aux travaux en cours sur les systèmes automatisés de production. Je ne peux que remercier Jean-Pierre Forestier, Chef du DERA, pour sa compréhension et l'attention qu'il a accordée au présent travail. L'intérêt qu'il semble avoir suscité chez certains collègues du CERT m'est aussi un précieux encouragement.

Le contenu de cette thèse est en grande partie le fruit d'une collaboration avec Henri Prade, collaboration qui date de longtemps déjà, et qui s'est poursuivie, soutenue par l'amitié, jusqu'à aujourd'hui, et semble appelée à se renforcer encore.

Merci enfin à Michelle Dupuy qui s'est chargée de dactylographier cet épais manuscrit.

Il serait important de démêler pourquoi, quand on parle d'un nez rouge, on se contente de l'affirmation fort imprécise qu'il est rouge, alors qu'il serait possible de le préciser au millième de millimètre près par le moyen des longueurs d'onde ; et pourquoi, au contraire, à propos de cette entité autrement complexe qu'est la ville où l'on séjourne, on veut toujours savoir exactement de quelle ville particulière il s'agit. Ainsi est-on distrait de questions plus importantes.

Robert Musil. L'homme sans qualités.

SOMMAIRE

AVANT-PROPOS

INTRODUCTION ET PRESENTATION DE LA THESE

PARTIE I : Outils mathématiques

CHAPITRE 1 : Mesures d'incertitude

- 1.1 - Mesures monotones de Sugeno
 - 1.2 - Cas particuliers : possibilité, nécessité, probabilité
 - 1.3 - Mesures d'incertitude décomposables
 - 1.3.1 - Caractérisation de l'opérateur de décomposition
 - 1.3.2 - Mesures d'incertitude duales
 - 1.3.3 - Mesures auto-duales
 - 1.3.4 - Conclusion
 - 1.3.5 - Extension au cas d'ensembles infinis : quelques remarques
 - 1.4 - Mesures d'incertain induites par une probabilité et une fonction multivoque
 - 1.4.1 - Probabilités supérieures et inférieures
 - 1.4.2 - Cas particuliers possibilités et probabilités
 - 1.4.3 - Cas particuliers : fonctions d'ensemble décomposables
 - 1.4.4 - Espérances mathématiques supérieure et inférieure
 - 1.4.5 - Itération du processus de génération des probabilités supérieures et inférieures
 - 1.5 - Combinaison d'informations incertaines
 - 1.5.1 - Conditionnement
 - 1.5.2 - Combinaison d'informations incertaines
 - 1.6 - Ponts entre possibilité et probabilité
 - 1.6.1 - Probabilité et possibilité cohérentes
 - 1.6.2 - Epreuves imprécises
 - 1.7 - Possibilité et probabilité qualitatives
 - 1.8 - Conclusion
- Références.

CHAPITRE 2 : Ensembles flous

- 2.1 - Représentations d'un ensemble flou
- 2.2 - Détermination des fonctions d'appartenance
 - 2.2.1 - Catégorie vague au sens d'un individu
 - 2.2.2 - Ensembles flous construits sur des données statistiques
- 2.3 - Opérateurs ensemblistes flous
 - 2.3.1 - Opérateurs de complémentation
 - 2.3.2 - Opérateurs d'intersection et d'union approche axiomatique
 - 2.3.3 - Opérateurs d'union et d'intersection : approches empiriques
 - 2.3.4 - Opérateurs ensemblistes spécifiques des ensembles flous
 - 2.3.4.1 - Opérateurs de moyenne
 - 2.3.4.2 - Opérateurs de somme symétrique
 - 2.3.5 - Opérateurs ensemblistes de différence et d'inclusion
 - 2.3.5.1 - Différence ensembliste et inclusion : 1ère construction
 - 2.3.5.2 - Différences ensemblistes induites par l'inclusion de Zadeh
 - 2.3.4.3 - Différence symétrique et équivalence
- 2.4 - Evaluations scalaires et comparaison d'ensembles flous
 - 2.4.1 - Evaluations scalaires d'un ensemble flou
 - 2.4.1.1 - Entropie ou indice de flou
 - 2.4.1.2 - Précision
 - 2.4.1.3 - Cardinalité
 - 2.4.1.4 - Evaluateurs existentiels et universels
 - 2.4.2 - Indices scalaires de comparaison d'ensembles flous
 - 2.4.3 - Mesures d'incertitude d'événements flous
 - 2.4.3.1 - Probabilité d'un événement flou
 - 2.4.3.2 - Possibilité d'un événement flou
 - 2.4.3.3 - Probabilité supérieure et inférieure d'événements flous
 - 2.4.3.4 - Mesure d'incertain décomposable d'un événement flou
 - 2.4.3.5 - Mesures générales d'incertitude d'un événement flou

Références.

CHAPITRE 3 : Principes d'extensions

- 3.1 - Transport d'une mesure d'incertitude par une application simple
- 3.2 - Cas d'une mesure de possibilité
- 3.3 - Relations floues et produit cartésien flou-noninteraction
- 3.4 - Fonction de possibilité définie sur un produit cartésien
- 3.5 - Transport d'une fonction de possibilité définie sur un produit cartésien
- 3.6 - Exemples d'utilisation du principe d'extension
 - 3.6.1 - Cardinalité floue d'un ensemble flou
 - 3.6.2 - Probabilité floue d'un événement flou
 - 3.6.3 - Compatibilité de deux ensembles flous
- 3.7 - Transport d'une mesure de possibilité par une fonction multivoque
 - 3.7.1 - Cas général
 - 3.7.2 - Inverse d'une application
 - 3.7.3 - Images inverses d'un ensemble flou par une fonction multivoque
- 3.8 - Transport d'une fonction de possibilité par une relation floue
 - 3.8.1 - Définitions
 - 3.8.2 - Intérêt de la composition des relations floues

Références.

CHAPITRE 4 : le calcul des intervalles flous

- 4.1 - Quantités, intervalles et nombres flous
 - 4.1.1 - Définitions
 - 4.1.2 - Interprétations
 - 4.1.3 - Valeur moyenne d'un intervalle flou
- 4.2 - Arithmétique floue non-interactive pessimiste
 - 4.2.1 - Résultat fondamental
 - 4.2.2 - Calcul pratique dans le cas général
 - 4.2.3 - Cas de quantités floues multimodales
 - 4.2.4 - Lien avec le calcul d'erreurs
 - 4.2.5 - Application aux opérations usuelles.
 - 4.2.5.1 - Fonctions d'une variable réelle
 - 4.2.5.2 - Extension des quatre opérations
 - 4.2.5.3 - Calcul pratique exact des 4 opérations

- 4.2.5.4 - Calcul approché des fonctions d'intervalles flous
- 4.2.5.5 - Extension du minimum et du maximum
- 4.2.6 - Le problème des représentations équivalentes d'une fonction
- 4.2.7 - Analogues possibilistes de la loi des grands nombres.
- 4.3 - Arithmétique floue pessimiste avec interaction des variables
 - 4.3.1 - Dépendance linéaire entre variables
 - 4.3.2 - Addition de nombres flous au sens d'une t-norme
- 4.4 - Comparaison d'intervalles flous
 - 4.4.1 - Comparaison d'un réel et d'une quantité floue
 - 4.4.2 - Comparaison de deux quantités floues
 - 4.4.3 - Classement de n quantités floues
 - 4.4.3.1 - Indices globaux de surclassement
 - 4.4.3.2 - Relations de surclassement floues
 - 4.4.4 - Lien avec la compatibilité de deux quantités floues
 - 4.4.5 - Domaines délimités par des bornes floues
- 4.5 - Arithmétique floue optimiste
 - 4.5.1 - Addition et soustraction de Minkowski
 - 4.5.2 - Equations d'intervalles flous
 - 4.5.3 - Solutions des équations d'intervalles flous
 - 4.5.4 - Equation $M \oplus X = N$
 - 4.5.5 - Equation $M \ominus X = N$
 - 4.5.6 - Autres équations d'intervalles flous
 - 4.5.7 - Autres types d'équations floues

Références

CHAPITRE 5 : Fonctions floues

- 5.1 - Quelques définitions possibles d'une fonction floue
- 5.2 - Construction de fonctions floues
 - 5.2.1 - Spécifications incomplètes de fonctions
 - 5.2.2 - Approximation d'une spécification linguistique par une fonction
 - 5.2.3 - Approximation d'une spécification linguistique par une relation floue
 - 5.2.4 - Régression linéaire floue
 - 5.2.5 - Lissage d'une application floue

5.3 - Eléments d'analyse floue

5.3.1 - Intégration de Riemann d'une application floue

5.3.2 - Dérivée d'une application floue

5.3.3 - Intégration sur un intervalle flou

5.4 - Variables aléatoires floues

Références

PARTIE II : Applications

CHAPITRE 6 : Modèles de processus décisionnels

6.1 - Modèles d'agrégation de critères décrits pour des ensembles flous

6.1.1 - Représentation d'un critère par un ensemble flou

6.1.2 - Ensemble flou et fonction d'utilité

6.1.3 - Méthodologie pour l'agrégation de critères

6.1.3.1 - Conditions minimales sur l'opération h

6.1.3.2 - Agrégation de deux critères

6.1.3.3 - Agrégation de n critères par combinaison symétriques des objectifs.

6.1.3.4 - Identification des opérations

6.1.3.5 - Critères d'importance inégale

6.1.3.6 - Liens avec d'autres approches

6.1.4 - Exemples d'application

6.2 - Aide à la décision en environnement incertain ou imprécisément perçu

6.2.1 - Une extension du concept d'utilité de Von Neumann et Morgenstern

6.2.2 - Environnement aléatoire imprécisément perçu

6.2.2.1 - Conséquences imprécises

6.2.2.2 - Probabilités imprécises des conséquences

6.2.2.3 - Cas mixte

6.2.2.4 - Anomalies engendrées par l'imprécision des probabilités

6.2.3 - Modèles non probabilistes d'environnement incertain

6.2.4 - Modèle possibiliste d'environnement incertain

6.3 - Classement d'actions potentielles imprécisément évaluées

6.3.1 - Comparaison d'évaluations imprécises

6.3.1.1 - Définitions d'ordres totaux sur A

6.3.1.2 - Méthodes de traitement de relations de surclassement floues

6.3.1.3 - Applications aux relations de surclassement d'intervalles flous

6.3.2 - Surclassement flou multicritère

Références

CHAPITRE 7 : L'intérêt des intervalles flous en Recherche Opérationnelle

7.1 - Problèmes de cheminement sur des graphes imprécisément valués

7.1.1 - Arithmétique floue optimiste ou pessimiste

7.1.2 - Calcul de chemins extrémaux dans un graphe imprécisément valué

7.1.3 - Application au problème central de l'ordonnancement

7.1.4 - Application à des problèmes d'ordonnancement à contraintes cumulatives

7.1.5 - Comparaisons avec les approches probabilistes.

7.2 - Problèmes d'optimisation en environnement imprécis.

7.2.1 - Formulation classique d'un problème d'optimisation en environnement imprécis

7.2.2 - Programmation linéaire avec contraintes souples

7.2.3 - Approche floue des problèmes linéaires multicritères

7.2.4 - Contraintes linéaires floues : approche par l'inclusion de Zadeh

7.2.5 - Contraintes linéaires floues : approche par les possibilités et les nécessités.

7.2.6 - Interprétation déterministe des contraintes floues

7.2.7 - Comparaison avec la programmation stochastique

7.2.8 - A problème d'optimisation floue, solution floue

7.2.9 - Un problème de réglage optimal d'un système de production.

7.3 - L'imprécision dans les modèles de files d'attente

7.3.1 - Modèle possibiliste de file d'attente

7.3.2 - Autres modèles stochastiques entâchés d'imprécision.

Références

AVANT-PROPOS

Ce mémoire constitue d'abord une synthèse de travaux précédemment publiés, dont une partie importante a été réalisée en collaboration avec Henri Prade. La parenté entre cette thèse et la sienne a d'ailleurs été volontairement soulignée au travers de la similarité de leurs titres. Dans les deux mémoires, on trouve comme idée directrice la formulation des concepts de possibilité et de nécessité, complémentaires de la notion de probabilité, et leur utilisation pour la représentation et le traitement d'informations imprécises et incertaines.

Néanmoins cette thèse se singularise par l'accent qu'elle met sur l'élucidation des liens et des différences entre ensemble flou et probabilité, la généralisation du calcul d'erreurs, et sur les méthodes d'agrégation d'informations imprécises. Notre contribution à une théorie de l'imprécision et de l'incertitude, actuellement en construction, concerne principalement les trois points suivants :

- . le calcul des intervalles flous et l'analyse des fonctions floues, qui en est un prolongement naturel (voir les chapitres 4 et 5).
- . l'étude de nouvelles mesures d'incertitude obtenues en généralisant l'axiome d'additivité des probabilités des événements disjoints. Cela permet de mieux situer ensemble flou et probabilité l'un par rapport à l'autre, et de définir des passages entre les deux représentations (voir le chapitre 1)
- . une approche axiomatique des opérations ensemblistes floues qui permet d'en dresser une nomenclature (voir le chapitre 2).

Nous présentons des applications du calcul d'intervalles flous à l'analyse de sensibilité dans les modèles de décision en environnement incertain, à l'étude des chemins critiques dans les graphes potentiels-tâches imprécis, à l'optimisation sous contraintes floues. Les opérations ensemblistes floues s'utilisent directement pour agréger des critères, et les mesures d'incertitude décrites au chapitre 1 suggèrent de nouvelles représentations d'environnement incertain, reflétant l'état des connaissances du décideur.

En plus d'une synthèse de travaux passés, ce mémoire contient l'esquisse de développements originaux qui pourront faire l'objet de recherches futures. En ce sens, on ne prétend pas ici fermer le sujet abordé, mais plutôt faire le point et indiquer de nouvelles voies de recherche.

On ne trouvera pas non plus un panorama complet de la théorie des ensembles flous et de ses applications⁽¹⁾.

Le mémoire est organisé comme suit : une longue introduction tente de motiver les recherches présentées, et résume le contenu des deux parties qui suivent. Elle se veut autonome, afin de fournir au lecteur une idée suffisamment précise du contenu et des enseignements que l'on peut tirer du travail présenté. Les deux autres parties, consacrées respectivement aux outils mathématiques et aux applications, constituent en quelque sorte le détail des arguments en faveur de la thèse. La partie mathématique contient cinq chapitres (N°1 à 5) qui suivent un ordre logique et dont le fil conducteur est la fonction de possibilité : mesure d'incertitude (chap.1), elle s'interprète en termes d'ensemble flou (chap.2), définit des variables floues (chap.3). Le calcul des fonctions de variable floue est explicité (chap.4) et débouche sur l'analyse (chap.5). La partie "applications" contient deux chapitres indépendants (n° 6 et 7) qui utilisent les principaux outils mathématiques présentés auparavant, sur des problèmes d'analyse de la décision et de recherche opérationnelle. La terminologie utilisée appelle quelques remarques. Au lieu de "sous-ensemble flou" et "valuation", termes employés par le Prof. Kaufmann, on a préféré parler d'ensemble flou (traduction de "fuzzy set") et de "mesure d'incertitude", sachant qu'un ensemble flou est un sous-ensemble et qu'une mesure d'incertitude n'est pas une mesure au sens de la théorie du même nom.

De plus, on trouvera des mots suspects tels que "possiblement" et "possibiliste". Il est étrange qu'en français, "possiblement" soit inusité, alors que "nécessairement" et "probablement" sont des vocables courants (en anglais "possibly" s'emploie). Cette dissymétrie d'usage a été volontairement négligée afin de préserver la dualité entre "possible" et "nécessaire", que l'emploi d'une périphrase voilerait. Quant à "possibiliste", il fait écho à "probabiliste".

(1) voir le livre de Dubois et Prade (Op. Cit. cf /11/ du chap.1)

INTRODUCTION
&
PRESENTATION DE LA THESE

Ce mélange subtil de savoir et de savoir-faire qu'on nomme "Sciences de l'Ingénieur" se trouve au confluent d'autres disciplines, réputées plus pures, qui vont des mathématiques et de la physique jusqu'aux sciences humaines, telles la psychologie ou la sociologie. Elles empruntent à celles-ci des connaissances et les exploitent à l'aide de ce nouveau-né surdoué de plus en plus envahissant qu'est l'informatique. L'informatique a bouleversé les sciences de l'ingénieur avant toutes les autres, c'est un lieu commun que de le dire. Mais l'avènement de l'ordinateur, en suscitant chez les ingénieurs l'ambition de résoudre des problèmes pratiques nouveaux, à partir de modèles de plus en plus complexes a précipité le besoin de recueillir des données toujours plus sophistiquées. On s'est aperçu qu'une grosse partie de cette information était souvent inaccessible sous la forme de nombres précis et définitifs. Que pour des raisons diverses - les instruments de mesure sont imparfaits, la mesure perturbe le phénomène, ou encore, et c'est souvent le cas, la seule source de renseignements est un être humain - cette information était imprécise, incohérente, incomplète en tout cas. De ce fait, l'élaboration de théories, de techniques de représentation, et de méthodes de traitement de l'imprécision et de l'incertitude y compris d'origine subjective, avec l'émergence de l'informatique, représente un enjeu important, et risque de conditionner en partie les progrès de cette dernière, et des disciplines qui en tirent parti.

Non seulement la précision est inaccessible à la mesure, mais encore il est des situations où elle s'avère peu utile, voire difficilement interprétable. C'est le cas lorsqu'on analyse le fonctionnement d'un système complexe, ou de grande dimension. Un modèle agrégé fournit parfois une information plus directement assimilable qu'un modèle très détaillé, et donc plus précis. Comme le souligne Zadeh : "à mesure que la complexité d'un système s'accroît, notre aptitude à formuler des affirmations précises, mais significatives, sur son comportement diminue jusqu'à un seuil au-delà duquel la précision et le sens deviennent mutuellement exclusifs" /15/. Ce principe d'incompatibilité est lié à la façon dont l'humain perçoit et raisonne ; il utilise principalement des représentations résumées de la réalité, et donc imprécises. Les techniques de simulation, par exemple, sont directement concernées par les propos de Zadeh : la puissance de l'ordinateur permet de simuler l'évolution de systèmes complexes, mais que faire des avalanches de chiffres qu'on est ainsi à même de produire, si on ne dispose pas de méthodes pour synthétiser

cette information, sous une forme assimilable par l'esprit humain ? On voit donc que l'imprécision est liée au recul que l'on peut prendre par rapport au système que l'on observe, et s'avère donc parfois bénéfique.

En fait un "bon modèle" doit réaliser un compromis évitant tout excès de précision (qui risque d'être arbitraire, donc incertaine) ou d'imprécision (qui peut devenir trop peu informative). Les propos qui précèdent ne prétendent pas critiquer la démarche scientifique : depuis la découverte des relations d'incertitude d'Heisenberg, l'imprécision irréductible est devenue un phénomène familier. Par ailleurs, la prise en compte de l'imprécision, même de celle qui provient des capacités limitées de l'entendement humain, n'exclut pas le souci de rigueur qui est, plus encore que la quête du précis, le propre de la démarche scientifique.

Imprécision, incertitude et concepts voisins

L'imprécision et l'incertitude peuvent être considérées comme deux points de vue antagonistes sur une même réalité qu'est l'imperfection de l'information. Si on représente une information sous la forme d'une proposition logique, l'imprécision réfère au contenu de cette proposition, l'incertitude réfère à sa vérité, entendue au sens de sa conformité à une réalité. Etant donné un corps de connaissances (un ensemble d'informations) l'antagonisme entre imprécision et incertitude s'exprime par le fait que si l'on rend plus précis le contenu d'une proposition, on tendra à accroître son incertitude. C'est une autre forme du principe d'incompatibilité énoncé plus haut.

Dans la langue française, il y a d'autres qualificatifs qui renvoient à l'imprécis, tels que "vague", "flou", "général", "ambigu". L'ambigu est une forme d'imprécision liée au langage, produite par effet d'homonymie parfois, une proposition est ambiguë dans la mesure où elle renvoie à plusieurs contextes ou référentiels possibles. Ce type d'imprécision n'est pas celui qui sera considéré dans ce mémoire : on suppose connu le référentiel où sont définies les propositions. Le "général" est une forme (bénéfique) d'imprécision liée au processus d'abstraction ; une proposition est générale si elle désigne un ensemble d'objets dont elle souligne une propriété commune. Mais le caractère vague, ou flou, d'une proposition réside dans l'absence de contour bien délimité de l'ensemble des objets qu'elle désigne. Beaucoup de qualificatifs du langage naturel sont flous, et présentent les caractères de

la généralité. A titre d'exemple on peut citer :

- une proposition imprécise non floue : $x = y$ à ϵ près
- une proposition imprécise floue : x est approximativement égal à y .

Le terme vague "approximativement" désigne plusieurs valeurs plus ou moins adéquates de ϵ .

Parallèlement, on juge de l'incertitude d'une proposition à l'aide de qualificatifs tels que "probable", "possible", "nécessaire", "plausible", "crédible" auxquels on va tenter de conférer un sens précis. La modalité du probable a été beaucoup étudiée depuis deux siècles. Elle a deux connotations différentes : l'une est physique, liée aux expériences statistiques et concerne la fréquence d'occurrence d'un événement. L'autre est épistémique : probable renvoie alors à un jugement subjectif. Les modalités du "possible" et du "nécessaire" datent d'Aristote, qui a souligné leur dualité (si un événement est nécessaire, c'est que son contraire est impossible). Curieusement, et contrastant avec le concept de probable, le possible et le nécessaire ont été souvent considérés comme des catégories de type tout ou rien. Mais, le possible a comme le probable deux interprétations : physique (mesure de la difficulté matérielle d'effectuer une action) et épistémique (jugement subjectif qui engage peu son auteur). Le nécessaire est au contraire une notion beaucoup plus affirmative, au sens physique, ou épistémique (la nécessité subjective est la certitude). Il est naturel d'admettre des degrés dans le possible et le nécessaire, comme dans le probable (la nuance se trouve déjà dans la langue puisqu'on peut dire "très possible", par exemple). Le plausible et le crédible ont des connotations spécifiquement épistémiques et renvoient respectivement au possible et au nécessaire. Chacun d'eux correspond à un mode de raisonnement à partir d'un corps de connaissance donné : le plausible est ce qui peut être induit du corps de connaissance, le crédible est ce qui en est déduit. La nuance apparaît dès que dans le corps de connaissance, certaines informations sont vagues ou elles-mêmes incertaines.

Des exemples de propositions incertaines sont :

- il est probable que Jean mesure au moins 1.70 m.
- il est vraisemblable (= plausible) qu'il tombe 10 mm d'eau demain.
- il est presque certain qu'il pleuvra beaucoup demain.

Il va de soi qu'une proposition peut être à la fois imprécise (ou vague) et incertaine, comme en témoigne la dernière proposition ci-dessus.

Il est facile de constater que l'imprécision d'une information entraîne l'incertitude de certaines des propositions qu'on peut énoncer à partir de cette connaissance. Mais le caractère incertain d'une information précise, oblige à conférer une certaine imprécision aux conclusions définitives qu'on en tirera. Cet antagonisme entre incertitude et imprécision a été modélisé par Zadeh /17/ dans ses travaux sur la représentation du sens dans les langages naturels (voir aussi Prade /6, pp.270-279/).

Modèles traditionnels de l'imprécision et de l'incertitude

On dispose traditionnellement de deux moyens de représenter l'imperfection des informations : la théorie des probabilités d'une part, et ce qu'on appelle le calcul d'erreurs d'autre part. Examinons brièvement les domaines d'utilisation de ces deux approches.

La théorie des probabilités est aujourd'hui une théorie mathématique respectable, dont les axiomes sont clairs et non contestés, et les développements très avancés. L'axiome fondamental est celui de l'additivité des probabilités d'événements disjoints. Les controverses relatives à la théorie des probabilités portent sur son interprétation : quelle réalité prétend-on traduire à l'aide de ce modèle mathématique ? Historiquement on s'en servait pour le "calcul des chances" dans les jeux de hasard notamment, la probabilité d'un événement étant définie par le rapport du nombre de cas favorables au nombre de cas possibles. Le caractère peu opérationnel de cette définition a donné naissance à l'école fréquentiste, qui considère les probabilités comme les valeurs asymptotiques des fréquences d'événements observés. La troisième école, dite subjectiviste, a tenté d'éviter les difficultés rencontrées par les fréquentistes pour appliquer la théorie (nombre d'observations suffisant, répétabilité des expériences ...) en suggérant les probabilités comme une mesure du sentiment d'incertitude. La valeur de probabilité est alors interprétée comme proportionnelle à la somme qu'un individu accepterait de payer au cas où la proposition qu'il affirme être vraie s'avèrerait fausse. Dans la mesure où cette somme existe, on montre que la mesure d'incertitude ainsi définie obéit aux axiomes de la théorie des probabilités, pourvu que le

comportement de l'individu satisfasse des conditions de "rationalité" (Savage /9/). A partir de là, les subjectivistes ont pu affirmer que les axiomes de Kolmogorov étaient les seuls raisonnables pour évaluer le sentiment d'incertitude. Cette attitude extrême peut être contestée, comme on le verra plus loin.

Quant au calcul d'erreurs, couramment pratiqué en physique, il ne fait que répercuter l'imprécision des appareils de mesure, exprimée sous forme d'intervalles, sur les grandeurs qu'on évalue à partir de ces mesures. Mathématiquement, on calcule l'image d'une application dont les arguments sont des sous-ensembles. Le calcul d'erreur est non nuancé : si on ignore la valeur exacte d'un paramètre, on connaît précisément les limites de son domaine de variation. On remarquera qu'étant donnée une mesure imprécise M de la grandeur X , les propositions du type " X appartient à l'intervalle I " seront naturellement qualifiées à l'aide des modalités du possible et du nécessaire car

- i) Si $M \cap I$ est non vide alors " $X \in I$ " est possible
- ii) Si $M \subseteq I$ alors " $X \in I$ " est nécessaire.

On met ici en évidence les liens entre ces modalités et la théorie des ensembles : le possible est évalué à partir d'une intersection ensembliste entre les contenus M et I des deux propositions " $X \in M$ ", " $X \in I$ "; le nécessaire est évalué à partir de l'inclusion ensembliste. Il est clair qu'en i) la proposition $X \in I$ est induite, tandis qu'en ii) elle est déduite.

Le caractère tout-ou-rien du calcul d'erreurs, alors que la théorie des probabilités est nuancée, introduit une dissymétrie entre eux, qu'on aimerait pouvoir gommer. Il est clair que la seconde ne généralise pas le premier, car une fonction de variables aléatoires à distribution uniforme (l'équivalent probabiliste de l'intervalle d'erreur) n'a pas généralement une distribution uniforme. Un des apports principaux de cette thèse est de proposer une généralisation canonique du calcul d'erreurs, qui soit nuancée.

Après cette brève analyse, les questions qui viennent à l'esprit sont : la théorie des probabilités est-elle le seul modèle mathématique de l'incertain qui soit défendable ? et aussi : Quelles réalités un calcul d'erreurs nuancé peut-il traduire ? et enfin : Quels liens entretiendraient ces deux théories ? C'est à ces questions que cette thèse tente de répondre,

en partie au moins. Mais d'abord, on discute deux hypothèses que nécessite l'utilisation du modèle probabiliste, et qui ne seront pas toujours vérifiées, notamment pour la représentation de jugements subjectifs incertains.

Remarques sur la structure de l'ensemble des événements réalisables

Une mesure de probabilité est construite sur un ensemble d'événements qui peuvent être vus comme l'ensemble des sous-ensembles d'un référentiel dit "événement certain". Ce qu'on a appelé précédemment proposition est une affirmation relative à l'occurrence d'un événement. Une autre notion fondamentale est celle d'épreuve qui est le résultat d'une expérience aléatoire destinée à mettre en évidence l'occurrence de certains événements, et la non-occurrence des autres. Une hypothèse capitale est celle que l'espace des épreuves peut être mis en bijection avec celui des événements : à tout événement on associe l'ensemble des épreuves qui le réalisent (il est non vide si l'événement n'est pas l'événement impossible), et pour chaque paire d'événement distincts, il existe au moins une épreuve qui réalise l'un à l'exclusion de l'autre. Cette dernière hypothèse permet de partitionner l'événement certain en événements élémentaires, chacun correspondant à une épreuve spécifique.

Examinons la validité de cette hypothèse en pratique. Dans le cas d'un recueil de données statistiques, cela revient à supposer qu'il existe toujours une partition non triviale de l'espace des réalisations telle que le résultat de chaque expérience puisse être affecté à l'un, et l'un seulement des éléments de cette partition. Autrement dit le résultat de chaque expérience est suffisamment précis pour caractériser la réalisation d'un événement élémentaire. Si on étudie la cohérence des épreuves successives, on constate qu'elles correspondent à des événements disjoints. Autrement dit, l'information recueillie est (suffisamment) précise mais dispersée. Le caractère aléatoire du phénomène réside d'ailleurs dans la fluctuation des observations successives (sinon on observe un phénomène déterministe).

Il est clair qu'on peut imaginer des exemples où cette hypothèse de partitionnement des épreuves ne tient pas. Si les résultats des expériences sont justement des intervalles d'erreurs, on aura en général peu de chance de pouvoir les affecter à des classes disjointes. Le physicien s'attend même à rencontrer la situation inverse ; à savoir une cohérence suffisante des intervalles d'erreurs provenant de mesures indépendantes, pour pouvoir, par recou-

pement, diminuer l'erreur de mesure !

On voit donc que même à propos de phénomènes répétitifs "objectifs" la théorie des probabilités ne peut être appliquée telle quelle. Le modèle probabiliste s'avère judicieux pour prendre en compte des informations précises et dispersées. Dès qu'on ne sait plus définir a priori un ensemble exhaustif d'événements élémentaires disjoints qui formeront l'événement certain, on sort de l'épure.

Considérons maintenant la notion de probabilité subjective. L'hypothèse fondamentale analysée ci-dessus, à savoir la structure de σ -algèbre de l'ensemble des événements réalisables doit rester valide. Plus encore que dans le cadre des applications de type statistiques, cette hypothèse peut être critiquée dans l'optique subjectiviste dont l'ambition est de pouvoir s'appliquer à des domaines tels que l'économie, le management, etc... Les critiques ne sont pas nouvelles (par exemple Shackle /10/) et sont fondées sur le fait qu'il est très difficile d'identifier les conséquences possibles (et disjointes) d'une décision aux événements élémentaires d'un espace de probabilité, en raison de la complexité des situations rencontrées.

Un exemple limite, mais révélateur est celui de la représentation de l'ignorance totale. De par la nécessité de devoir fractionner l'intervalle unité en poids de probabilité portant sur les événements élémentaires, aucune mesure de probabilité n'est susceptible de représenter l'ignorance totale. Cet état peut se rencontrer au début d'un processus d'apprentissage. On le modélise en général par une répartition uniforme de poids de probabilité sur les événements élémentaires. Il est bien connu qu'une telle distribution, même si elle est, par rapport à toutes les autres mesures de probabilité, d'entropie maximale, représente encore de l'information puisque, au sens de cette mesure, certains événements seront plus probables que d'autres. L'ignorance totale, définie comme l'état de connaissance où tous les événements imaginables sont également incertains, échappe donc au modèle probabiliste, et ne nécessite pas l'identification d'une σ -algèbre non triviale d'événements réalisables. On peut même ignorer quels sont les événements qui sont réalisables (puisque on n'a pas encore fait d'épreuves !). En particulier une fonction d'ensemble qui représente l'ignorance totale doit être insensible au

fractionnement d'événements jugés jusqu'alors élémentaires. La condition de normalisation des probabilités est un obstacle à cette insensibilité.

L'axiome d'additivité dans l'approche subjectiviste

Il est bien connu que l'axiome d'additivité des probabilités d'événements disjoints reflète l'additivité des fréquences d'occurrence de ces événements.

Quand on considère des probabilités qui modélisent un jugement subjectif d'incertitude, cette propriété d'additivité est plus difficile à interpréter. La seule justification que l'on connaisse est celle de la situation de pari évoquée plus haut. Il semble difficile de soutenir que tout jugement incertain correspond à une situation de pari. En particulier, les subjectivistes qualifient d'irrationnel tout individu qui, d'accord pour payer une somme proportionnelle à $P(A)$ si l'événement A ne se réalise pas, refuse de s'engager proportionnellement à $P(\bar{A}) = 1 - P(A)$ au cas où c'est A qui advient.

Ce qui n'est pas clair dans cette interprétation, c'est la manière dont les sommes engagées dans le pari sont reliées à l'état des connaissances du parieur. Si c'est un parieur chevronné, on conçoit facilement qu'il répartisse la somme en jeu, soit proportionnellement aux fréquences d'occurrence des événements concernés telles qu'elles ont été perçues par lui dans le passé, soit d'après ce que lui suggère une analyse a priori du phénomène sur le nombre de cas favorables. Néanmoins, le modèle probabiliste ne permet pas de distinguer l'attitude d'un parieur confirmé qui joue la même somme sur les deux événements contraires parce qu'il sait qu'ils ont toutes chances de se produire aussi fréquemment l'un que l'autre, de celle d'un néophyte qui agit de même parce que c'est la stratégie la plus prudente quand on ignore tout. Ce que voudrait pouvoir exprimer ce dernier, c'est que, pour lui tous les paris sont - risque financier mis à part - également envisageables. On aimerait disposer d'une méthode de représentation de l'incertitude qui reflète l'état des connaissances du parieur et puisse ainsi distinguer entre le néophyte et l'habitué. Si on recueille les degrés de croyance en mettant l'individu devant une situation de pari, ils seront biaisés par le sentiment de risque financier qui l'emporte alors sur celui du manque de connaissance :

avouer qu'on sait trop peu de choses devient un acte coûteux.

On peut au contraire admettre que le parieur choisisse lui-même la valeur du total des sommes, qu'il jouera sur chaque événement, entre 0 et un maximum convenu. Cela revient à supposer que les degrés de croyance $G(A)$ et $G(\bar{A})$ vérifient $G(A) + G(\bar{A}) \leq 1$. La quantité $1-G(A)-G(\bar{A})$ est à la mesure du manque de connaissance de l'individu. On peut interpréter $G(A)$ comme la plus petite somme qu'il soit justifiable de parier sur A , et $1-G(\bar{A})$ la plus grande de ces sommes, au sens d'un parieur expérimenté (pour qui $G(A)+G(\bar{A})=1$). Le terme "justifiable" s'entend ici non en termes de risque financier, mais au sens de "qui reflète l'état de la connaissance". Dans ce cadre, l'ignorance totale amène à poser $G(A) = G(\bar{A}) = 0$; en effet cela signifie que tous les paris ($P(A)$, $P(\bar{A})=1-P(A)$) sont envisageables. On note au passage que ce type de degré de croyance rend compte du fait que les indices que l'on possède en faveur d'un événement ne sont pas liés à ceux que l'on possède en faveur de son contraire. Autrement dit on peut très bien douter de l'occurrence d'un événement et de celle de l'événement contraire (faute d'indices). On voit que $G(A)$ est interprétable en termes de degré de certitude (ou de nécessité, de crédibilité) de A . $G(A) = G(\bar{A}) = 0$ correspond à l'absence totale de certitude. On notera que l'école subjectiviste n'échappe pas à sa propre critique de l'école fréquentiste quand à la nécessité de répéter les épreuves, puisque le modèle probabiliste est justifié surtout pour un individu qui a l'habitude de parier !

Un autre problème est celui du niveau de précision des probabilités subjectives ainsi obtenues. Dire "la probabilité (subjective) de A est $P(A)$ " est certainement en pratique équivalent à la proposition imprécise "la probabilité de A est $P(A) \pm \epsilon$ " pour un ϵ suffisamment petit mais fini. L'école subjectiviste semble faire fi de cette imprécision, qui rend sujette à caution, voire difficile à interpréter la formule $P(A)+P(\bar{A})=1$. Un dernier point de litige concerne le problème de changement d'échelle : s'il est clair que, sous les hypothèses classiques, les sommes mise en jeu sur deux événements disjoints s'ajoutent en vue d'un pari sur la réunion de ces événements, il est faux de croire que le seul modèle de fonction d'ensemble correspondant soit le modèle probabiliste : on peut très bien faire un changement d'échelle non linéaire qui transforme le cumul des sommes pariées en une autre façon de combiner les probabilités.

Donc l'approche par les probabilités subjectives se heurte à trois problèmes cités par ordre décroissant d'importance :

- Elle ne rend compte que d'un éventail limité d'attitudes de l'individu devant l'incertitude
- Elle suppose que les poids de probabilité émis par l'individu sont précis
- Le modèle additif n'est pas le seul possible même dans l'hypothèse du pari.

Un peu d'histoire

Cette mise en question de l'opportunité des probabilités additives dans les modèles de jugement incertain n'est en fait pas nouvelle. Dans un remarquable article d'histoire des sciences, Shafer /12/ rappelle que jusqu'à la fin du XVII^{ème} siècle les concepts de chance (liée au hasard) et de probabilité (attribut de l'opinion) sont perçus indépendamment l'un de l'autre. Il indique que le lien entre ces notions semble avoir été pour la première fois suggéré par Arnauld et Nicole dans leur ouvrage "L'art de penser", en 1662. A l'origine la théorie des chances et celle des "probabilités" ont été développées séparément, ces dernières sans l'axiome d'additivité. Shafer /12/ a montré qu'une partie non négligeable des travaux de Jacques Bernoulli traitent des probabilités non-additives. Néanmoins, ce n'est pas cette partie qui a attiré l'attention de ses successeurs, lesquels auraient un peu vite assimilé chance et probabilité. C'est la loi des grands nombres que l'histoire a retenu, et c'est un théorème sur le comportement du hasard plutôt que sur les degrés de croyance. Au XVIII^{ème} siècle, le seul chercheur ayant persisté à distinguer chance et probabilité et à développer une théorie non-additive semble avoir été Lambert dont Shafer montre que les travaux, bien que précurseurs des siens (cf. chap.1), étaient déjà inintelligibles pour la communauté scientifique de son époque. Le succès des probabilités additives est aussi lié à l'essor des sciences physiques qui a relégué au second plan les problèmes de modélisation du jugement. Quand ces préoccupations, au XX^{ème} siècle, sont réapparues les premiers travaux auxquels elles ont donné lieu portent la marque du préjugé en faveur de l'additivité des degrés de croyance. Avec les progrès de l'informatique, et le développement des techniques d'intelligence artificielle, le besoin de disposer d'une théorie du jugement incertain qui déborde le cadre probabiliste s'est fait soudain plus pressant comme en témoignent des systèmes experts tels que MYCIN (Shortliffe & Buchanan /13/). Les récents travaux de

chercheurs tels que Shackle, Zadeh, Dempster, Shafer, etc... auxquels cette thèse doit beaucoup, semblent avoir renoué avec l'intuition qui guidait la démarche de Bernoulli et Lambert. Ils tentent non seulement de généraliser l'axiome d'additivité, mais aussi évitent de recourir à la nécessité de disposer de la liste exhaustive des événements élémentaires a priori.

Après avoir tenté de motiver la recherche de nouveaux modèles de l'imprécision et de l'incertitude, on va maintenant résumer les résultats contenus dans ce mémoire, en suivant l'exposé des chapitres, afin d'en guider la lecture.

Fonctions de possibilité et de nécessité (chap.1)

Le premier chapitre de ce mémoire tente d'explorer diverses approches possibles qui sortent du cadre probabiliste pour la représentation de l'incertitude. On définit d'abord une collection minimale de conditions que doit satisfaire une fonction d'ensemble à valeurs sur l'intervalle unité pour être utilisable dans cette perspective. La propriété fondamentale est la monotonie au sens de l'inclusion ; elle exprime que le degré de croyance en un événement est au moins aussi grand que le degré de croyance en ceux qui l'impliquent. A partir de là, on voit que le degré de croyance en l'union de deux événements quelconques est inférieurement limité par le plus grand des degrés de croyance en chacun d'eux. Si on refuse de dépasser cette borne inférieure, on atteint un cas limite de ces fonctions d'ensemble, appelées par Zadeh /16/ mesures (ou fonctions) de possibilité Π telle que

$$\Pi(A \text{ ou } B) = \max(\Pi(A), \Pi(B)) \quad \forall A, B$$

Il est facile de voir que ce type de fonction d'ensemble parvient à capturer la modalité du possible. L'affirmation la plus prudente, relative à l'occurrence d'un événement, est bien de dire qu'il est possible ; c'est cohérent avec l'axiome ci-dessus, qui propose un degré minimum de croyance pour juger de l'événement A ou B. On vérifie que l'une des deux quantités $\Pi(A)$ et $\Pi(\bar{A})$ (\bar{A} est le contraire de A) vaut 1, ce qui est en accord avec l'idée que l'un au moins de deux événements contraires est possible, et parfois les deux. Par ailleurs l'axiome des fonctions de possibilité est conforme à l'interprétation physique de cette modalité, puisqu'il dit que pour réaliser l'événement A ou B, il suffit de réaliser le moins difficile des deux.

La situation d'ignorance totale est représentable par de telles fonctions, si on pose pour tout événement A non vide $\Pi(A) = 1$. La modalité duale (nécessité) est capturée par une fonction d'ensemble notée N dont la donnée équivaut à celle d'une fonction Π , la relation de dualité se traduisant par l'égalité $N(A) = 1 - \Pi(\bar{A})$. On remarque que pour posséder toute l'information relative à un événement A , il faut connaître les deux valeurs $\Pi(A)$ et $\Pi(\bar{A})$ (ce dernier exprime le doute sur l'occurrence de A) qui ne se déduisent pas forcément l'une de l'autre.

L'axiome des fonctions de possibilité permet de calculer le degré de croyance d'un événement en fonction de ceux d'événements qui l'impliquent. Cette propriété constructive permet de définir une fonction de possibilité à l'aide d'une distribution de poids sur une partition en événements élémentaires de l'événement certain. Mais ces poids ne sont plus liés par la contrainte probabiliste de normalisation. Ce qui fait que l'adjonction de nouveaux événements élémentaires peut n'avoir aucune influence sur les degrés de possibilité déjà alloués aux autres. C'est cette distribution de poids qui définit ce que Zadeh a nommé ensemble flou (fuzzy set).

On notera que la précision des degrés de possibilité est moins cruciale que celle des degrés de probabilité. En effet l'axiome des fonctions de possibilité implique qu'on n'opère que des comparaisons (ce qui est prudent quand on a affaire à des quantités fournies par un individu). L'erreur qu'on peut commettre au cours de manipulations de fonctions de possibilité ne dépassera jamais, comme l'a remarqué Prade /7/ récemment, celle qu'on a commise au départ. Il n'en est pas de même si on utilise des manipulations de type probabiliste où interviennent sommes et produits.

De nouvelles classes de mesures d'incertitude (chap.1)

A part les cas limites que représentent les fonctions de possibilité et de nécessité, il est possible de définir des classes plus riches de fonctions d'ensemble monotones qui puissent servir de support numérique aux jugements d'incertitude. On considère principalement deux approches : la première, récemment étudiée par Prade et l'auteur analyse les substituts possibles de l'axiome d'additivité ; la seconde, initialisée par Dempster /1/ et reprise par Shafer /11/, ne suppose plus qu'une application existe de l'ensemble des épreuves vers celui des événements élémentaires, mais seulement une correspondance.

La première approche identifie les fonctions d'ensemble dites décomposables, au sens où elles peuvent être définies par la donnée d'une opération qui combine les degrés de croyance d'événements disjoints, et d'une allocation de poids sur l'ensemble des événements élémentaires. Elle a le mérite de mettre en évidence l'existence de deux sous-classes de fonctions d'ensembles, disjointes, aux mesures de Dirac près :

- celles dont la valeur pour un événement détermine celle de l'événement contraire. Elles sont isomorphes aux mesures de probabilité, et sont pour cela appelées pseudo-probabilités. Ce sont des probabilités distordues.
- celles dont les valeurs pour un événement et son contraire ne sont pas liées. Leurs archétypes sont les fonctions de possibilité et de nécessité. Certaines sont isomorphes à des mesures non finies.

Cette approche exploite les solutions, trouvées indépendamment par des mathématiciens, d'équations fonctionnelles sur les réels. (Cf. Annexe).

La seconde classe de fonctions d'ensemble présentée est obtenue en ne supposant plus réalisables certains événements élémentaires. Les poids de probabilité sont affectés à un ensemble quelconque d'événements, disjoints ou non, qui forme le corps de connaissance. Cette information est équivalente à la donnée de l'une quelconque de deux fonctions d'ensembles notées P^* et P_* , duales au sens où $P^*(A) = 1 - P_*(\bar{A})$, respectivement sous-additives et sur-additives. Ces fonctions délimitent un ensemble de mesures de probabilité compatibles avec le corps de connaissance. C'est pourquoi Dempster les nomme probabilités supérieures et inférieures respectivement ($P^*(A) \geq P_*(A)$). $P^*(A)$ (resp : $P_*(A)$) est induit (resp : déduit) du corps de connaissances, en sommant les poids de probabilité des événements qui sont compatibles avec (resp : impliquent) A. C'est pourquoi Shafer les nomme fonctions de plausibilité et de crédibilité (belief) respectivement. Les événements formant le corps de connaissances sont appelés "éléments focaux".

Si le corps de connaissance est formé d'événements élémentaires (disjoints) seulement, alors $P^* = P_*$ est une mesure de probabilité. Si au contraire le corps de connaissance est formé d'une chaîne implicative d'événements (ensembles emboîtés) alors P^* et P_* sont des fonctions de possibilité et de nécessité respectivement. Cette approche place les fonctions de possibilité aux antipodes des mesures de probabilité : les premières traduisent une information imprécise (les événements sur lesquels portent les poids de probabilité ne sont pour l'essentiel pas élémentaires) mais complètement

cohérente, tandis que les secondes sont le reflet d'informations précises mais dispersées. On notera que les fonctions de possibilité sont en ce sens plus naturelles pour la représentation du sentiment d'incertitude : d'un individu on n'attend pas forcément une information très précise, mais on souhaitera la plus grande cohérence possible dans ses propos. En revanche des données précises et incohérentes sont plus couramment issues de l'observation d'un phénomène physique. Dans le cas général, les probabilités supérieures et inférieures sont construites sur un corps de connaissances à la fois imprécises et plus ou moins incohérentes.

Parmi les points forts de l'approche de Shafer, on peut citer l'extension canonique du concept d'espérance mathématique qui éclate en espérances inférieures et supérieures. De plus ce modèle semble non seulement pertinent pour représenter le jugement incertain, mais encore peut avoir des applications en statistique. Son point faible est la nécessité de spécifier, pour représenter une probabilité supérieure ou inférieure, autant de poids de probabilité qu'il y a d'événements imaginables. C'est l'avantage des fonctions d'ensemble décomposables - dont certaines, autres que les probabilités, possibilité ou nécessités, sont des fonctions de Shafer - de pouvoir être complètement définies par leurs restrictions à l'ensemble des événements élémentaires.

Une autre question abordée par le premier chapitre concerne le passage éventuel entre une représentation probabiliste et une représentation possibiliste d'un même corps de connaissance. Il est clair que pour être simultanément acceptées, ces deux représentations doivent vérifier certaines relations de compatibilité. En particulier un événement est d'abord possible avant d'être probable, et probable avant d'être nécessaire. D'où une inégalité entre les degrés de possibilité, probabilité et nécessité issus d'un même corps de connaissance. Dans le cas fini, une bijection entre les mesures de probabilité et les fonctions de possibilité est construite, respectant cette inégalité, et fondée sur l'hypothèse que la nécessité d'un événement par rapport à un autre reflète l'excès de probabilité du premier par rapport au second. Cette transformation, intuitive dans le sens probabilité \rightarrow possibilité, a pour inverse le calcul d'une mixture de mesures de probabilité à distribution uniforme, définies sur chaque élément focal, ce qui lui confère un caractère intuitif.

Le premier chapitre se clôt sur l'esquisse d'une théorie qualitative des possibilités à l'instar des probabilités qualitatives définies par Savage, dont on montre que les fonctions (numériques) de possibilité ne sont pas un modèle.

La suite de la thèse traite essentiellement de notions basées sur les fonctions de possibilité et de leurs applications.

La théorie des ensembles flous (chap.2)

Suggéré en 1965 par Zadeh, le concept d'ensemble flou /14/ veut rendre compte du fait qu'un objet peut plus ou moins correspondre à une certaine catégorie à laquelle on voudrait l'assimiler. D'où l'idée de graduer l'appartenance sur une échelle numérique, l'intervalle unité, ce qui permet de généraliser la fonction caractéristique d'un ensemble. Bien que justifié par une intuition très forte, ce concept a connu des fortunes diverses, soulevant l'enthousiasme des uns, éveillant la méfiance des autres. Parmi les raisons qui ont motivé une certaine incompréhension, l'une des plus importantes est l'impression qu'il pouvait s'agir de probabilités travesties. Et ce d'autant plus qu'un examen de la littérature sur le sujet révèle l'existence de multiples points de vue parfois contradictoires, ou peu justifiés.

Le concept d'ensemble flou ne pourra faire l'unanimité que lorsque seront éclaircis les rapports qu'il entretient avec celui de probabilité. C'est pourquoi dans cette thèse on a préféré étudier d'abord les mesures d'incertitude avec des hypothèses suffisamment affaiblies pour fournir un cadre commun aux probabilités et aux ensembles flous. Ces derniers, comme Zadeh l'a lui-même suggéré en 1978 /16/, sont très liés aux fonctions de possibilité et de nécessité, rencontrées plus haut comme cas limites de mesures d'incertitude. Pour le voir il suffit de remarquer qu'une fonction de possibilité tout-ou-rien, définie sur un ensemble, en caractérise de façon unique un sous-ensemble non vide, et réciproquement. De même la distribution de poids qui caractérise, en général, une fonction de possibilité, est interprétable comme la fonction d'appartenance d'un ensemble flou. Il est important de noter qu'une distribution de probabilité n'est pas susceptible d'une telle interprétation. En effet les mesures tout ou rien de probabilité (les fonctions de Dirac) caractérisent seulement les singltons du référentiel. Donc une mesure de probabilité doit être vue comme un "point flou" et non un ensemble flou.

Le chapitre 2 tente de faire le point sur la notion d'ensemble flou, sous l'angle ensembliste proprement dit. On énumère diverses représentations possibles d'un ensemble flou : la fonction d'appartenance, une suite d'ensembles emboîtés et pondérés, notamment. Puis on discute brièvement les techniques qui permettent de déterminer des degrés d'appartenance. Cette question, elle-aussi, est à l'origine du sentiment de méfiance à l'égard des ensembles flous. A notre avis, la réponse dépend du type de situation que l'on veut modéliser :

- soit le résultat d'épreuves statistiques : dans ce cas on peut estimer une fonction de possibilité si les épreuves fournissent des informations imprécises qui ont un minimum de cohérence. On peut aussi affaiblir une mesure de probabilité par la transformation évoquée au chapitre 1, section 6.1.
- soit une mesure subjective d'incertitude, pour la représentation de catégories vagues du langage naturel. Zadeh a suggéré l'utilisation de fonctions de possibilité dans ce but /17/. Dans une phrase telle que "Jean est grand" le terme "grand" est vu comme un ensemble flou de tailles, qui définit une distribution de possibilité sur la taille de Jean. Le choix d'une mesure de possibilité (plutôt que de probabilité par exemple) répond au souci d'utiliser la moins arbitraire possible des représentations, donc la plus prudente. On a vu que parmi les fonctions d'ensembles monotones, les fonctions de possibilité satisfaisaient ce critère. Une fois ceci accepté, l'estimation des degrés d'appartenance relève des techniques de la psychométrie - tout au moins de celles qui ne contiennent pas d'hypothèses probabilistes.

Lorsqu'on a reconnu un ensemble dans une mesure de possibilité, l'étape suivante est d'étudier la structure algébrique que forment ces ensembles. Cette question rejoint certaines préoccupations de chercheurs tels que Lukasiewicz et Tarski, qui, entre les deux guerres, ont étudié les logiques multivalentes. En effet les logiques multivalentes sont aux ensembles flous ce que la logique binaire est à la théorie des ensembles. Le problème abordé dans le chapitre 2 est celui de l'extension multivalente des tables de vérité, ou, de manière équivalente, la définition des opérations ensemblistes floues à l'aide de fonctions de valuation. Grâce à la théorie des équations fonctionnelles, ce problème est en grande partie résolu au chapitre 2 : on pose comme axiomes certaines propriétés des treillis booléens, et on résoud les équations auxquelles elles aboutissent. On peut tirer de cette analyse

les enseignements suivants :

- la structure de treillis booléen est incompatible avec l'appartenance nuancée d'un élément à un ensemble
 - on retrouve comme cas particulier des modèles de calculs logiques multivalents proposés dans le passé, comme la "logique du probable" de Reichenbach, mais aussi certains systèmes proposés par Lukasiewicz (cf. Rescher /8/).
 - les opérations "max" et "min" proposées par Zadeh pour faire l'union et l'intersection des ensembles flous ne sont pas arbitraires : elles dotent les ensembles flous de la structure la plus riche possible (celle de treillis vectoriel)
 - il y a des opérations ensemblistes floues qui n'ont pas de contrepartie classique tout en ayant une interprétation intuitive : il en est ainsi des moyennes, et des opérations invariantes par transformation de De Morgan.
 - il y a deux façons (au moins) de concevoir l'inclusion des ensembles flous : l'une qui se fonde sur l'opération "non A ou B", l'autre qui suppose a priori que l'inclusion s'exprime par une inégalité des fonctions d'appartenance. Elles correspondent à deux familles d'implications multivalentes ayant en commun celle suggérée par Lukasiewicz.
- La fin du chapitre répertorie différentes évaluations scalaires d'un ensemble flou. Ces évaluations permettent de déterminer à quel point un ensemble flou possède telle ou telle propriété, à savoir :
- à quel point les frontières d'un ensemble flou sont-elles mal localisées (indice de flou)
 - quel est le niveau de précision d'un ensemble flou (on définit l'équivalent de l'entropie probabiliste, bien que dans la littérature des ensembles flous on nomme "entropie" un indice de flou).
 - qu'est-ce que le cardinal d'un ensemble flou ? (un nombre réel)
 - à quel point un ensemble flou est-il vide (ou au contraire remplit-il tout le référentiel) ? ce dernier type d'évaluation débouche sur des indices de comparaison entre ensembles flous.
 - quel est le degré d'incertitude d'un événement flou ? (probabilité, possibilité, nécessité, ...). Le concept d'événement flou s'introduit naturellement

dans la mesure où un événement est représenté sous la forme d'un ensemble. Ce dernier point s'avère très important pour la représentation des propositions qui sont à la fois vagues (d'où l'événement flou) et incertaines.

Le principe d'extension (chap.3)

Une fonction de possibilité est en général associée à une variable dont elle délimite le domaine. On l'appelle variable possibiliste ou variable floue. Ce qu'on appelle principe d'extension permet d'étudier les fonctions de variables floues. C'est le sujet du chapitre 3. Connaissant une mesure de possibilité Π sur un ensemble Ω , une fonction f , de Ω vers un autre ensemble U , induit sur U une mesure de possibilité Π_f . Quand Π est de type tout ou rien, elle revient à la donnée d'un sous-ensemble A de Ω , et Π_f est la mesure de possibilité associée à l'ensemble $f(A)$, image de A par f . En ce sens, on peut s'attendre à ce que l'étude des fonctions de variables floues généralise le calcul d'erreurs. Quand la fonction f a plus d'un argument, on est amené à manipuler la distribution de possibilité jointe, qu'on appelle encore relation floue. Dans l'étude des fonctions de variables aléatoires, on fait intervenir la notion d'indépendance, ce qui permet de simplifier les calculs. Dans le cas de variables floues on introduit la "non-interaction". Le domaine alloué à un ensemble de variables floues non-interactives est un produit cartésien d'ensembles flous. Cette remarque indique que la "non-interaction" des variables possibilistes correspond à l'absence de liaison fonctionnelle ou relationnelle entre variables et n'a que peu de rapport avec l'indépendance stochastique. Afin de montrer l'intérêt du principe d'extension pour forger de nouveaux concepts, on donne deux illustrations :

- le cardinal flou d'un ensemble flou : on construit ainsi des entiers flous, mesures de possibilité sur les entiers. Par extension on peut donner un sens à une probabilité d'un événement flou, qui soit une quantité floue.

- la compatibilité d'un ensemble flou par rapport à un autre. C'est une généralisation de la valeur d'appartenance d'un élément à un ensemble. Cela permet de visualiser de façon synthétique, sous la forme d'un ensemble flou sur le segment unité la comparaison point par point de deux ensembles flous. La quantité floue ainsi obtenue peut s'interpréter comme la valeur de vérité d'une proposition vague par rapport à une autre qui représente la réalité. Le chapitre 3 se clôt sur la question du transport d'une mesure de possibilité

par une correspondance, ou une relation floue, et signale les liens entre cette question et des problèmes d'algèbre linéaire dans les dioïdes /2/.

Le calcul des intervalles flous (chapitre 4)

Une partie importante de ce mémoire est consacré à l'étude des quantités floues, et surtout de celles qui généralisent la notion d'intervalle. Ces travaux ont fait l'objet de publications dont les résultats sont synthétisés dans le chapitre 4. En appliquant le principe d'extension aux fonctions de variables réelles, on obtient un calcul des intervalles flous qui est une généralisation du calcul d'erreur. On prouve que ce nouveau calcul se ramène à de l'analyse de sensibilité, effectuée en parallèle sur une infinité d'intervalles contenant les arguments d'une fonction. Néanmoins, connaissant deux intervalles flous M et N , le calcul de leur image $f(M,N)$ par une fonction f s'avère simple à mener dans de nombreux cas, en tout cas beaucoup plus simple que sa contrepartie probabiliste.

On s'intéresse particulièrement à l'extension des quatre opérations de l'arithmétique à des intervalles flous. On montre que si deux intervalles flous se déduisent l'un de l'autre par une translation et une affinité orthogonale, alors la somme de ces deux nombres flous peut se déduire de l'un d'eux par les mêmes transformations. Ce résultat permet à l'aide d'une représentation paramétrée générale d'un intervalle flou de réduire le calcul des quatre opérations sur de tels intervalles à une simple combinaison arithmétique des paramètres. Dans le cas général on propose des méthodes de calcul approché. L'extension des opérations minimum et maximum se fait très simplement, et s'interprète très naturellement. On étudie les structures algébriques de l'ensemble des intervalles flous, muni des diverses opérations étendues. Enfin on fait remarquer que si on connaît deux expressions équivalentes d'une fonction, leurs extensions à des intervalles flous ne sont plus forcément équivalentes.

Les résultats ci-dessus supposent la non-interaction des variables mises en jeu. La suite du chapitre montre des cas d'interaction où on est encore capable de calculer aisément le résultat d'opérations arithmétiques. On considère notamment le cas où l'interaction se traduit par une contrainte linéaire et où l'on veut calculer une somme pondérée. On traite aussi le cas où la relation floue entre des variables à additionner est une forme affaiblie de produit cartésien d'ensembles flous.

L'arithmétique des intervalles flous, basée sur le principe d'extension, est appelée pessimiste car elle permet, connaissant les domaines flous que parcourent des variables x, y, \dots de calculer le plus grand domaine flou que puisse parcourir la variable $f(x, y, \dots)$. On introduit dualement une arithmétique optimiste qui, essentiellement, recherche le plus petit domaine possible pour la variable $f(x, y, \dots)$. Si l'arithmétique floue pessimiste ne suppose aucune compensation des erreurs, la contrepartie optimiste impose une compensation maximale des erreurs. Cette notion, qui généralise la soustraction d'ensembles de Minkowski, s'avère utile pour la résolution des équations d'intervalles flous.

Une section du chapitre 4 est consacrée à la comparaison des intervalles flous, en vue d'étendre l'analyse des positions relatives de deux intervalles. On fournit un ensemble canonique et complet d'indices de comparaison dans le cadre de la théorie des possibilités. Ces indices permettent de définir des relations floues de préordre entre intervalles flous dont les propriétés facilitent la définition de procédures naturelles de classement.

Fonctions floues (chap. 5)

Une fois acquises certaines propriétés des quantités floues, généralisant celles des nombres, l'étape suivante est l'étude des fonctions floues en vue de définir des éléments d'analyse. Il y a plus d'une façon de concevoir une fonction floue : on en connaît principalement quatre :

- une fonction à domaines de départ et d'arrivée flous
- une relation floue entre deux ensembles
- un ensemble flou de fonctions entre deux ensembles
- une fonction qui à un ensemble flou fait correspondre un autre ensemble flou.

C'est le second et le troisième point de vue qui retiennent notre attention. En effet la première définition n'est pas assez générale, et la quatrième l'est trop. Alors que dans le cas classique, une correspondance est équivalente à un faisceau de fonctions, ce n'est plus le cas pour les relations floues. Une relation floue peut être considérée comme une fonction à valeurs floues, et on peut construire l'ensemble flou des fonctions qui sont compatibles avec une relation floue. Néanmoins on ne peut généralement pas, connaissant cet ensemble flou de fonctions reconstituer la relation floue, sauf pour

certaines fonctions floues appelées applications floues fortes.

Avant d'étudier l'analyse des correspondances floues, on tente de préciser les situations dans lesquelles on peut être amené à définir et utiliser des fonctions floues. Il y en a principalement deux. Une fonction floue peut représenter un ensemble de règles logiques approchées, fournies sous forme linguistique par un individu. Une méthode pour synthétiser une fonction floue à partir de ce type de données imprécises est proposée. Une fonction floue peut aussi fournir une représentation à la fois synthétique et fidèle d'un "nuage de points"; des techniques de régression linéaire floues sont présentées, qui mènent à une représentation paramétrée de la relation floue ainsi construite.

La section 3 résume une série de trois publications relatives à l'intégration (de Riemann) des correspondances floues, l'intégration d'une fonction réelle sur un intervalle flou, et la différentiation. On montre qu'on peu étendre assez facilement des résultats dûs à Aumann sur la convexité de l'intégrale par exemple. De plus la linéarité de l'intégration est préservée, au sens de l'addition d'intervalles flous. L'analyse des fonctions floues peut être un point de départ indispensable pour l'étude des variables aléatoires floues, qui peuvent synthétiser des données statistiques imprécises (telles que des résultats de sondages d'opinion par exemple).

Applications de la théorie des possibilités et des ensembles flous

Le propre des fonctions de possibilité et des ensembles flous semble être de saisir des informations imprécises mais cohérentes ; c'est là qu'à notre avis se situe leur différence majeure avec les mesures de probabilité dont le domaine d'utilisation concerne les phénomènes précisément observables mais fluctuants. Cet aspect rend les fonctions de possibilité particulièrement aptes à prendre en compte le type d'informations que fournit un individu quand il décrit l'état de ses connaissances, quand il spécifie un cahier des charges, ou encore quand il explique sa manière de procéder pour résoudre un problème.

Les outils mathématiques présentés dans cette thèse peuvent donc être précieux dans le cadre de disciplines dont l'informatique a favorisé

le développement et qui cherchent à exploiter le savoir et l'expérience des individus quand ils résolvent des problèmes ou opèrent des choix, afin de leur allier la puissance de calcul de l'ordinateur. On pense ici en priorité aux méthodes d'analyse de décision et à l'intelligence artificielle.

Concernant ce dernier domaine, la thèse de Prade /7/ a démontré l'intérêt des fonctions de possibilité pour représenter certains modes de raisonnement naturel, qui sont à la base de ce que l'on appelle les systèmes-experts, et aussi pour l'intégration d'informations imprécises ou incomplètes dans les bases de données, y compris les aspects de communication homme-machine.

Les deux derniers chapitres tentent de préciser la manière dont on peut utiliser les opérations ensemblistes et l'arithmétique floues dans les problèmes d'analyse de la décision et de recherche opérationnelle.

Application à l'analyse de la décision (chap.6)

Les théories de la décision se trouvent confrontées à trois problèmes : la modélisation des préférences, la représentation de l'incertitude, et l'agrégation des critères. Le chapitre 6 s'intéresse principalement aux deux derniers problèmes, mais traite aussi de méthodes de classement pertinentes pour le premier.

- Agrégation de critères

Un objectif peut être vu comme un ensemble flou de bonnes décisions. Cette remarque ne débouche sur une méthode d'aide à la décision que si l'objectif en question est représentable dans un espace dessiné par les attributs que l'on souhaite contrôler. Ainsi, la recherche d'une automobile bon marché et qui consomme peu, sera visualisé dans l'espace coût d'achat-consommation kilométrique sous la forme d'un domaine dans lequel on souhaite se placer.

Ce domaine n'a en général pas de frontières précises, mais on peut penser l'appréhender en exprimant pour chaque attribut la zone dans laquelle il doit se trouver. Cette zone peut s'exprimer sous la forme d'un ensemble flou qui est par exemple la traduction numérique d'un qualificatif linguistique désignant l'objectif partiel (par exemple "bon marché"). En fait

l'objectif global forme une relation floue dans l'espace des attributs, dont les objectifs partiels sont en quelque sorte, les projections. L'approche proposée pour agréger les critères cherche à reconstituer cette relation floue à partir de ces objectifs partiels. L'étude extensive des opérations ensemblistes floues faite au chapitre 2 trouve là une utilisation naturelle. L'identification d'une opération d'agrégation est possible en faisant remplir par le décideur une table de vérité "qui cristallise la façon dont il relie les objectifs les uns aux autres. En particulier, on ne suppose pas a priori que l'on recherche un compromis entre les objectifs ; les liens étroits qu'entretient cette approche avec la logique permettent de prendre en compte la conjonction d'objectifs ou leur redondance éventuelle⁽¹⁾. La question de l'importance relative des critères est abordée. On distingue les cas où un objectif partiel est plus pointu qu'un autre, plus pertinent qu'un autre, et celui où il y a une dissymétrie structurelle dans la façon dont les deux objectifs sont agrégés.

Le caractère pointu d'un objectif est exprimé au travers de la forme de sa fonction d'appartenance, la dissymétrie de l'agrégation est reflétée par la structure de l'opération choisie pour lier les objectifs. La pertinence des objectifs est exprimée à l'aide d'une mesure d'incertitude sur l'ensemble des critères. Dans ce cas, l'évaluation globale peut être représentée par la mesure d'incertitude d'un événement flou. Le modèle classique d'agrégation de critères d'importance inégales à l'aide d'une combinaison convexe des évaluations correspond au choix d'une mesure d'incertitude probabiliste.

Ce type d'approche permet de représenter le sens de concepts qui mettent en jeu un certain nombre d'attributs physiques observables. Elle est donc utile dans le cadre de modèles de raisonnement approché pour réaliser l'interface entre la description microscopique d'une situation et les catégories macroscopiques que l'individu utilise pour raisonner à partir de cette situation ; ces catégories traduisent la fonction "résumante" de l'esprit humain face à un donné complexe. En revanche ces méthodes ne sont pas adaptées à l'agrégation de critères d'un niveau d'abstraction trop élevé : dans ce dernier cas un modèle de raisonnement, de type système-expert, semble plus adapté.

(1) La méthode n'est pas sans lien avec les techniques de synthèse de circuits logiques fondées sur l'utilisation des "tables de Karnaugh"

- Décision en environnement incertain

L'existence de représentations non-probabilistes de l'incertitude, et les fonctions de possibilité permettent d'introduire l'imprécision dans les modèles classiques de décision en environnement incertain, et aussi d'imaginer de nouveaux modèles.

Une analyse du concept d'utilité au sens de Von Neuman et Morgenstern indique que la composition des probabilités le long d'un chemin, dans un arbre de décision procède d'une hypothèse probabiliste d'indépendance. Dans ce mémoire on propose une esquisse de ce que devient la théorie de Von Neumann, si on omet cette hypothèse, et qu'on ne présuppose rien sur l'opération de composition des probabilités le long d'un chemin dans un arbre de décision. On se contente de poser quelques axiomes qui paraissent naturels, mais qui s'avèrent caractériser la forme de l'opération de composition. On peut ainsi généraliser la formule de l'espérance mathématique de l'utilité.

Pour prendre en compte l'imprécision dans les valeurs de probabilité traduisant l'incertitude du décideur quant aux conséquences de sa décision, on suggère d'utiliser des probabilités linguistiques, représentables par des fonctions de possibilité sur l'intervalle unité. On donne des méthodes simples pour évaluer l'espérance mathématique de l'utilité, qui devient, en général, un intervalle flou.

Si on n'admet plus que les degrés d'incertitude fournis par l'individu, sont, stricto-sensu, des probabilités, on peut montrer que, dans le cadre d'une interprétation de type "pari", le degré d'incertitude de l'union de deux événements incompatibles doit se situer dans un intervalle dont les bornes sont l'évaluation possibiliste et l'évaluation probabiliste. Un certain nombre de fonctions d'incertitude décomposables, étudiées au chapitre 1, satisfont ce critère. Elles ont l'avantage de pouvoir être définies à l'aide d'une distribution de poids sur l'ensemble des conséquences prévisibles, mais sans que la contrainte de normalisation soit a priori fixée. Ceci devrait donner plus de facilité d'expression à l'individu qui décrit l'état de ses connaissances. Cette approche n'est qu'à un stade préliminaire de développement, et on prétend ici seulement donner quelques éléments susceptibles de motiver une investigation théorique et pratique plus complète, au cours de recherches futures.

- Classement de décisions imprécisément valuées

Une des raisons qui motivent l'introduction d'ensembles flous en théorie de la décision, est que de nombreux modèles d'analyse de la décision présupposent l'existence de meilleures décisions que l'analyse permettra de déterminer. C'est le cas de la plupart des techniques basées sur de l'information cardinale. D'autres modèles plus récents ont admis que certaines décisions pouvaient se révéler impossible à départager. Il en est ainsi de certaines approches multicritères qui refusent la possibilité d'agréger numériquement des évaluations de nature différente. C'est aussi vrai pour celles qui tentent de classer des objets à partir d'une relation de préférence obtenue par comparaison deux à deux des objets. L'introduction des ensembles flous dans ce domaine veut souligner qu'une des causes possibles d'incomparabilité réside dans l'imprécision de l'information dont on dispose pour évaluer les diverses alternatives. Dans ce cas, un des rôles du modèle est, plutôt que de fournir un classement arbitraire, de détecter ce type de situation. Autrement dit, au lieu de chercher quelle est la meilleure décision, on essaie de déterminer s'il est licite d'en proposer une comme étant la meilleure.

En pratique l'imprécision se traduit naturellement par des évaluations floues ; l'imprécision peut avoir son origine dans la méconnaissance de la valeur des attributs de l'objet que l'on considère, dans le choix d'une opération d'agrégation, dans la description de l'incertitude quant aux conséquences d'une décision. Le classement d'objets imprécisément évalués revient au problème de comparaison d'intervalles flous traité au chapitre 4. On définit, à partir des évaluations imprécises, plusieurs relations floues de crédibilité de surclassement, et on parvient à l'aide de techniques adéquates, à construire plusieurs classements dont la confrontation permet de mieux juger des positions respectives obtenues par les divers objets. Cette approche n'est pas sans lien avec certaines méthodes d'arbitrage de relations de préférence obtenues directement par questionnaires.

- Applications en recherche opérationnelle (chap.7)

Si l'apport des mesures d'incertitude et des fonctions de possibilité paraît de nature à susciter de nouveaux modèles d'analyse de la décision, on ne prétendra pas ici que les ensembles flous sont à même de révolutionner le paradigme de la recherche opérationnelle. Ce qu'on voudrait suggérer, néanmoins, c'est que l'arithmétique des intervalles flous peut, dans

certaines situations, prendre le relai de la théorie des probabilités. C'est notamment vrai lorsque l'utilisation de cette dernière mène à formuler des problèmes trop difficiles à résoudre, ou s'avère difficile à justifier faute de données statistiques suffisantes. L'emploi de l'arithmétique floue ne se justifiera que si elle fournit des résultats dont l'imprécision ne soit pas trop grande. Elle correspond toujours à une forme élaborée d'analyse de sensibilité.

Dans le chapitre 7, on propose deux exemples où le calcul des intervalles flous peut dans des situations pratiques donner naissance à des modèles de complexité raisonnable (beaucoup plus simples que leurs contreparties probabilistes pour un même niveau de généralité), et d'un certain pouvoir informatif : la détermination de chemins extrémaux dans un graphe imprécisément valué ; l'optimisation (notamment la programmation linéaire) avec des contraintes floues.

Pour ce qui est du premier exemple on a les résultats suivants :

- les algorithmes de plus courts chemins en environnement déterministe s'étendent directement au cas de l'environnement flou, et restent simples à mettre en oeuvre (contrairement aux versions probabilistes).
- il n'y a plus de chemins extrémaux à proprement parler, néanmoins on peut définir des degrés de criticité des tâches et des chemins, dans un graphe potentiel - tâche décrivant les diverses étapes d'un projet, par exemple.

Pour ce qui est de la programmation linéaire, on prouve qu'une contrainte linéaire dont les coefficients (y compris le second membre) sont des intervalles flous se ramène à un ensemble de deux contraintes linéaires une fois fixé un paramètre qui mesure le degré de satisfaction de la contrainte. Ce degré de satisfaction est une fonction de possibilité ou une fonction de nécessité selon le sens que l'on attribue à la contrainte floue. Le problème de la maximisation simultanée des degrés de satisfaction des contraintes se ramène à une suite (rapidement convergente) de programmes linéaires. Des éléments de comparaison entre "programmation floue" et programmation stochastique sont fournis.

Afin d'illustrer l'intérêt de l'arithmétique floue pour certains problèmes d'optimisation, une étude de cas, menée par l'auteur, relative

à un problème de réglage de machines-outils est présentée brièvement. On y montre notamment l'intérêt de laisser imprécise la solution d'un problème d'optimisation floue.

Pour éviter de donner à entendre au lecteur que l'arithmétique floue peut se substituer au calcul des probabilités dans toutes les situations, on esquisse brièvement l'analyse d'une file d'attente où la durée de service au guichet, et l'intervalle de temps entre deux arrivées sont représentées par des fonctions de possibilité. On peut constater que l'information qu'on en déduit est très pauvre, et ne justifie guère l'emploi d'un tel modèle. Néanmoins on peut utiliser de manière conjointe la théorie des probabilités et celle des possibilités, pour modéliser des phénomènes tels qu'un processus de Poisson de taux d'arrivée imprécis, etc...

Conclusion : les ensembles flous existent-ils ?

L'abondance des articles et rapports se réclamant des ensembles flous depuis 1965 amène l'observateur à poser deux questions :

- s'agit-il effectivement d'une nouvelle théorie, au sens mathématique du terme
- de nouvelles méthodes pour modéliser le réel et résoudre des problèmes concrets ont-elles vu le jour ?

Pour ce qui est de la première question, l'auteur, qui n'est pas mathématicien même s'il lui arrive d'utiliser des mathématiques, se sent quelque peu embarrassé pour répondre. Dans un récent article de critique, - Manes /5/ oppose du point de vue de leurs statuts épistémologiques la théorie des probabilités - une "vraie" théorie selon lui - à celle des ensembles flous, qui ne ferait que s'afficher comme telle. On peut admettre de qualifier du nom de théorie une démarche conceptuelle et un ensemble de résultats en découlant, seulement lorsque cette démarche a fait la preuve de son originalité en se situant par rapport aux autres théories existantes et a reçu un développement suffisant pour lui assurer des fondements stables. Selon ces critères on reconnaîtra volontiers que la théorie des ensembles flous se cherche encore une identité. En effet le besoin d'une théorie non probabiliste de l'imprécision et de l'incertitude a surgi chez des ingénieurs plutôt que chez les mathématiciens, et ce n'est que récemment que certains parmi ces derniers ont commencé à étudier sous l'angle théorique, les concepts suggérés par l'intuition

de praticiens. On veut penser - et ce mémoire, tout imparfait qu'il soit voudrait contribuer à étayer cette thèse - que ces concepts peuvent donner matière à des développements mathématiques nouveaux. Les ensembles flous semblent avoir une place bien distincte de celle des probabilités au sein d'une théorie générale des mesures d'incertitude (encore balbutiante), qui entretiendrait des liens étroits avec les logiques multivalentes, et permettrait l'extension de certains concepts en théorie des ensembles (voir les travaux réalisés à l'université de Lyon) et en topologie (par exemple les articles de Lowen /4/ ; dans /3/, Höhle montre qu'un filtre est un cas particulier de mesure d'incertitude).

On peut souligner une analogie de statut entre la théorie des graphes et celle des ensembles flous. En particulier, les points suivants méritent d'être soulignés.

. De même qu'un graphe n'est qu'une relation binaire sur un ensemble (généralement fini), un ensemble flou n'est qu'une fonction d'un ensemble dans un autre ensemble (généralement le segment unité). En ce sens les mathématiciens peuvent considérer comme suspecte la nouveauté de ces concepts, a priori.

. Les graphes se sont révélés un outil commode et précieux pour la modélisation de structures dans les problèmes d'organisation, les réseaux de transports etc... La possibilité d'exprimer un graphe à l'aide d'un dessin a joué un rôle primordial dans la diffusion et l'acceptation du concept. Les ensembles flous semblent attirer plus facilement les praticiens que les théoriciens, grâce à leur pouvoir évocateur. L'intervalle flou est une représentation simple et riche d'une quantité imprécise, telle qu'elle est perçue par un individu. On peut lui demander de fournir une donnée imprécise en esquissant le profil d'une fonction d'appartenance. Bien sûr, l'ensemble flou synthétise la sémantique d'une information alors qu'un graphe en saisit la structure, et cela fixe certaines limites à la comparaison.

. La théorie des graphes a permis de résoudre des problèmes concrets avec des méthodes simples découvertes à cette occasion. Dans ce mémoire on cherche à montrer que les calculs qu'occasionnent l'utilisation des ensembles flous ne sont pas prohibitifs : c'est notamment le cas avec l'arithmétique floue. La théorie de Zadeh a permis de poser et de résoudre des problèmes qu'on ne savait pas - ou mal - résoudre auparavant : dans le domaine de l'intelligence

artificielle, Prade /7/ a démontré l'utilité des ensembles flous pour les systèmes experts et les bases de donnée. Cette thèse tente de souligner leur apport dans les modèles d'analyse de la décision et indique que certains problèmes de recherche opérationnelle en environnement imprécis peuvent être commodément formulés et résolus.

. Le concept de graphe a prouvé sa fécondité sur le plan mathématique, et il est probable que sans lui (avec la possibilité de visualisation qu'il suppose) de nombreux résultats n'auraient pu être découverts par pure spéculation abstraite. De même le concept d'ensemble flou semble fournir l'occasion d'un réexamen des fondements de la théorie des probabilité, et cette analyse peut déboucher sur certaines généralisations des deux concepts, comme il a été suggéré plus haut. A un niveau plus modeste, l'arithmétique floue généralise le calcul d'erreurs, et une meilleure compréhension de la façon dont on peut agréger des informations imprécises a été obtenue, grâce à l'utilisation conjointe des opérations ensemblistes et des résultats de la théorie des équations fonctionnelles.

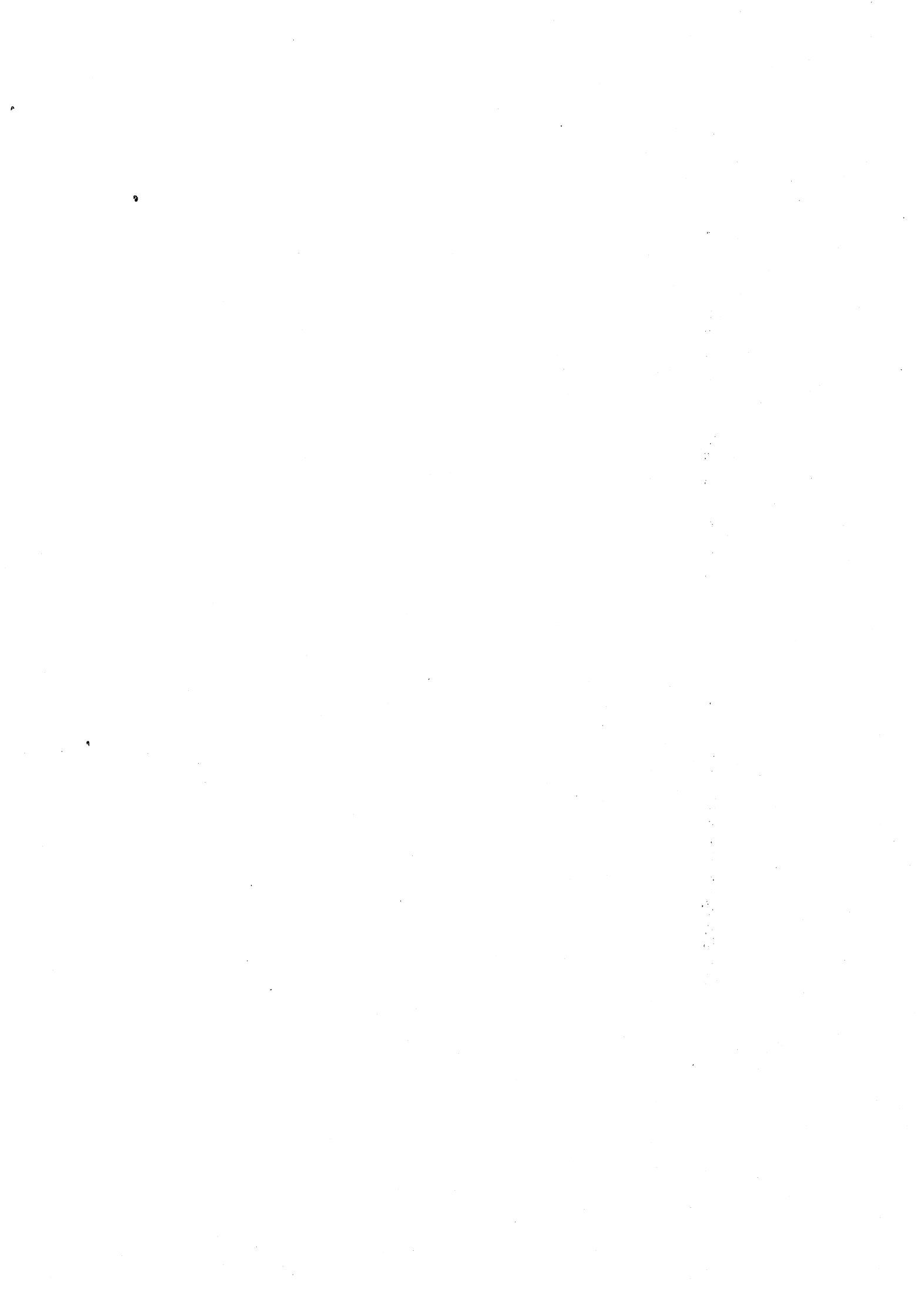
Le danger des concepts nouveaux est de provoquer des épidémies de généralisations. Dans la littérature des ensembles flous, les constructions gratuites ne manquent pas, qui risquent de fourvoyer le néophyte. Néanmoins il ne semble pas souhaitable de vouloir contrôler la prolifération des constructions formelles, car seul le temps permet de les juger, et ce n'est pas parce que la théorie des hypergraphes semble demeurer sans grand écho qu'on a eu tort de l'élaborer. Un des problèmes de la théorie des ensembles flous est de répondre, dans la pratique, en fournissant de nouvelles méthodes de résolution de problèmes réels, à l'attente qu'une philosophie séduisante peut susciter. Ce souci a guidé l'auteur de cette thèse, et on a cherché à ce que les mots se concrétisent en outils mathématiques et algorithmiques, dans la mesure du possible.

Il est clair que le travail nécessaire pour situer définitivement l'apport des concepts imaginés par Zadeh et étudiés par un nombre croissant de chercheurs à sa suite, aux mathématiques elles-mêmes et à des disciplines touchant à l'informatique, l'automatique et la recherche opérationnelle, est loin d'être terminé. Néanmoins de nombreux signes laissent penser que les ensembles flous conserveront une place importante dans les modèles de l'imprécis et de l'incertain qui seront élaborés dans le futur.

REFERENCES

-
- /1/ A.P. DEMPSTER (1967)
Op. cit. cf /4/ du chapitre 1
- /2/ GONDRAN M., MINOUX M. (1977)
Valeurs propres et vecteurs propres dans les dioïdes et leur interprétation en théorie des graphes. Bulletin de la direction des Etudes et recherches EDF - Série C n°2, 25-41.
- /3/ HÖHLE U. (1982)
Op. cit. cf. /33/ du chapitre 2
- /4/ LOWEN R. (1983)
Hyperspaces of fuzzy sets, Fuzzy Sets and Systems 9, 287-311.
- /5/ MANES E.G. (1982)
Critique du livre "Fuzzy Sets and Systems : theory and applications" par D. Dubois et H. Prade - Bulletin of the American Mathematical Society - Vol.7 n°3, 603-612.
- /6/ PRADE H. (1982)
Modèles mathématiques de l'imprécis et de l'incertain en vue d'applications au raisonnement naturel, Thèse d'Etat Université Paul Sabatier - Toulouse.
- /7/ PRADE H. (1983)
BUSEFAL N°14, p.127.
- /8/ RESCHER N. (1969)
Many-valued logics. Mc Graw-Hill
- /9/ SAVAGE L.J. (1972)
Op. Cit. Cf. /26/ du chap. 1
- /10/ SHACKLE G.L.S. (1961)
Op. cit. cf /28/ du chap. 1
- /11/ SHAFER G. (1976)
Op. cit. cf. /29/ du chap. 1

- /12/ SHAFER G. (1978)
Non-additive probabilities in the works of Bernoulli and Lambert, Archives for the History of Exact Sciences 19, 309-370
- /13/ SHORTLIFFE E.H., BUCHANAN B.G. (1975)
A model of inexact reasoning in medicine. Mathematical Biosciences 23, 351-379
- /14/ ZADEH L.A. (1965)
Op. cit. cf. /81/ du chap. 2
- /15/ ZADEH L.A. (1973)
Op. cit. cf /34/ du chap. 3
- /16/ ZADEH L.A. (1978)
Op. cit. cf /34/ du chap. 1
- /17/ ZADEH L.A. (1978)
Op. cit. cf /39/ du chap. 3.



PARTIE I

OUTILS MATHÉMATIQUES

CHAPITRE 1

MESURES D'INCERTITUDE

*"S'il y a un sens du réel,
il doit aussi y avoir un sens du possible"*

Robert Musil, "l'homme sans qualités".

Cette section passe en revue un certain nombre de fonctions d'ensemble à valeur dans le segment unité $[0,1]$ qui sont des modèles possibles pour représenter la connaissance que l'on possède concernant l'occurrence ou la non-occurrence d'un événement. A partir des diverses approches présentées on verra apparaître à chaque fois, comme cas particuliers limites les mesures de probabilité d'une part, et les fonctions de possibilité de Zadeh /34/ d'autre part.

1.1 - MESURES MONOTONES DE SUGENO

Considérons un ensemble d'événements que l'on assimile aux sous-ensembles d'un ensemble Ω dit "événement toujours certain". Sauf précision contraire Ω est supposé fini. Soit g une fonction qui à chaque événement $A \subseteq \Omega$ fait correspondre un nombre $g(A) \in [0,1]$. $g(A)$ mesure la croyance que possède un individu quant à l'occurrence de l'événement A , et est d'autant plus grand que cette croyance augmente, par convention. En particulier si A est un événement certain $g(A) = 1$, si A est un événement impossible $g(A) = 0$. Les réciproques de ces deux propositions ne sont pas forcément valides. On a en particulier :

$$g(\Omega) = 1 \quad ; \quad g(\emptyset) = 0 \quad (1)$$

où \emptyset dénote l'événement toujours impossible.

Si on sait qu'un événement A implique un événement B , ce que l'on note $A \subseteq B$, alors on croit toujours au moins autant en l'occurrence de B qu'en celle de A , ce qui justifie l'axiome général des mesures d'incertitude

$$A \subseteq B \rightarrow g(A) \leq g(B) \quad \forall A, B \subseteq \Omega \quad (2)$$

Les fonctions d'ensemble satisfaisant (1) et (2) ont été étudiées par Sugeno /30/ sous le nom de "mesures floues". Lorsque Ω n'est pas fini, on doit rajouter un axiome de continuité, de nature technique, à savoir : pour toute suite monotone $\{A_i\}_i$ d'événements

$$\lim_{i \rightarrow \infty} g(A_i) = g(\lim_{i \rightarrow \infty} A_i) \quad (3)$$

1.2 - CAS PARTICULIERS : POSSIBILITE, NECESSITE, PROBABILITE

Une conséquence immédiate de (1) et (2) est

$$\forall A, B \quad g(A \cap B) \leq \min(g(A), g(B)) \quad (4)$$

$$g(A \cup B) \geq \max(g(A), g(B)) \quad (5)$$

où l'occurrence de $A \cap B$ (resp $A \cup B$) signifie l'occurrence simultanée de A et B (resp, d'au moins l'un des deux événements).

On trouve donc naturellement comme cas limite de mesures d'incertitude des fonctions notées Π telles que

$$\forall A, B \quad \Pi(A \cup B) = \max(\Pi(A), \Pi(B)) \quad (6)$$

L'axiome (6) correspond à une déduction très prudente puisque $\Pi(A \cup B)$ est une borne inférieure sur la mesure d'incertitude de $A \cup B$ connaissant $\Pi(A)$ et $\Pi(B)$. On construit facilement une fonction Π à partir de la connaissance d'un événement certain E en posant

$$\begin{aligned} \Pi_E(A) &= 1 & \text{si } A \cap E \neq \emptyset \\ &= 0 & \text{sinon} \end{aligned} \quad (7)$$

Sachant que E est certain, $\Pi_E(A) = 1$ signifie clairement A est possible. D'où l'interprétation naturelle des fonctions Π comme fonctions de possibilité. L'axiome (6) ainsi que son interprétation ont été suggérés par Zadeh /34/.

Soit $\omega_1 \dots \omega_n$ les éléments de Ω (ou événements élémentaires) ; de (6) on déduit facilement

$$\forall A \subseteq \Omega \quad \Pi(A) = \max_{\omega_i \in A} \Pi(\{\omega_i\}) \quad (8)$$

Donc toute fonction de possibilité sur Ω est caractérisée par un ensemble de valeurs $\pi_i = \Pi(\{\omega_i\})$, $i = 1, n$, appelé distribution de possibilité. Si $\pi_i \in \{0, 1\}$ on voit facilement que cette distribution n'est autre que la fonction caractéristique μ_E de l'ensemble E qui apparaît en (7), i.e.

$$\mu_E(\omega_i) = \pi_i, \forall i \iff \Pi = \Pi_E \quad (9)$$

Plus généralement si $\pi_i \in [0,1] \forall i$, cette distribution correspond à un ensemble dont les limites sont imprécises, c'est-à-dire un ensemble flou. On notera que $\Pi(\Omega) = 1$ implique

$$\exists \omega_i, \pi_i = 1 \quad (10)$$

Enfin (8) est cohérent avec l'idée de "réalisabilité" d'un événement, puisqu'on cherche le plus facile à réaliser parmi les événements élémentaires qui impliquent A.

NB : Dans le cas où Ω est infini l'axiome (6) est généralisé par

$$\forall \{A_i | i \in I\}, \quad \Pi\left(\bigcup_{i \in I} A_i\right) = \sup_{i \in I} \Pi(A_i). \quad (11)$$

Comme signalé par Puri et Ralescu /25/, si on définit Π à l'aide d'une distribution de possibilité $\pi : \Omega \rightarrow [0,1]$ par

$$\Pi(A) = \sup_{\omega \in A} \pi(\omega) \quad (12)$$

Π n'est généralement pas une mesure floue au sens strict parce que la continuité est violée pour des suites $(A_i)_i$ d'ensembles décroissantes au sens de l'inclusion. De même (11) n'implique pas forcément l'existence de π satisfaisant (12). Néanmoins les fonctions Π définies à partir de (12) seront encore appelées fonction de possibilité.

L'autre cas limite de "mesure floue" est obtenu à partir de l'équation (4). Il s'agit des fonctions N telles que

$$\forall A, \forall B, \quad N(A \cap B) = \min(N(A), N(B)) \quad (13)$$

Cet axiome correspond à une attitude risquée puisqu'en recoupant deux informations A et B, on affecte au résultat la valeur de confiance la plus élevée possible. On construit facilement une fonction N à partir de la connaissance d'un événement certain E en posant

$$\begin{aligned} N_E(A) &= 1 && \text{si } E \subseteq A \\ &= 0 && \text{sinon} \end{aligned} \quad (14)$$

$N_E(A) = 1$ signifie clairement que A est certain. De plus, on voit facilement que N vérifie (13) si et seulement si la fonction d'ensemble Π définie par

$$\forall A, \quad \Pi(A) = 1 - N(\bar{A}) \quad (15)$$

où \bar{A} est le complémentaire de A, vérifie (6), c'est-à-dire est une fonction de possibilité. (14) et (15) justifient parfaitement pour les fonctions N le qualificatif de "nécessité". (15) est une transcription numérique d'une relation de dualité entre les catégories "possible" et "nécessaire" bien connue en logique modale (Hugues et Creswell, /17/). Ici "nécessaire" s'entend bien sûr au sens d'"inévitabile" et non d'"indispensable". D'un point de vue subjectif, "possible" et "nécessaire" renvoient respectivement aux concepts de "vraisemblance" et de "certitude". La relation (15) a pour conséquence le fait que dans le cas où Ω est fini, N s'exprime en fonction de la même distribution de possibilité π par:

$$N(A) = \inf_{\omega_i \in \bar{A}} 1 - \pi_i \quad (16)$$

Les fonctions de nécessité ont été suggérées indépendamment de la théorie des ensembles flous par l'économiste anglais G.L.S. Shackle /28/ qui les utilisait en remplacement des probabilités subjectives (Savage /26/). Shackle appelait "degré de surprise potentielle" de A la quantité $N(\bar{A})$ (cf Dubois /5/). Les fonctions de possibilité et de nécessité satisfont les relations remarquables :

$$\min(N(A), N(\bar{A})) = 0 \quad (17)$$

$$\max(\Pi(A), \Pi(\bar{A})) = 1 \quad (18)$$

(17) exprime que deux événements contraires ne peuvent être certains simultanément, mais peuvent être incertains tous deux. (18) exprime que ces événements peuvent être simultanément possibles, mais qu'un seul des deux peut être impossible à la fois.

Quand la connaissance que l'on possède sur l'occurrence d'événements est disponible sous la forme de fréquences mesurées, la fonction d'ensemble P construite sur ces fréquences satisfait naturellement l'axiome d'additivité :

1.3 - MESURES D'INCERTITUDE DECOMPOSABLES

Un point commun entre possibilité et probabilité est que, connaissant l'évaluation portant sur des événements mutuellement exclusifs, on peut calculer l'évaluation portant sur l'union de tels événements. Cette propriété de décomposabilité d'une fonction d'ensemble peut s'exprimer par l'axiome suivant :

Il existe une opération interne $*$ sur le segment unité telle que g vérifie :

$$\forall A, B, A \cap B = \emptyset \rightarrow g(A \cap B) = g(A) * g(B) \quad (*)$$

Dans le paragraphe précédent on a rencontré les cas $*$ = + (probabilité) et $*$ = max (possibilité). Une investigation systématique des mesures d'incertitude décomposables a été menée à bien dans Prade /24/ Dubois /7/.

On donne ici un résumé des résultats exposés dans la version définitive (Dubois & Prade, /10/).

1.3.1 - Caractérisation de l'opérateur de décomposition

On vérifie aisément que l'opération $*$ doit satisfaire les conditions suivantes : notons $a = g(A)$, $b = g(B)$, $c = g(C)$, $A, B, C \subseteq \Omega$.

- associativité $a * (b * c) = (a * b) * c$ (compatibilité avec l'union ensembliste)
- commutativité $a * b = b * a$ (idem)
- $a * 0 = a$ (car $A \cup \emptyset = A$)
- $a \geq b \rightarrow a * c \geq b * c$ (mesure monotone, cf. (2))

Donc $*$ est une opération de semi-groupe commutatif de $[0,1]$, non décroissante selon chaque argument, et d'élément neutre 0.

NB : Schwyhla /27/ a fait remarquer que puisque Ω est fini, l'ensemble des valeurs $\{g(A) \mid A \subseteq \Omega\}$ est fini. Donc la connaissance des propriétés de $*$ sur un nombre fini d'éléments de $[0,1]$ ne suffit pas à caractériser cette opération. On ne peut conclure que si on considère toutes les fonctions g satisfaisant (*). C'est ce qui est implicite ici. De toute façon si on ne considère qu'une seule mesure floue décomposable selon $*$, les conditions de cohérence permettent de conclure à l'existence d'un semi-groupe ad hoc

$$\forall A, \forall B, \quad A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B) \quad (19)$$

i.e. est une mesure de probabilité, qui bien sûr, est monotone. (19) est l'équivalent probabiliste de (6) et (13). On notera qu'en vertu de (4), il est inutile d'introduire la condition $A \cap B = \emptyset$ dans (6). De même dans (13) la condition duale $A \cup B = \Omega$ est sans objet. L'équivalent de (8) et (16) est

$$P(A) = \sum_{\omega_i \in A} P(\{\omega_i\}) \stackrel{\Delta}{=} \sum_{\omega_i \in A} p_i \quad (20)$$

où les p_i satisfont $\sum_{i=1}^n p_i = 1$, condition analogue à (10). Enfin le pendant de (17) et (18) est:

$$P(A) + P(\bar{A}) = 1 \quad (21)$$

tandis que (16-17) implique

$$N(A) + N(\bar{A}) \leq 1 \quad (22)$$

$$\Pi(A) + \Pi(\bar{A}) \geq 1 \quad (23)$$

On perçoit ici une des différences majeures entre possibilité et probabilité : la probabilité d'un événement détermine complètement celle de son contraire. La possibilité d'un événement et celle de son contraire ne sont que faiblement liées. Dans le cas où l'on cherche à modéliser la croyance, il paraît naturel de ne pas vouloir rendre rigide le lien entre l'évidence que l'on possède en faveur d'un événement et celle que l'on possède en faveur de son contraire. Dans ce cas, la notion de probabilité apparaît moins souple que celle de possibilité.

On peut lire (22) et (23) sous la forme suivante

$$\forall A, \quad \Pi(A) \geq N(A) \quad (24)$$

c'est-à-dire qu'un événement est d'abord possible avant d'être nécessaire. On a d'ailleurs les relations plus fortes déduites de (17), (18).

$$N(A) > 0 \rightarrow \Pi(A) = 1 \quad (25)$$

$$\Pi(A) < 1 \rightarrow N(A) = 0 \quad (26)$$

coïncidant avec \star sur l'ensemble fini de valeurs concernées.

Le type de semi-groupe de $[0,1]$ considéré ici à déjà été étudié par le passé par des mathématiciens en géométrie probabiliste et en théorie des équations fonctionnelles sous le nom de co-normes triangulaires. De l'annexe, il ressort que l'opérateur \star vérifie les inégalités

$$\forall a, b \in [0,1] \quad \max(a,b) \leq a \star b \leq T_W^{\star}(a,b) \quad (27)$$

où

$$T_W^{\star}(a,b) = 1 \quad \forall (a,b) \in]0,1[$$

Les co-normes de base sont, outre "max" et T_W^{\star}

- la somme probabiliste $a + b - ab$
- la somme bornée $\min(1, a+b)$

On peut montrer que (\star) est équivalent à

$$\forall A, B, \quad g(A \cap B) \star g(A \cup B) = g(A) \star g(B) \quad (28)$$

Quand $B = \bar{A}$ on généralise (18) et (21) en

$$g(A) \star g(\bar{A}) = 1 \quad (29)$$

Par décompositions successives, on voit facilement que si $A = \{\omega_{i_1} \dots \omega_{i_p}\}$

$$g(A) = g(\{\omega_{i_1}\}) \star \dots \star g(\{\omega_{i_p}\}) \stackrel{\Delta}{=} \star_{\omega \in A} g(\{\omega\}) \quad (30)$$

Les termes $g_i \stackrel{\Delta}{=} g(\{\omega_{i_j}\})$ sont l'analogue d'une distribution de possibilité ou d'une densité de probabilité. Ils vérifient la condition de normalisation :

$$g_1 \star \dots \star g_n = 1 \quad (31)$$

(30) et (31) sont très intéressants pour savoir mener des calculs pratiques avec des mesures monotones.

On rappelle que (cf annexe) les co-normes continues et archimédiennes ($a \star a > a \quad \forall a \in]0,1[$) possèdent une représentation à l'aide de fonctions continues croissantes : $[0,1] \rightarrow [0,+\infty)$ sous la forme

$$a \star b = \varphi^{\star}(\varphi(a) + \varphi(b)) \quad (32)$$

où φ^* est la pseudo-inverse de φ . φ est appelé générateur additif de $*$.
En particulier

→ Si $\varphi(0) = 0$; $\varphi(1) = +\infty$ on caractérise ainsi les co-normes strictes qui sont strictement croissantes et isomorphes à la somme probabiliste $(a+b-ab)$

→ Si $\varphi(0) = 0$; $\varphi(1) < +\infty$ on obtient les co-normes nilpotentes, isomorphes à la somme bornée, et telles que pour toute suite de valeurs $a_n \in]0, 1[$,
 $\exists k : a_1 * \dots * a_k = 1$

On a les résultats suivants (Dubois et Prade /10/).

Proposition 1.1 : Si $*$ satisfait l'une des deux conditions suivantes :

- i) $*$ est une co-norme stricte
- ii) \exists co-norme stricte $\bar{*}$ telle que $a * b \leq a \bar{*} b$, $\forall a, b$
alors si g vérifie $(*)$, il vérifie (18) c'est-à-dire

$$\max(g(A), g(\bar{A})) = 1 \quad (33)$$

Ce qui implique que la condition de normalisation de la densité s'écrit :

$$\exists i : g_i = 1 \quad (34)$$

En d'autres termes les mesures monotones conformes à la proposition 1 ont un comportement similaire aux fonctions de possibilité. On les appellera pour cette raison des pseudo-possibilités. Bien sûr les fonctions de possibilité elles-mêmes font partie de cette classe ($*$ = max).

En élargissant les hypothèses de la proposition 1, on peut facilement voir que si $*$ est plus petite ou égale à la somme bornée (i.e. $a * b \leq \min(1, a+b) \forall a, b$, alors on a

$$g(A) + g(\bar{A}) \geq 1 \quad (35)$$

qui généralise (21) et (23). Une extension de (35) est obtenue par le résultat suivant

Proposition 1.2 : Si $*$ est une co-norme nilpotente de générateur additif φ alors

$$\forall A, \varphi \circ g(A) + \varphi \circ g(\bar{A}) \geq \varphi(1) \quad (36)$$

La condition de normalisation des mesures décomposables basées sur une co-norme nilpotente s'écrit donc

$$\sum_{i=1}^n \varphi(g_i) \geq \varphi(1) \quad (37)$$

En posant $\varphi =$ identité, on trouve $\ast =$ somme bornée, et en imposant l'égalité dans (37) on retrouve les mesures de probabilité.

A l'aide de la notion de générateur additif, on peut engendrer de nombreuses familles paramétrées de mesures flous, telles les mesures g_λ de Sugeno [30] qui sont définies par $\lambda > -1$ et :

$$A \cap B = \emptyset \rightarrow g_\lambda(A \cup B) = g(A) + g(B) + \lambda g(A) \cdot g(B) \quad (38)$$

Pour $\lambda \leq 0$, elles satisfont (35) et sont une combinaison linéaire entre mesures de probabilité et mesures basées sur la somme probabiliste. (38) suppose que (37) est vérifié avec l'égalité et

$$\varphi(a) = \frac{1}{\lambda} \ln(1+\lambda a) \quad (\lambda \neq -1, 0) \quad (39)$$

Pour $\lambda > -1$ la co-norme $\min(1, a+b+\lambda ab)$ est nilpotente.

Quand l'égalité est vérifiée dans (37), il est clair que les termes

$$\frac{\varphi(g_i)}{\varphi(1)} \triangleq \hat{\varphi}(g_i)$$

peuvent être assimilés avec des valeurs de probabilité : soit P_φ la mesure de probabilité associée. On a alors

$$\forall A, g(A) = \varphi^{-1} \left(\sum_{\omega_i \in A} \varphi(g_i) \right) = \hat{\varphi}^{-1} (P_\varphi(A)) \quad (40)$$

C'est-à-dire que g est une mesure de probabilité à une transformation près.

1.3.2 - Mesures d'incertitude duales

On a vu que lorsqu'on utilise des fonctions de possibilité, il est nécessaire, pour posséder toute l'information relative à un événement A de connaître $\Pi(A)$ et $N(A) = 1 - \Pi(\bar{A})$, les deux informations n'étant pas redondantes. La dernière égalité exprime que le degré de nécessité d'un

événement n'est autre que le degré d'impossibilité de l'événement contraire. Connaissant le degré de possibilité, le degré d'impossibilité s'en déduit à l'aide d'un opérateur de négation C dont le choix n'est pas forcément réduit à $1-u = C(u)$. Des conditions naturelles qui définissent C , application de $[0,1]$ sur lui-même, sont :

- . $C(1) = 0$
- . C est strictement décroissant
- . C est involutif : $C \circ C = \text{Identité}$.

Soit g une mesure floue. La mesure floue duale \bar{g}_C au sens de la négation C se définit par :

$$\bar{g}_C(A) = C(g(\bar{A})) \quad \forall A \subseteq \Omega \quad (41)$$

Si g vérifie $(*)$ pour la co-norme $*$, alors \bar{g}_C vérifie l'axiome suivant :

$$\forall A, B, \quad A \cup B = \neg \neg (A \cap B) \quad \bar{g}_C(A \cap B) = \bar{g}_C(A) \perp g_C(B) \quad (4)$$

où l'opération \perp est définie par

$$a \perp b = C(C(a) * C(b)) \quad (42)$$

On peut vérifier que $([0,1], \perp)$ forme un semi-groupe commutatif non décroissant, d'élément neutre 1; \perp est une norme triangulaire (cf annexe) ou t-norme. (42) est une transformation de De Morgan qui transforme les co-normes en t-normes et inversement.

Les mesures floues \bar{g} satisfaisant (4) sont définies indépendamment de la négation. L'opération \perp est toujours telle que

$$T_W(a,b) \leq a \perp b \leq \min(a,b)$$

où $T_W(a,b) = 0 \quad \forall a,b \in [0,1[$

Les t-normes de base sont outre \min et T_W ,

- le produit $a.b$ qui engendre les t-normes strictes par isomorphisme
- la t-norme $T_m(a,b) = \max(0, a+b-1)$ qui engendre les t-normes nilpotentes par isomorphisme.

On trouve pour ces mesures floues des propriétés analogues à (28) - (31) à savoir

$$(1) \leftrightarrow \bar{g}(A \cup B) \perp \bar{g}(A \cap B) = \bar{g}(A) \perp \bar{g}(B) \quad (43)$$

$$\bar{g}(A) \perp \bar{g}(\bar{A}) = 0 \quad (44)$$

$$\text{if } A = \{\omega_{i_1}, \dots, \omega_{i_p}\}, \bar{g}(A) = \bar{g}_{i_{p+1}} \perp \dots \perp \bar{g}_{i_n} \triangleq \bigwedge_{\omega_i \notin A} \bar{g}_i \quad (45)$$

où $\bar{g}_i = \bar{g}(\Omega - \{\omega_i\})$ et

$$\bar{g}_1 \perp \dots \perp \bar{g}_n = 0 \quad (46)$$

pour $\perp = \min$ on obtient les fonctions de nécessité. On peut démontrer des résultats analogues aux propositions 1 et 2 (cf. Dubois et Prade /10/).

1.3.3 - Mesures auto-duales

On peut se demander dans quel cas une mesure floue g basée sur une co-norme, et la mesure duale \bar{g}_C coïncident, i.e.

$$\bar{g}_C(A) = g(A) = C(g(\bar{A})), \forall A \quad (47)$$

Les propositions suivantes (Dubois et Prade /10/) répondent à cette question

Proposition 3. : si g vérifie $(*)$ où $*$ est soit une co-norme stricte soit une co-norme comprise entre max et une co-norme stricte, alors les deux affirmations suivantes sont équivalentes.

- g satisfait (\perp) pour une t-norme
- g est une mesure de Dirac

Cette proposition se contente d'entériner le fait que pour les pseudo-possibilités (conformes à la proposition 1) les deux informations $g(A)$ et $g(\bar{A})$ ne peuvent pas se déduire l'une de l'autre, et (47) est faux en général.

Proposition 4. : si g vérifie $(*)$ où $*$ est une co-norme nilpotente de générateur φ telle que

$$\sum_{i=1}^n \varphi(g_i) = \varphi(1) \quad (48)$$

alors $g = \bar{g}_C$ pour $C = \varphi^{-1}(\varphi(1) - \varphi(\cdot))$

Ceci est dû au fait que toute négation s'exprime en fonction d'un générateur additif de co-norme nilpotente (Trillas, /31/) (cf. Annexe). En particulier, si $\varphi(x) = x, \forall x$, on retrouve les probabilités, et (47) exprime $P(A) + P(\bar{A}) = 1$.

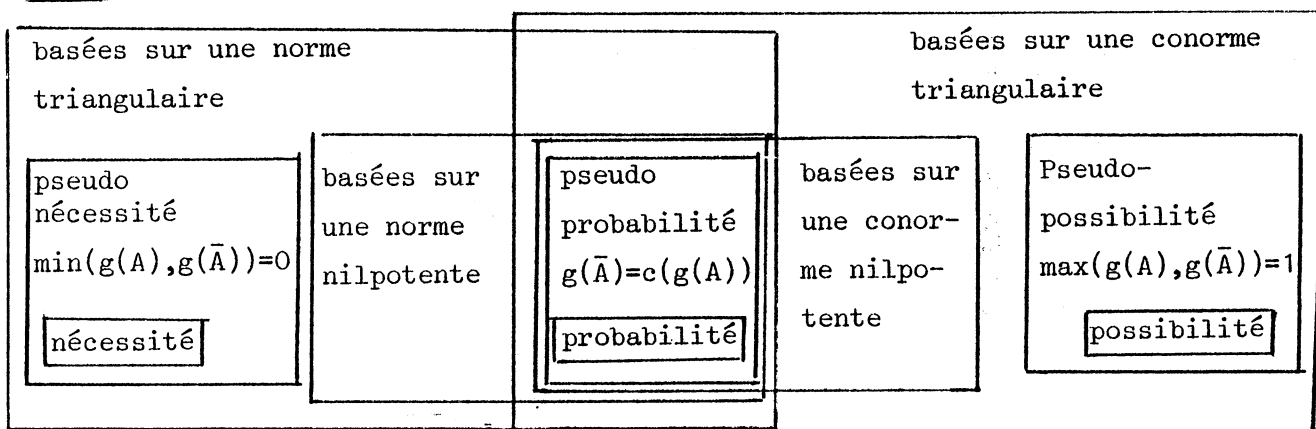
Donc la proposition 4 caractérise une classe de mesures monotones telles que $g(\bar{A})$ s'exprime en fonction de $g(A)$, et qui sont des probabilités à un effet de distortion près. On peut les appeler des pseudo-probabilités. Les mesures g_λ de Sugeno sont des pseudo-probabilités ($\lambda > -1$).

1.3.4 - Conclusion

L'étude systématique de toutes les fonctions d'ensemble monotones décomposables est très instructive car elle met en évidence les fonctions de possibilité et de nécessité d'une part et les mesures de probabilité d'autre part comme les deux notions basiques, irréductibles l'une à l'autre, pour la représentation de la connaissance incertaine. Les classes de fonctions rencontrées sont représentées ci-dessous:

mesures monotones décomposables (sauf mesures de Dirac) sur des domaines

finis



Toutes les fonctions décomposables g telles que $g(A)$ et $g(\bar{A})$ se déduisent l'un de l'autre par une transformation C sont réductibles à des probabilités (par (40)). Toutes les pseudo-possibilités étendent au cas multivalué la définition classique du concept "possible" (voir équation (7)), les "pseudo-nécessités" étendent le même le concept classique de "nécessaire" (équation (14)).

1.3.5 - Extension au cas d'ensembles infinis : quelques remarques.

L'extension des pseudo-probabilités et des pseudo-possibilités à des ensembles quelconques est loin d'avoir été menée à bien. Sugeno /30/ Tsukamoto et al. /32/, Kruse /20/ ont étendu les mesures g_λ à l'évaluation d'intervalles réels. Dubois /5/ (cf. aussi Dubois et Prade /10/) a suggéré des définitions de $g(A)$ où A est un intervalle réel fermé dans le cas où g admet une représentation à l'aide d'un générateur additif (de t-norme stricte ou nilpotente). Schwyhla /27/, considère un ensemble mesurable (X, \mathcal{A}) et une mesure de Sugeno $m = \mathcal{A} \rightarrow \bar{\mathbb{R}}$ continue ; moyennant une hypothèse forte sur la façon dont sont distribuées les valeurs $g(A)$, à savoir

$$\forall A \in \mathcal{A}, \forall b < g(A) \quad \exists B \subset A, g(B) = b \quad (49)$$

et si g est décomposable (eqn (x)) alors g est isomorphe à une mesure additive tronquée. Ce résultat est valide pour les mesures g_λ de Sugeno. Schwyhla peut alors étendre diverses notions classiques dont la mesure de Lebesgue d'un intervalle, la mesure produit, l'intégrale d'une fonction arbitraire etc... au sens de la mesure g_λ . On notera que la condition (49) exclut que l'ensemble Ω soit fini, et aussi qu'un ensemble fini ait une évaluation non nulle. Les fonctions de possibilité et de nécessité ne sont pas concernées pour cette approche, par conséquent (d'ailleurs elles ne sont pas continues). Le cas d'un référentiel infini dénombrable n'est pas non plus très bien cerné par l'axiome (49). Néanmoins le travail de Schwyhla, bien que limité par cet axiome, concerne aussi bien les mesures de Sugeno basées sur une co-norme nilpotente que sur une co-norme stricte. Il faut signaler que dans le cas de référentiel infini, la condition de normalisation d'une fonction d'ensemble basée sur une co-norme stricte ne s'écrit plus selon (34) forcément.

REMARQUE : Pour des pseudo-probabilités, on peut aussi, dans le cas fini étendre beaucoup de notions classiques des probabilités. Par exemple si

$$V = \sum_{i=1}^m v_i \mu_{U_i}$$

est une variable aléatoire définie sur une partition $U_1 \dots U_m$ de Ω , l'espérance mathématique $E_g(X)$ au sens d'une pseudo-probabilité g basée sur la co-norme X vérifiant (34) peut se définir par /27/ :

$$E_g(V) = \hat{\varphi}^{-1} \left(\sum_{i=1}^m \hat{\varphi}(v_i) \cdot P_\varphi(U_i) \right) \quad \text{pour } v_i \in \text{Domaine de } \hat{\varphi}, \forall i$$

Ce type d'espérance mathématique distordue n'est pas définie pour des pseudo-possibilités.

1.4 - MESURES D'INCERTAIN INDUITES PAR UNE PROBABILITE ET UNE FONCTION MULTIVOQUE

Au lieu de généraliser la loi d'additivité d'une fonction de probabilité, on peut, pour sortir du cadre des fonctions d'ensemble additives, transporter une mesure de probabilité sur un ensemble mesurable (X, \mathcal{A}) dans un autre ensemble (Ω, \mathcal{B}) mesurable, par l'intermédiaire d'une fonction multivoque. C'est ce qui a été réalisé par Dempster /4/. Les êtres mathématiques que l'on obtient ainsi sur Ω ont été étudiés indépendamment par Choquet /3/ sous le nom de "capacités alternantes" et "capacités monotones", et par Kendall /19/, Matheron /22/ sous le nom d'ensembles aléatoires. La classe de fonctions d'ensemble ainsi obtenues contient probabilités, possibilités et nécessités. L'intérêt pratique de cette approche réside peut-être principalement dans les liens qu'elle souligne entre mesures non-additives et statistique permettant ainsi d'interpréter les ensembles flous non plus seulement comme modèle commode de catégories vagues du langage, mais aussi comme résultat objectif d'un processus de recueil de données statistiques. Il est à noter que Shafer /29/ réinterprète les "probabilités supérieures" et "inférieures" de Dempster en termes de "plausibilité" et de "crédibilité" suggérant ainsi un modèle de mesure subjective d'incertitude, alors que les travaux de Dempster sont exempts de telles préoccupations.

1.4.1 - Probabilités supérieures et inférieures - cas fini.

Soit X un ensemble fini à m éléments et P_x une mesure de probabilité définie par une allocation de poids

$$\{p_i \mid i=1, m; \sum_{i=1}^m p_i = 1\}$$

Soit Γ une application $X \rightarrow 2^\Omega$. X est une liste d'épreuves et $\Gamma(x)$ est l'interprétation de l'épreuve c'est-à-dire l'événement correspondant au succès de l'épreuve x . Γ est multivoque au sens où $\Gamma(x)$ n'est pas forcément un événement élémentaire (atome) de Ω , $\{\Gamma(x) \mid x \in X\}$ n'est pas forcément une partition de Ω .

Soit un événement $A \subseteq \Omega$, l'image réciproque de A par Γ se compose d'une image minimale

$$A_* = \{x \in X \mid \Gamma(x) \subseteq A, \Gamma(x) \neq \emptyset\} \quad (50)$$

et d'une image maximale

$$A^* = \{x \in X \mid \Gamma(x) \cap A \neq \emptyset\} \quad (51)$$

A_* est l'ensemble des épreuves dont le succès rend nécessaire l'occurrence de A , A^* l'ensemble des épreuves dont le succès rend possible l'occurrence de A . Autrement dit, si γ est une fonction $X \rightarrow \Omega$ compatible avec Γ ($\forall x, \gamma(x) \in \Gamma(x)$) on a

$$A_* \subseteq \gamma^{-1}(A) \subseteq A^* \quad (52)$$

Dempster définit, pour tout $A \subseteq \Omega$, et si $\forall x \Gamma(x) \neq \emptyset$, (ce qu'on supposera sauf précision contraire):

$$\text{la probabilité supérieure de } A : P^*(A) = P_X(A^*)$$

$$\text{la probabilité inférieure de } A : P_*(A) = P_X(A_*)$$

au sens où $P^*(A)$ resp($P_*(A)$) est l'allocation maximale (resp : minimale) de probabilité que l'on peut affecter à l'événement A , connaissant P sur X . On vérifie que P_* (resp P^*) est monotone (resp : alternante) d'ordre infini, ce qui s'écrit

$$\forall k \in \mathbb{N}, \forall A_1 \dots A_k \in 2^\Omega$$

$$\text{(monotone d'ordre } k) \quad P_*(A_1 \cup \dots \cup A_k) \geq \sum_{I \subseteq [1, k]} (-1)^{|I|+1} P_*\left(\bigcap_{j \in I} A_j\right) \quad (53)$$

$$\text{(alternante d'ordre } k) \quad P^*(A_1 \cup \dots \cup A_k) \leq \sum_{I \subseteq [1, k]} (-1)^{|I|+1} P^*\left(\bigcap_{j \in I} A_j\right) \quad (54)$$

avec $[1, k] = \{1, 2, \dots, k-1, k\}$, et $I \neq \emptyset$

Ecrites à l'ordre 2 (53) et (54) donnent

$$P_*(A \cup B) \geq P_*(A) + P_*(B) - P_*(A \cap B) \quad (55)$$

$$P^*(A \cup B) \leq P^*(A) + P^*(B) - P^*(A \cap B) \quad (56)$$

On voit que les probabilités inférieures sont sur-additives et les probabilités inférieures sont sous-additives. Il est facile de voir que ce sont des mesures floues de Sugeno, au sens de (2).

On appelle éléments focaux les éléments de l'ensemble

$$\mathcal{F} = \{\Gamma(x_i) \mid p_i > 0, x_i \in X\} \quad (57)$$

C'est l'ensemble des événements sur lesquels on possède des présomptions relatives à leur occurrence. Si l'ensemble vide appartient à \mathcal{F} , cela signifie que l'ensemble des événements considérés n'est pas suffisant, où qu'une partie de l'information que l'on possède est ininterprétable. Dans ce cas on peut modifier les définitions des probabilités supérieures et inférieures en normalisant par la quantité $P_X(\Omega^*) = P_X(\Omega_*)$; sinon on travaille sur X avec une mesure de probabilité non-normalisée; $P_X(\Omega^*)$ évalue la qualité de l'information disponible. La donnée de (X, P_X, Γ) revient, dans le cas fini à celle d'une fonction $m = 2^\Omega \rightarrow [0, 1]$ définie par

$$\forall A \in \mathcal{F}, \quad m(A) = p_i \quad \text{t.q.} \quad A = \Gamma(x_i) \quad (58)$$

$$\forall A \notin \mathcal{F} \quad m(A) = 0$$

et qui vérifie

$$\begin{cases} \sum_{A \subseteq \Omega} m(A) = 1 \\ m(\emptyset) = P(X - \Omega^*) \end{cases} \quad (59)$$

$m(\emptyset) = 0$ si et seulement si tous les x_i ont une image non vide ($\Gamma(X) \subseteq 2^\Omega - \{\emptyset\}$). m est appelée probabilité de base. On a le résultat remarquable suivant (Shafer, 1976) : P_* et P^* satisfont (53) et (54) si et seulement si

$$\forall A, P_*(A) = \sum_{\substack{B \subseteq A \\ B \neq \emptyset}} m(B) \quad (60)$$

$$P^*(A) = \sum_{\substack{B \subseteq A \\ B \neq \emptyset \\ B \cap A^* \neq \emptyset}} m(B) \quad (61)$$

on connaît les formules inverses définissant m à partir de P^* ou P_* (cf Shafer, /29/). Il est facile de voir, de plus, que

$$\forall A, P_*(A) = 1 - P^*(\bar{A}) \quad (62)$$

$$P^*(A) \geq P_*(A) \quad (63)$$

D'où l'interprétation "subjectiviste" de P^* comme fonction de plausibilité (mesure le vraisemblable, ce qui peut arriver, en accord avec (51) et (61) et de P^* comme fonction de crédibilité (ce qui doit arriver, en accord avec (50) et (60)). (62) exprime qu'un événement est plausible lorsque son contraire est peu crédible. (63) que la crédibilité est une notion plus dure que la plausibilité.

1.4.2 - Cas particuliers : Probabilités et possibilités

Une mesure de probabilité P est à la fois une fonction de plausibilité et une fonction de crédibilité. On a de façon équivalente

$$P^* = P_{\star} = P \text{ (mesure de probabilité)}$$

. (53) ou (54) est vérifiée avec l'égalité (dans ce cas l'ordre 2 suffit)

. \mathcal{F} ne contient que des singletons.

Une fonction de possibilité est une fonction de plausibilité. P^* est une fonction de possibilité si et seulement si l'une des conditions suivantes est remplie.

$$- \quad \forall A, B \quad P^*(A \cup B) = \max(P^*(A), P^*(B)) \quad (64)$$

- \mathcal{F} est une suite monotone d'ensembles emboîtés

On peut vérifier facilement que si $\Omega = \{\omega_1 \dots \omega_n\}$ tel que

$$Pl(\omega_1) = \pi_1 \geq Pl(\omega_2) = \pi_2 \geq \dots Pl(\omega_n) = \pi_n,$$

en notant

$$A_1 = \{\omega_1\}, \quad A_2 = \{\omega_1, \omega_2\}, \quad \dots A_j = \{\omega_1 \dots \omega_j\} \quad \text{etc...}$$

on a :

$$\mathcal{F} \subseteq \{A_1 \dots A_n\} \quad (65)$$

$$\forall i = 1, n \quad \pi_i = \sum_{j=i}^n m(A_j) \quad (66)$$

ou

$$m(A_i) = \pi_i - \pi_{i+1} \quad (\text{avec } \pi_{n+1} = 0) \quad (67)$$

Une fonction de nécessité est une fonction de crédibilité, de par (62). On voit que probabilité et possibilité correspondent une fois de plus à deux cas limites radicalement opposés :

- les probabilités correspondent à une information précise (les éléments focaux sont des singletons) mais contradictoire, dissonante (éléments focaux disjoints)
- les possibilités correspondent à une information cohérente (éléments focaux emboîtés) mais (= donc) imprécise.

Mais dans tous les cas, on voit que le couple (P^*, P_*) définit des valeurs limites de probabilité des événements $A \subseteq \Omega$. Les mesures de probabilité correspondantes sont définies par la classe

$$\mathcal{C} = \{P \mid \exists \gamma = X \rightarrow \Omega : \forall x, \gamma(x) \in \Gamma(x), P(A) = \sum_{x \in \gamma^{-1}(A)} p(x)\} \quad (68)$$

c'est-à-dire les mesures de probabilités images de P_X par une application compatible avec Γ . On a

$$\forall P \in \mathcal{C}, P^*(A) \geq P(A) \geq P_*(A) \quad (69)$$

La classe

$$\hat{\mathcal{C}} = \{P \mid \forall A, P^*(A) \geq P(A) \geq P_*(A)\}$$

peut être caractérisée à l'aide de la probabilité de base, en partageant les "poids d'évidence" : On peut définir $k(\omega, A)$, $\omega \in A$ tel que

$$\forall A \in \mathcal{F}, m(A) = \sum_{\omega \in A} k(\omega, A) \quad (70)$$

et

$$P(\omega) = \sum_{A \ni \omega} k(\omega, A) \quad (71)$$

Autrement dit on répartit les poids d'évidence sur les singletons des éléments focaux. (68) revient à choisir $\omega \in A$ tel que $k(\omega, A) = m(A)$. Dempster /4/ a montré que $\hat{\mathcal{C}}$ était la fermeture convexe de \mathcal{C} .

Une façon très naturelle d'approximer un couple (P^*, P_*) par

une mesure de probabilité est d'équirépartir les poids d'évidence $m(A)$ sur les singletons des éléments focaux, ce qui fournit $P \in \mathcal{E}$ telle que (Prade, /23/)

$$\forall \omega \quad P(\omega) = \sum_{A \ni \omega} \frac{m(A)}{|A|} \quad (72)$$

Cette approximation est similaire à celle que l'on fait en inférence bayésienne lorsqu'en l'absence d'information, on utilise une distribution uniforme comme probabilité a priori ⁽¹⁾

On peut aussi approximer une probabilité supérieure (resp : inférieure) par une fonction de possibilité (resp : nécessité). En effet, on a (Prade, /23/), (Dubois et Prade /11/).

$$\begin{aligned} \forall A, P^*(A) &= \sum_{B \subseteq A} m(B) \max_{\omega \in A} \mu_B(\omega) \\ &\geq \max_{\omega \in A} \sum_{B \subseteq \Omega} m(B) \mu_B(\omega) = \max_{\omega \in A} P^*({\omega}) \stackrel{\Delta}{=} \Pi_P(A) \end{aligned} \quad (73)$$

$$P_*(A) = 1 - P^*(\bar{A}) \leq \inf_{\omega \in \bar{A}} 1 - P^*({\omega}) \stackrel{\Delta}{=} N_P(A) \quad (74)$$

Autrement dit, (73) et (74) reviennent à ne conserver de P^* et P_* que leurs valeurs sur les singletons de Ω , et à considérer ces valeurs comme une distribution de possibilité (ou un ensemble flou, cf. chapitre 2). On a alors les inégalités

$$\forall A, P^*(A) \geq \Pi_P(A) \geq N_P(A) \geq P_*(A) \quad (75)$$

Pour que Π_P et N_P soient normalisées il est nécessaire et suffisant que les éléments focaux de (P^*, P_*) soient globalement non disjoints, i.e.

$$\bigcap_{A \in \mathcal{F}} A \neq \emptyset \quad (76)$$

Parmi les fonctions d'ensemble définies par une probabilité de base,

(1) La définition d'une probabilité compatible avec P^* et P_* peut d'ailleurs s'interpréter dans une optique Bayésienne : la quantité $k(\omega, A)$ dans (70) peut s'écrire

$$k(\omega, A) = p(x) \cdot p(\omega|x) \quad \text{où} \quad A = \Gamma(x) \quad \text{et} \quad \sum_{\omega \in A} p(\omega|x) = 1$$

$p(\omega|x)$ est la probabilité de déduire ω d'une observation x . On voit alors $\Gamma(x)$ comme une fonction aléatoire. (71) s'écrit alors

$$p(\omega) = \sum_x p(\omega|x) p(x)$$

on trouve les mesures de Dirac, les seules qui soient à la fois des probabilités des possibilités, et des nécessités. On trouve aussi la fonction d'ignorance totale, construite à partir de :

$$m(\Omega) = 1 \quad ; \quad m(A) = 0 \quad \text{si} \quad A \subset \Omega \quad (77)$$

Elle définit une fonction de possibilité Π_0 telle que $\Pi_0(A) = 1 \quad \forall A \neq \emptyset$, et une fonction de nécessité : $N_0(A) = 0 \quad \forall A \neq \Omega$. Il est notable que la représentation de l'ignorance totale, utile pour définir des distributions a priori, sorte du cadre probabiliste (cf. Banon, /2/).

1.4.3 - Cas particuliers : fonctions d'ensemble décomposables

On peut se demander quels sont les points de rencontre entre l'approche esquissée au 1.3 et celle de Dempster-Shafer. La question n'est pas totalement résolue, mais on possède quelques éléments de réponse.

Les deux familles ont en commun les mesures de probabilité, et les fonctions de possibilité et de probabilité.

Il est démontré par Dubois et Prade /10/ que les fonctions d'ensemble basées sur des t-normes (resp : co-normes) distributives sont des probabilités inférieures (resp : supérieures) (voir annexe pour cette notion). Ce sont toutes des pseudo-nécessités (resp : possibilités).

Les deux t-normes (co-normes) distributives basiques sont minimum (maximum) et produit (somme probabiliste) ; les autres se construisent à partir de celles-là par somme ordinale (voir annexe).

Donc les fonctions telles que :

$$\forall A, B, A \cup B = \Omega, \quad \underline{g}(A \cap B) = \underline{g}(A) \cdot \underline{g}(B) \quad (78)$$

$$\forall A, B, A \cap B = \emptyset \quad \bar{g}(A \cup B) = \bar{g}(A) + \bar{g}(B) - \bar{g}(A)\bar{g}(B) \quad (79)$$

sont des probabilités inférieures (resp. supérieures). On a

$$\bar{g}(A) = 1 - \underline{g}(\bar{A}) \quad (80)$$

c'est-à-dire que (\bar{g}, g) correspond à une probabilité de base unique. Elle est calculée explicitement dans Dubois et Prade /10/. On sait que les deux fonctions (\bar{g}, g) se définissent à partir d'une densité $\{g_i | \max_i g_i = 1\}$ d'après la proposition 1. On montre que les éléments focaux de (\bar{g}, g) sont de la forme

$$\mathcal{F} = \{A \mid K \stackrel{\Delta}{=} \{\omega_i \mid g_i = 1\} \subseteq A\} \quad (81)$$

On remarque que \mathcal{F} est un filtre au sens de la topologie. Néanmoins la structure de \mathcal{F} ne suffit pas à caractériser la nature de la fonction d'ensemble, comme c'est le cas pour probabilités (singletons) et possibilités-nécessités (suite monotone), puisque $\{m(A) \mid A \in \mathcal{F}\}$ est défini à partir des n valeurs g_1, \dots, g_n , et le cardinal de \mathcal{F} est en général plus grand que celui de Ω ($=n$). Ces fonctions ne semblent donc pas basiques.

Banon /1/ a montré que les mesures g_λ de Sugeno sont des probabilités supérieures pour $\lambda \in]-1, 0]$ et inférieures pour $\lambda \in [0, +\infty)$, en exhibant la probabilité de base. Si l'on remarque que pour $\lambda < 0$, g_λ est une combinaison convexe de probabilité et de fonction \bar{g} définie par (79), et que de plus toute combinaison convexe de probabilités supérieures est une probabilité supérieure, le résultat devient évident pour $\lambda \leq 0$. Si $\lambda > 0$, il suffit de signaler, pour conclure, que

$$\forall A \quad g_\lambda(A) = 1 - \frac{g_{-\lambda}}{1+\lambda} \quad (\bar{A}) \quad (82)$$

Les éléments focaux d'une fonction g_λ sont de la forme $\mathcal{F} = 2^{\Omega^+}$ où $\Omega^+ = \{\omega_i \mid g_i > 0\}$ si g_1, \dots, g_n est la densité normalisée de g_λ au sens de la proposition 2, i.e.

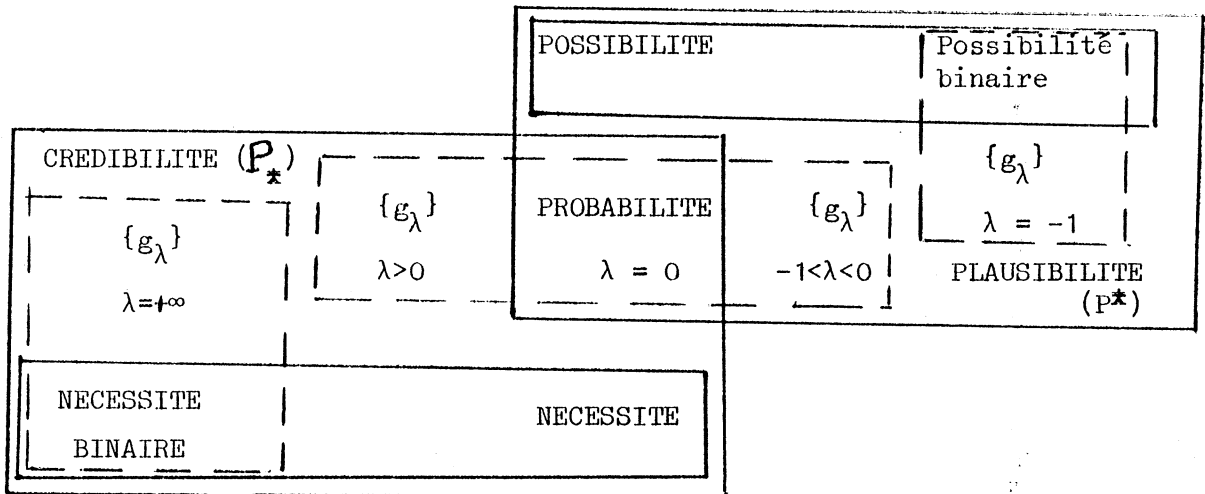
$$\prod_{i=1}^n (1 + \lambda g_i) = 1 + \lambda \quad (83)$$

En revanche certaines mesures floues décomposables ne sont pas des probabilités supérieures ou inférieures puisque

$$\forall A, B \quad \max(P^*(A), P^*(B)) \leq P^*(A \cup B) \leq \min(1, P^*(A) + P^*(B))$$

Si on choisit une co-norme supérieure à la somme bornée on est sûr d'obtenir une fonction d'ensemble qui n'est pas une probabilité supérieure.

On trouvera ci-dessous un panorama des probabilités inférieures et supérieures.



Probabilités supérieures et inférieures, sauf mesures de Dirac, sur des domaines finis.

REMARQUE : Soit \mathcal{P}_Ω l'ensemble des mesures de probabilité sur Ω , π_Ω l'ensemble des fonctions de possibilité sur Ω et \mathcal{M} l'ensemble des probabilités supérieures sur Ω . On sait que \mathcal{P}_Ω et \mathcal{M} sont convexes. π_Ω ne l'est pas. On peut considérer son enveloppe convexe $\hat{\pi}_\Omega$. Etant donné le caractère basique de \mathcal{P}_Ω et π_Ω on peut se demander si \mathcal{M} ne peut pas être généré comme enveloppe convexe de $\mathcal{P}_\Omega \cup \hat{\pi}_\Omega$; cette conjecture valide si $|\Omega| = 2$ reste vérifier dans le cas général.

1.4.4 - Espérances mathématiques supérieures et inférieures

Soit $\{U_1 \dots U_m\}$ une partition de Ω et μ_{U_i} la fonction caractéristique de U_i , $\forall i = 1, m$. On considère une variable aléatoire

$$V = \sum_{i=1}^m v_i \mu_{U_i} \quad ; \quad v_1 < v_2 < \dots < v_m \in \mathbb{R}$$

que l'on appellera variable d'incertitude.

Connaissant un couple de probabilités inférieures et supérieures (P_*, P^*) sur Ω , on peut définir à l'aide d'une intégrale de Lebesgue-Stieljes des fonctions de répartition supérieures et inférieures F^* et F_* par

$$\forall r \in \mathbb{R}, \quad F^*(r) = P^*(V \leq r) \quad (84)$$

$$F_*(r) = P_*(V \leq r) \quad (85)$$

et des espérances mathématiques inférieures et supérieures par

$$E_{\star}(V) = \int_{-\infty}^{+\infty} r dF_{\star}(r) \quad (86)$$

$$E^{\star}(V) = \int_{-\infty}^{+\infty} r dF^{\star}(r) \quad (87)$$

Ces concepts ont été introduits par Dempster /4/. Si \mathcal{E} est l'ensemble des mesures de probabilités P compatibles avec (P_{\star}, P^{\star}) (voir section précédente) et si $E(V)$ est l'espérance mathématique au sens de P

$$E_{\star}(V) = \inf_{P \in \mathcal{E}} E(V) \quad ; \quad E^{\star}(V) = \sup_{P \in \mathcal{E}} E(V) \quad (88)$$

De plus les résultats suivants sont démontrés dans Dubois et Prade /12/ supposons $v_1 < v_2 \dots < v_m$:

$$E_{\star}(V) = \sum_{i=1}^m v_i \left[P_{\star} \left(\bigcup_{j=i}^m U_j \right) - P_{\star} \left(\bigcup_{j=i+1}^m U_j \right) \right] \quad (89)$$

$$= \sum_{A \subseteq \Omega} m(A) \max \{v_k \mid A \cap U_k \neq \emptyset\} \quad (90)$$

$$E^{\star}(V) = \sum_{i=1}^m v_i \left[P^{\star} \left(\bigcup_{j=i}^m U_j \right) - P^{\star} \left(\bigcup_{j=i+1}^m U_j \right) \right] \quad (91)$$

$$= \sum_{A \subseteq \Omega} m(A) \min \{v_k \mid A \cap U_k \neq \emptyset\} \quad (92)$$

Höhle /16/ a fait remarquer que (91) est valide pour des mesures monotones de Sugeno quelconques. Par ailleurs il est facile de voir que

$$E_{\star}(-V) = -E^{\star}(V) \quad (93)$$

$$E_{\star}(a+bV) = a + bE_{\star}(V) \quad (b > 0) \quad (94)$$

$$= a + bE^{\star}(V) \quad (b < 0)$$

$$E_{\star}(V) + E_{\star}(V') \leq E_{\star}(V+V') \leq E_{\star}(V) + E^{\star}(V') \quad (95)$$

$$E_{\star}(V) + E^{\star}(V') \leq E^{\star}(V+V') \leq E^{\star}(V) + E^{\star}(V') \quad (96)$$

Ce concept d'espérance mathématique généralisé paraît très intéressant à étudier plus en détail, car il semble qu'on puisse étendre à la théorie des probabilités supérieures et inférieures certains concepts et résultats de la théorie classique des probabilités; par exemple on peut facilement,

dans le cas fini démontrer une inégalité de Tchebitchev sous la forme

$$\forall \varepsilon > 0 \quad P_{\star}(|V| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} E_{\star}(V^2) \quad (97)$$

De plus ce concept inventé par Dempster /4/ indépendamment de la théorie des ensembles flous s'avère très utile pour unifier certains résultats dans cette dernière. Si Ω est fini et $U_i = \{\omega_i\} \forall i$, si $\{\pi_i\}_{i=1,n}$ est une distribution de possibilité avec $\pi_{i\star} = 1$ alors l'espérance "possibiliste" de V est définie par les deux bornes

$$E_{\star}^{\star}(V) = \sum_{i=1}^n v_i (\max_{k>i} \pi_k - \max_{k>i} \pi_k) \quad (98)$$

$$E_{\star}^{\star}(V) = \sum_{i=1}^n v_i (\max_{k \leq i} \pi_k - \max_{k < i} \pi_k) \quad (99)$$

1.4.5 - Itération du processus de génération des probabilités supérieures et inférieures

Il est tentant de vérifier si la méthode de génération des probabilités supérieures et inférieures par transport d'une mesure de probabilité à l'aide d'une fonction multivoque est indépendante de la façon dont on définit cette fonction, en particulier si Γ est définie par composition d'applications multivoques.

Soit

$$\Gamma_1 = X \rightarrow 2^{\Omega_1} \quad \text{et} \quad \Gamma_2 : \Omega_1 \rightarrow 2^{\Omega_2}.$$

On note $\Gamma_{1\star}^{-1}(A)$ et $\Gamma_1^{-1\star}(A)$ les images réciproques minimales et maximales de $A \subseteq \Omega_1$ par Γ_1 . On montre (Dubois et Prade, /14/), $\forall A \subseteq \Omega_2$

$$\Gamma_1^{-1\star}(\Gamma_2^{-1\star}(A)) = (\Gamma_2 \circ \Gamma_1)^{-1\star}(A) \quad (100)$$

$$\Gamma_{1\star}^{-1}(\Gamma_{2\star}^{-1}(A)) = (\Gamma_2 \circ \Gamma_1)_{\star}^{-1}(A) \quad (101)$$

(on rappelle que $\Gamma_1(x) \neq \emptyset$, $\Gamma_2(\omega_1) \neq \emptyset$)

Une fonction de plausibilité supérieure (resp = inférieure) $P^{\star\star}$ (resp : P_{\star}^{\star}) est définie sur Ω_2 par

$$\left. \begin{aligned} P^{**}(A) &= P^*(\Gamma_2^{-1*}(A)) \\ P_{**}^*(A) &= P^*(\Gamma_{2*}^{-1}(A)) \end{aligned} \right\} \quad (102)$$

On voit tout de suite que, par (100)

$$P^{**}(A) = P((\Gamma_2 \circ \Gamma_1)^{-1*}(A)) \quad (103)$$

De même, une fonction de crédibilité supérieure (resp : inférieure) P_{**}^* (resp : P_{**}) est définie sur Ω_2 par

$$P_{**}^*(A) = P_{**}(\Gamma_2^{-1*}(A)) \quad (104)$$

$$P_{**}(A) = P_{**}(\Gamma_{2*}^{-1}(A)) \quad (105)$$

Et on a de même, par (101)

$$P_{**}(A) = P((\Gamma_2 \circ \Gamma_1)_{**}^{-1}(A)) \quad (106)$$

(103) (resp (106)) exprime qu'une fonction de plausibilité supérieure (resp = crédibilité inférieure) reste une fonction de plausibilité (resp : crédibilité). Ce résultat reste valable à l'ordre n et indique que le processus de génération des probabilités supérieures (resp : inférieures) est fermé.

On vérifie que

$$P^{**}(A) = 1 - P_{**}(\bar{A}) \geq P_{**}^*(A) \quad (107)$$

$$P_{**}^*(A) = 1 - P_{**}^*(\bar{A}) \quad (108)$$

Mais P_{**}^* et P_{**} ne sont en général pas des fonctions de crédibilité ou de plausibilité, parce qu'on peut trouver des exemples tels que :

$\exists A : P_{**}^*(A) > P_{**}^*(A)$ et $\exists B, P_{**}^*(B) > P_{**}^*(B)$. Néanmoins ce sont encore des mesures floues de Sugeno.

Si on applique ce processus à des fonctions de possibilité et de nécessité, on montre qu'une fonction de possibilité supérieure (resp : nécessité inférieure) est une fonction de possibilité (resp : nécessité).

C'est évident si l'on remarque que $\forall \Gamma_2 : \Omega_1 \rightarrow 2^{\Omega_2}$. Si $\{\Gamma_1(x)\}_{x \in X}$ est une suite monotone, alors $\{\Gamma_2 \circ \Gamma_1(x)\}_{x \in X}$ est aussi une suite monotone.

Soit Π une fonction de possibilité sur Ω_1 et soit

$$\phi(\Omega_2) = \{ \Gamma_2(\omega_1) \mid \omega_1 \in \Omega_1 \} \quad (109)$$

et $\emptyset \notin \Phi(\Omega_2)$. Alors si on définit la "possibilité de base" par $\beta : 2^{\Omega_2} \rightarrow [0,1]$ telle que

$$\begin{aligned} \beta(A) &= \Pi(\{\omega_1 \mid \Gamma_2(\omega_1) = A\}), A \in \Phi(\Omega_2) \\ &= 0 \text{ sinon} \end{aligned}$$

on a

$$\Pi^*(A) = \max_{B: A \cap B \neq \emptyset} \beta(B) \quad (110)$$

$$\Pi_*(A) = \max_{\substack{B \subseteq A \\ B \neq \emptyset}} \beta(B) \quad (111)$$

et

$$N^*(A) = \min_{B \subseteq \bar{A}} 1 - \beta(B) = 1 - \Pi_*(\bar{A}) \quad (112)$$

$$N_*(A) = \min_{B = \bar{A} \cap B \neq \emptyset} 1 - \beta(B) = 1 - \Pi^*(\bar{A}) \quad (113)$$

On voit facilement que $\max(\Pi^*(A), \Pi_*(\bar{A})) = 1$

Π^* peut se calculer à partir de $\{\Pi^*(\omega) = \max_{\omega \in B} \beta(B)\}$, dans le cas fini.

Ce n'est bien sûr pas le cas de Π_* : par exemple si $\{\Gamma_2(\omega) \mid \omega \in \Omega_1\}$ ne contient pas de singleton, alors $\Pi_*(\omega) = 0 \quad \forall \omega \in \Omega_2$.

De par (110) on voit qu'on peut toujours générer une distribution de possibilité à partir d'une famille \mathcal{F} de sous-ensembles d'un référentiel fini Ω valués par des poids β tels que $\max_{A \in \mathcal{F}} \beta(A) = 1$

1.5 - COMBINAISON D'INFORMATIONS INCERTAINES

Un aspect essentiel de toute théorie de l'incertitude concerne la façon dont on construit une fonction d'incertitude à partir d'informations émanant de sources indépendantes exprimées elles aussi sous la forme de fonctions d'incertitude. Ce domaine a été fort peu exploré. Le seul résultat classique bien connu, c'est le conditionnement des probabilités et le théorème de Bayes. Dempster /4/ et Shafer /29/ ont généralisé le conditionnement au cas de la combinaison de deux probabilités supérieures et inférieures. On propose ici une revue des résultats connus et des directions de recherche.

1.5.1 - Conditionnement

Le concept de conditionnement revient au processus suivant : transformer une mesure d'incertitude g définie sur Ω en une mesure $g(.|U)$ définie sur un sous-ensemble U de Ω qui, de source indépendante, contient de façon certaine l'événement que l'on cherche à caractériser.

Ce processus revient à savoir renormaliser g sur U de façon à conserver le même être mathématique. Dans le cas d'une mesure de probabilité P sur Ω on a

$$\forall A \quad P(A|U) = \frac{P(A \cap U)}{P(U)} \quad (114)$$

$P(.|U)$ est une mesure de probabilité sur U .

Dans le cas d'une pseudo-probabilité g de générateur normalisé $\hat{\varphi}$, il faut une version "distordue" de (114) qui est

$$\begin{aligned} \forall A, \quad g(A|U) &= \hat{\varphi}^{-1}(P_{\varphi}(A \cap U) / P_{\varphi}(U)) \\ &= \hat{\varphi}^{-1}(\varphi(g(A \cap U)) / \varphi(g(U))) \end{aligned} \quad (115)$$

$g(.|U)$ et g sont de même nature.

Dempster (1967) et Shafer ont montré que (114) était encore valide pour une probabilité supérieure, i.e.

$$P^*(A|U) = \frac{P^*(A \cap U)}{P^*(U)} \quad (116)$$

ce qui donne pour la probabilité inférieure

$$P_{*}(A|U) = 1 - P^*(\bar{A}|U) = \frac{P_{*}(A \cup \bar{U}) - P_{*}(\bar{U})}{1 - P_{*}(\bar{U})} \quad (117)$$

(116) et (117) peuvent s'utiliser pour conditionner des fonctions de possibilité et de nécessité. (116) et (117) revient à définir la distribution $\pi(.|U)$ par

$$\pi(\omega_i | U) = \frac{\pi_i}{\Pi(U)} \quad (118)$$

Si on applique (116) à une mesure g_{λ} de Sugeno ($\lambda \in]-1, 0]$) on obtient une mesure g_{λ} , avec $\lambda' = \lambda g_{\lambda}(U)$; mais les possibilités supérieures basées sur

la co-norme $\star = a+b-ab$ ne sont pas invariantes par (116). Plus généralement, les pseudo-possibilités autres que les possibilités se prêtent mal à l'utilisation de (114) ou de (115). On peut, dans le cas fini renormaliser la densité par (118) en remplaçant Π par la mesure appropriée.

1.5.2 - Combinaison d'informations incertaines

Dans le cadre des fonctions d'incertitude induites par une fonction multivoque et une mesure de probabilité, Dempster /4/ a suggéré une règle de combinaison de deux probabilités de base m_1 et m_2 en une nouvelle probabilité de base m telle que

$$\forall A \neq \emptyset, m(A) = \frac{1}{K} \sum_{B_i \cap C_j = A} m_1(B_i) m_2(C_j) \quad (119)$$

où K est un coefficient de normalisation valant :

$$1 - \sum_{B_i \cap C_j = \emptyset} m_1(B_i) m_2(C_j)$$

on peut vérifier les propriétés suivantes :

→ si m_1 est telle que $m_1(U) = 1$ pour $U \subseteq \Omega$ alors la règle de Dempster redonne (116) et (117), c'est-à-dire que (119) généralise le conditionnement ;
si $U = \Omega$, alors $m = m_2$.

→ la règle est symétrique; on ne privilégie pas une information par rapport à l'autre.

→ si m_1 ou m_2 est défini sur des singletons seulement (e.g. $P_{1\star} = P_1^*$ = probabilité) alors m définit aussi une fonction de probabilité. Si m_1 et m_2 sont des probabilités alors les probabilités des événements élémentaires sont :

$$\forall i, p_i = p_i^{(1)} \cdot p_i^{(2)} / \sum_{j=1}^n p_j^{(1)} \cdot p_j^{(2)} \quad (120)$$

→ si m_1 et m_2 définissent des fonctions de possibilité alors ce n'est plus le cas en général pour m ; en effet si $\{A_i\}_i$ et $\{B_j\}_j$ sont deux familles d'ensembles monotones, $\{A_i \cap B_j\}_{i,j}$ n'en est pas une (cf /11/)

La renormalisation dans (119) n'est pas nécessaire pour la théorie. Elle a d'ailleurs été critiquée par Zadeh /35/. En effet elle suppose qu'on ignore les éventuels conflits entre m_1 et m_2 synthétisés par la quantité K qui si elle est trop grande, rend douteuse l'équation (119). En revanche on peut très bien admettre, dans la définition de la probabilité de base que $m(\emptyset) \geq 0$ mesure une certaine défiance à l'égard de la fonction d'incertitude. On peut alors poser (Dubois et Prade /11/).

$$\forall A, \bar{m}(A) = \sum_{B_i \wedge C_j = A} i_j m_1(B_i) m_2(C_j) \quad (121)$$

Avec cette définition ; et dans le cas de deux distributions de possibilité $\pi^{(1)}$ et $\pi^{(2)}$, on peut montrer que (121) génère une probabilité supérieure telle que :

$$\forall \omega_i, P^*(\{\omega_i\}) = \pi_i^{(1)} \cdot \pi_i^{(2)} \quad (122)$$

Enfin dans le cas de mesures de possibilité définies par une fonction $\beta: 2^\Omega \rightarrow [0, 1]$ quelconque, éventuellement normalisée ($\max_A \beta(A) = 1$), on peut penser à un analogue de la règle modifiée de Dempster qui construirait β à l'aide β_1 et β_2 par

$$\forall A, \beta(A) = \max_{A=B_i \wedge C_j} \min(\beta_1(B_i), \beta_2(C_j)) \quad (123)$$

si $\pi^{(1)}$ et $\pi^{(2)}$ sont les distributions de possibilité associées à β_1 et β_2 respectivement (cf. 1.4.5), et π est associé à β , alors

$$\forall i, \pi_i = \min(\pi_i^{(1)}, \pi_i^{(2)}) \quad (124)$$

1.6 - PONTS ENTRE POSSIBILITE ET PROBABILITE

1.6.1 - Probabilité et possibilité cohérentes

Probabilité et possibilité sont deux mesures d'incertitude dont la sémantique est profondément différente. Néanmoins ces deux notions ne sont pas complètement indépendantes. Si on ajoute des contraintes sur l'occurrence d'un événement, il tendra à arriver moins fréquemment. Avant d'être probable, un événement doit être possible, et il doit être probable pour être certain.

Un index de cohérence entre une mesure de probabilité et une mesure de possibilité a été suggéré par Zadeh /34/ dans le cas fini, soit $p = (p_1 \dots p_m)$ avec $\sum p_i = 1$, et $\pi = (\pi_1 \dots \pi_n)$ avec $\max_i \pi_i = 1$ alors le degré de cohérence entre π et p vaut

$$\gamma = \sum_{i=1}^n \pi_i p_i \in [0, 1] \quad (125)$$

La cohérence maximale $\gamma = 1$ signifie $p_i > 0 \rightarrow \pi_i = 1$ ce qui est très rigide. Une condition moins dure mais assez naturelle au vu de la section 1.4, et qui coïncide avec le sens commun, est

$$\forall A \subseteq \Omega \quad N(A) = 1 - \pi(\bar{A}) \leq P(A) \leq \pi(A) \quad (126)$$

Cette condition est vérifiée par toutes les mesures de probabilité compatibles avec $P^* = \Pi$.

Dubois /6/ a cherché à caractériser les fonctions de possibilité satisfaisant (126) pour une mesure de probabilité donnée, dans le cas fini. Une condition suffisante est

$$\forall i, \pi_i \geq P(C(p_i)) = \sum_{\omega_k \in C(p_i)} p_k \quad (127)$$

où

$$C(p_i) = \{\omega_k \mid p_k \leq p_i\}$$

Si par ailleurs on impose que π et p aient une forme semblable, c'est-à-dire, $\forall i, \forall k$

$$p_i \geq p_k \leftrightarrow \pi_i \geq \pi_k$$

alors (126) et (127) sont équivalents.

Pour construire π à partir de p on trouve deux approches principalement :

Shafer /29/ a suggéré la transformation ponctuelle suivante

$$\pi_i = \frac{p_i}{\max_k p_k} \quad (128)$$

Malheureusement cette méthode ne fournit pas toujours une fonction de possibilité cohérente avec la probabilité (voir Dubois et Prade /7/, /11/) pour des contre-exemples).

Une autre transformation plus satisfaisante est décrite par Dubois et Prade /8/. Elle définit la nécessité d'un événement A comme l'excès total de probabilité par rapport à celle du plus probable événement élémentaire disjoint de A soit :

$$N(A) = \sum_{\omega \in A} \max(p(\omega) - p_A, 0) \quad (129)$$

où

$$p_A = \max_{\omega \notin A} p(\omega)$$

On montre que la fonction N vérifie bien les axiomes d'une fonction de nécessité. La distribution de possibilité correspondante est

$$\pi_i = \sum_{j=1}^n \min(p_i, p_j) \quad (130)$$

De plus π et p sont compatibles au sens de (126) et de forme semblable. Cette transformation est bijective, et son inverse n'est autre que la méthode d'approximation par une possibilité, d'un couple de probabilités supérieure et inférieure telle qu'elle est décrite en section 1.4.2, à savoir l'équité répartition des poids d'évidence $m(A)$, $\forall A$, sur les singletons des éléments focaux. Ceci confère un caractère canonique à la transformation (130).

Un argument supplémentaire peut être encore avancé. Kaufmann /18/ et Yager /33/ ont suggéré une façon d'effectuer des tirages aléatoires au sens d'une fonction de possibilité qui revient à la procédure suivante :

- tirer un élément focal au hasard, selon la probabilité de base
- tirer au hasard un représentant de l'élément focal ainsi choisi

On peut montrer que si p et π sont liés par (130) les résultats obtenus par tirage au sens de la probabilité p et de la possibilité π sont les mêmes.

L'intérêt de cette transformation est de fournir un moyen de construire des distributions de possibilité à partir d'histogrammes. Les mesures d'incertitudes sur les événements que l'on obtient ainsi sont des bornes supérieures ou inférieures des valeurs de probabilité. Dans ce sens l'information qu'on en retire est plus pauvre que celle que l'on peut espérer d'un calcul probabiliste. Néanmoins, les variables "possibilistes" se prêtent beaucoup plus facilement au calcul (par exemple au calcul arithmétique) que les variables aléatoires. Ce dernier point sera beaucoup développé dans la suite. (cf. chap. 4)

1.6.2 - Epreuves imprécises

L'élaboration d'histogrammes sous-tend toujours l'hypothèse que le résultat d'un tirage aléatoire est précis. Cette hypothèse n'est pas toujours vérifiée : ce que fournit un instrument de mesure, ce sont généralement des intervalles d'erreur ; les sondages d'opinion eux aussi donnent lieu à des réponses imprécises. On décrit ici comment, à partir d'informations statistiques imprécises, on construit très naturellement une fonction de possibilité et non une fonction de probabilité, quand les informations disponibles sont cohérentes.

L'exposé résume Dubois et Prade /13/; voir aussi l'annexe 1 de Dubois et Prade /14/.

On suppose qu'on reçoit l'information sous la forme d'un ensemble d'intervalles $\{[a_k, b_k] \mid k=1, q\}$. En général, puisque l'information est imprécise, on lui demandera d'être un minimum cohérente c'est-à-dire :

$$\exists [a, b] = \bigcap_{k=1, q} [a_k, b_k] \neq \emptyset \quad (131)$$

Il est naturel de recueillir alors les données sous la forme suivante : on se définit un ensemble d'intervalles $\{I_i \mid i = 1, n\}$ emboîtés tels que

$$[a, b] \subseteq I_1 \subset I_2 \dots \subset I_n = \bigcap_{k=1, q} [a_k, b_k]$$

Ces intervalles servent de gabarit pour les $[a_k, b_k]$, et jouent le même rôle que les intervalles disjoints utilisés pour l'élaboration d'histogrammes. Chaque résultat d'expérience $[a_k, b_k]$ est affecté à un unique intervalle I_j qui est le plus petit I_j qui puisse le contenir. On définit alors la quantité

$$m(I_j) = \frac{\text{nombre de résultats affectés à } I_j}{\text{nombre de résultats } (=q)}, \forall j \quad (132)$$

il est clair que m est une probabilité de base, que $\{I_i | i=1, n\}$ sont des éléments focaux emboîtés. On définit donc par là des fonctions de possibilité et de nécessité qui fournissent des bornes aux valeurs de probabilité que l'on aurait obtenues si les résultats des expériences avaient été précis.

On retrouve donc l'idée fondamentale que les probabilités sont un modèle d'informations incertaines précises et dispersées, tandis que les fonctions de possibilité rendent mieux compte d'informations imprécises et cohérentes.

1.7 - POSSIBILITE ET PROBABILITE QUALITATIVES

Savage /26/ reconnaissant que les probabilités subjectives (= La théorie des probabilités utilisée comme modèle de l'incertitude subjective) n'étaient pas satisfaisantes. Au lieu de relaxer la propriété d'additivité il a suggéré une approche purement ordinales probabilités qualitatives. Il s'agit de définir sur l'ensemble des événements une relation, notée $\underline{\geq}$, telle que $A \underline{\geq} B$ signifie A est plus probable que B .

Avec les conventions

$$\begin{aligned} A > B &\leftrightarrow A \underline{\geq} B, \quad \text{et} \quad B \underline{\geq} A \quad \text{est faux} \\ A \sim B &\leftrightarrow A \underline{\geq} B, \quad \text{et} \quad B \underline{\geq} A \end{aligned}$$

Les axiomes suivants ont été suggérés:

- C0) $\Omega > \emptyset$ (non trivialité)
- C1) $A \underline{\geq} B$ ou $B \underline{\geq} A$ (comparabilité)
- C2) $A \underline{\geq} B, B \underline{\geq} C \rightarrow A \underline{\geq} C$ (transitivité)
- C3) $A \underline{\geq} \emptyset$ (l'événement impossible est le moins probable)
- C4) $A \cap (B \cup C) = \emptyset \rightarrow (B \underline{\geq} C \leftrightarrow A \cup B \underline{\geq} A \cup C)$

L'axiome (C4) permet des déductions par simplification.

Une fonction numérique d'ensemble g est compatible avec $\underline{\geq}$ si et seulement si

$$\forall A, \forall B \quad A \underline{\geq} B \leftrightarrow g(A) \underline{\geq} g(B) \quad (133)$$

Une conséquence de C4 est

$$A \underline{\geq} B \rightarrow \bar{B} \underline{\geq} \bar{A} \quad (134)$$

$$\text{On en déduit } A \sim \Omega \leftrightarrow \bar{A} \sim \emptyset \quad (135)$$

Cette dernière propriété montre que ce qui est modélisé a bien un comportement probabiliste ; puisque (135) est compatible avec $P(A) + P(\bar{A}) = 1$.

Même dans le cas fini, il n'existe pas toujours une mesure de probabilité compatible avec $\underline{\geq}$. C0-C4 sont insuffisants et on n'a pas encore trouvé d'axiome supplémentaire indiscutable (voir Fine /15/).

En revanche, une probabilité qualitative n'est généralement pas compatible avec une fonction de possibilité. C'est l'axiome C4 qui est à mettre en cause puisque, même si $A \wedge (B \cup C) = \emptyset$

$$\Pi(A \cup B) \underline{\geq} \Pi(A \cup C) \not\Rightarrow \Pi(B) \underline{\geq} \Pi(C)$$

Pour élargir la notion de probabilité qualitative, on définit la notion de mesure floue qualitative en remplaçant C4 par une version affaiblie :

$$C'4) \forall A, B \underline{\geq} C \rightarrow A \cup B \underline{\geq} A \cup C$$

Dans ce cas C0-C3 et C'4 suffisent à montrer

$$A \underline{\subset} B \rightarrow B \underline{\geq} A \quad (136)$$

Donc, dans le cas fini, il existe toujours une mesure floue compatible avec une mesure floue qualitative.

Pour obtenir une relation $\underline{\geq}$ cohérente avec l'idée de possibilité on peut adjoindre à C0-C3, C'4 l'axiome:

$$C5) A \underline{\geq} B \rightarrow A \cup B \sim A$$

Comme conséquence, on déduit $A \sim \Omega$ ou $\bar{A} \sim \Omega$, et, dans le cas fini, $\forall A \neq \emptyset \exists \omega \in A, \{\omega\} \sim A$; cet élément est noté $\text{sup}(A)$. En particulier, $\exists \omega \in \Omega, \{\omega\} \sim \Omega$. De plus il est facile de montrer que C5 implique C'4 de sorte que les possibilités qualitatives sont entièrement caractérisées par C0-C3 et C5, et sont des mesures floues qualitatives.

Dans le cas fini toute possibilité qualitative est compatible avec une fonction de possibilité.

La relation duale de la possibilité qualitative se définit par $\underline{\geq}$ telle que

$$A \underline{\geq} B \leftrightarrow \bar{B} \geq \bar{A} \quad (137)$$

On en déduit facilement que l'équivalence associée est

$$A \underline{\sim} B \leftrightarrow \bar{A} \sim \bar{B}$$

La relation $\underline{\geq}$ vérifie C0-C3 et l'axiome dual de C5

$$C6) \quad A \underline{\geq} B \rightarrow A \wedge B \underline{\sim} B$$

Et on déduit $\emptyset \underline{\sim} A$ ou $\emptyset \underline{\sim} \bar{A}, \forall A$, et par dualité $\exists \omega \in \Omega : \{\omega\} \underline{\sim} \emptyset$

On vérifie par dualité que C6 implique C'4 et que la relation $\underline{\geq}$ est bien une mesure floue qualitative, que l'on peut légitimement appeler "nécessité qualitative".

Dans le cas fini, l'équivalent de la distribution de possibilité sera la relation $\underline{\geq}$ restreinte aux éléments de Ω , telle que

$$\omega \underline{\geq} \omega' \leftrightarrow \{\omega\} \underline{\geq} \{\omega'\}$$

Cette relation est un ordre total sur Ω . On vérifie $A \underline{\geq} B \leftrightarrow \sup A \underline{\geq} \sup B$. Pour construire une distribution de possibilité qui génère des fonctions de possibilité (resp : nécessité) compatibles avec $\underline{\geq}$ (resp : $\underline{\geq}$) il suffit de poser

$$\begin{aligned} \pi(\omega) &= 1 && \forall \omega \sim \Omega \\ \pi(\omega) &= \pi(\omega') && \forall \omega \sim \omega' \\ \pi(\omega) &> \pi(\omega') && \forall \omega > \omega' \\ \pi(\omega) &= 0 && \forall \omega \sim \emptyset \end{aligned}$$

La relation = "un événement est d'abord possible avant d'être nécessaire" s'écrit ici

$$A \underline{\sim} \Omega \rightarrow A \sim \Omega$$

On voit donc que similairement à Savage /26/ on peut définir la relation "plus possible (= plus facile à réaliser) que" abstraitement sans faire référence à une échelle numérique. Le caractère fondamentalement ordinal des fonctions de possibilité rend même cette démarche plus aisée.

1.8 - CONCLUSION

Cette section a donc mis en évidence les fonctions de probabilité et de possibilité comme deux concepts distincts ayant chacun leur sémantique dans l'optique de la modélisation de l'incertitude. On a mis en évidence deux familles de mesures monotones qui généralisent à la fois probabilité et possibilité : les mesures décomposables qui conservent l'idée d'information ponctuelle mais où la façon de combiner les informations sur des événements disjoints n'est plus forcément additive. Les mesures de Dempster et Shafer qui rejettent l'idée d'information ponctuelle mais préservent un caractère additif. Les premières peuvent servir de modèles descriptif à la façon (non forcément additive) dont l'esprit humain combine des informations incertaines : tantôt de façon probabiliste mais avec des distorsions, tantôt à l'aide de catégories de type possibilité/nécessité qui n'obéissent plus du tout aux mêmes règles. Les mesures de Dempster rendent compte d'informations incertaines objectives et permettent de relier les fonctions de possibilité à des expériences statistiques, même si, comme Shafer l'indique, elles peuvent elles-aussi être à la base de modèles de probabilités subjectives. Les fonctions de possibilité émergent donc à la fois comme modèles d'information subjective (raisonnements à partir des catégories "possible" et "nécessaire") ou objective (traitement de données statistiques imprécises et cohérentes). Une telle fonction de possibilité définit implicitement un ensemble mal délimité de valeurs possibles que peut prendre une variable. C'est de ce type d'ensemble que traitera la section suivante.

REFERENCES

- /1/ BANON G. (1981)
Distinction between several subsets of fuzzy measures.
Fuzzy Sets and Systems. Vol. 5 n°3, pp.291-306.
- /2/ BANON G. (1979)
Utilisation de familles de mesures de possibilité comme modèles d'information a priori subjective dans l'estimation paramétrique. Note interne 79-I-33, LASS-CNRS, Toulouse.
- /3/ CHOQUET G. (1953)
Theory of capacities, Ann. Institut Fourier, 5, 131-295.
- /4/ DEMPSTER A.P. (1967)
Upper and lower probabilities induced by a multivalued mapping.
Ann. Math. Statist. Vol.38, pp.325-339.
- /5/ DUBOIS D. (1980)
Un économiste précurseur de la théorie des possibilités : G.L.S. Shackle. BUSEFAL n°2, Printemps 1980, L.S.I. Univ. P. Sabatier, Toulouse, pp.70-73.
- /6/ DUBOIS D. (1980)
Sur les liens entre les notions de probabilité et de possibilité (quelques remarques). Table Ronde CNRS "Quelques Applications Concrètes Utilisant les Derniers Perfectionnements de la Théorie du Flou", Lyon, 23-24-25 Juin 1980 (10 p.).
- /7/ DUBOIS D. (1980)
Triangular norms for fuzzy sets. Proc. 2d. International Seminar on Fuzzy set Theory, (E.P. Klement, Ed.) 39-68, J. Kepler Universität, Linz, Autriche.
- /8/ DUBOIS D., PRADE H. (1981)
Unfair coins and necessity measures. Proc. 3rd Int. Seminar On Fuzzy Set Theory, Johannes Kepler Univ. Linz, Austria, (E.P. Klement, ed.) pp.223-231, Version révisée dans Fuzzy Sets and Systems, vol. 10, n°1, 1983, 15-20

- /9/ DUBOIS D., PRADE H. (1980)
Fuzzy Sets and Systems : Theory and Applications Academic Press, New-York.
- /10/ DUBOIS D., PRADE H. (1982)
A class of fuzzy measures based on triangular norms - A general framework for the combination of uncertain information. Int. J. of General Systems, Vol.8, n°1, pp.43-61.
- /11/ DUBOIS D., PRADE H. (1982)
On several representations of an uncertain body of evidence. Dans "Fuzzy Information and Decision Processes" (M.M. Gupta, E. Sanchez, eds.) North Holland. 167-182.
- /12/ DUBOIS D., PRADE H. (1982)
Upper and lower possibilistic expectations and their applications. Proc. 4th International Seminar on fuzzy Set Theory, J. Kepler Universität, Linz, Autriche.
- /13/ DUBOIS D., PRADE H. (1982)
Fuzzy Sets and Statistical Data, présenté à EURO V, Lausanne, Suisse.
- /14/ DUBOIS D., PRADE H. (1983)
Upper and lower possibilities induced by a multivalued mapping, IFAC Symposium on Fuzzy Information Knowledge Representation and Decision Analysis, Marseille
- /15/ FINE T. (1973)
Théories of Probability Academic Press, New York.
- /16/ HÖHLE H. (1982)
Integration with respect to fuzzy measures. IFAC Symposium on Theory and Applications of Digital Control, New Delhi.
- /17/ HUGHES G.E., CRESWELL M.J. (1972)
An introduction to modal logic. Methuen, Londres.
- /18/ KAUFMANN A. (1980)
La simulation des ensembles flous, Table ronde CNRS, Université de Lyon, France.

- /19/ KENDALL D.G. (1974)
Foundations of a theory of random sets. In : Stochastic Geometry
(E.F. Harding, D.G. Kendall, eds) J. Wiley & Sons, 322-376.
- /20/ KRUSE R. (1982)
A note on λ -additive fuzzy measures. Fuzzy Sets and Systems,
8, 218-221.
- /21/ LOEVE M. (1960)
Probability theory Van Nostrand, Princeton - N.J.
- /22/ MATHERON G. (1975)
Random Sets and Integral Geometry, J. Wiley
- /23/ PRADE H. (1979)
Nomenclature of fuzzy measures Proc. 1st Int. Seminar on
Fuzzy Set Theory (E.P. Klement Ed.) J. Kepler Universität
 Linz, Autriche - 7-25.
- /24/ P DE H. (1980)
Une approche des mesures floues basée sur les normes traingou-
laires, Table Ronde CNRS, Université de Lyon, France.
- /25/ PURI M.L., RALESCU D. (1982)
A possibility measure is not a fuzzy measure. Fuzzy Sets and
Systems - 7, 311-314.
- /26/ SAVAGE L.J. (1972)
The foundations of statistics, Dover, New-York
- /27/ SCHWYHLA W. (1980)
About the isomorphism between some sugeno-measures and classical
measures. Proc. 2d Int. Seminar on Fuzzy Set Theory, J. Kepler
Universität, Linz, Autriche 143-176.
- /28/ SHACKLE G.L.S. (1961)
Decision, order, and time in human affairs. Cambridge University
 Press, U.K. (Traduction : *Décision, Déterminisme et temps* Dunod
 1971).
- /29/ SHAFER G. (1976)
A mathematical Theory of evidence. Princeton University Press,
 Princeton N.J.

- /30/ SUGENO M. (1974)
Theory of Fuzzy Integral and Its Applications. Ph. D. Thesis,
 Tokyo Inst. of Technology, Japon.
- /31/ TRILLAS E. (1979)
 Sobre funciones de negacion en la teoria de conjuntos
 borrosos, Stochastica III n°1, 47-59.
- /32/ TSUKAMOTO Y., GUPTA M.M., NIKIFORUK P.N. (1982)
 On density of λ -fuzzy measure, in : "Fuzzy Set and Possibility
 Theory" (R. Yager, ed.) Pergamon Press 419-428.
- /33/ YAGER R.R. (1982)
 Level-sets for membership evaluation of fuzzy sets, *ibidem*,
 90-97.
- /34/ ZADEH L.A. (1978)
 Fuzzy Sets as a basis for a theory of possibility. Fuzzy sets
 and systems 1, 3-27.
- /35/ ZADEH L.A. (1979)
On the validity of Dempster's rule of combination of evidence.
 Memo. UBC/ERL n°79/24, Univ. of California Berkeley.

AUTRES REFERENCES

- /36/ FERON R. (1976)
 Ensembles aléatoires flous. C.R.A.S. Paris, Vol.282 série A,
 903-906 (une extension des ensembles aléatoires au cas d'élé-
 ments focaux eux-mêmes flous).
- /37/ HISDAL H. (1978)
 Conditional possibilities, independence and non-interaction
Fuzzy Sets and Systems 1, 283-298 (Remarques sur les analogies
 de structure entre possibilité et probabilité).
- /38/ HÖHLE H. (1982)
 A mathematical theory of uncertainty in Fuzzy Set and Possi-
 bility Theory : recent developments (R.R. Yager ed.) Pergamon
 press 344-356 (La théorie de Shafer /29/ reformulée dans le
 contexte abstrait des algèbres booléennes).

/47/ WIERZCHON T.Z. (1982)

On fuzzy measure and fuzzy integral ibidem 79-86 (Etude des fonctions g_λ de Sugeno dans le cadre de la théorie de Shafer /29/).

CHAPITRE 2

ENSEMBLES FLOUS

On peut définir le concept d'ensemble flou en dehors de toute référence à une mesure d'incertitude en modifiant la traditionnelle fonction caractéristique d'un ensemble, de façon à introduire des degrés dans la relation traditionnelle d'appartenance. C'est une approche possible. Néanmoins la donnée d'une mesure d'incertitude revient à vouloir localiser la valeur d'une variable x en exprimant pour chaque sous-ensemble A d'un référentiel la connaissance que l'on possède sur la relation $x \in A$. La famille de sous-ensembles pertinents pour la représentation de la variable x va induire une fonction caractéristique généralisée d'un ensemble flou, les deux représentations étant strictement équivalentes seulement dans le cas de mesures de possibilité.

2.1 - REPRESENTATIONS D'UN ENSEMBLE FLOU

Un ensemble flou F est équivalent à la donnée d'un référentiel Ω et d'une application, notée μ_F , de Ω dans $[0, 1]$, l'intervalle unité (Zadeh./80/1965). $\mu_F(\omega)$, pour $\omega \in \Omega$, est interprété comme le degré d'appartenance de ω à l'ensemble flou F . Cette définition est la plus directe. Elle permet une modélisation simple de catégories vagues du langage naturel ("grand" par exemple) définies sur un support objectif qui peut être une échelle numérique ($\Omega =$ ensemble de tailles) ou l'ensemble des objets qui seront qualifiés à l'aide de ces catégories ($\Omega =$ ensemble de personnes). $\mu_F(\omega)$ exprime alors à quel point la valeur (ou l'objet) ω est compatible avec le concept F . Si $\Omega = \mathbb{R}$ (l'ensemble des réels), F est une quantité floue. Dans la suite $F = G$ veut dire $\mu_F = \mu_G \quad \forall F, G \in [0, 1]^\Omega$.

Il est naturel de chercher des représentations ensemblistes usuelles de F . Quand $\mu_F(\omega) \in \{0, 1\} \quad \forall \omega$, F coïncide avec un sous-ensemble usuel de Ω . Si ce n'est pas le cas on peut choisir un seuil $\alpha \in]0, 1]$ et définir l'ensemble

$$F_\alpha = \{\omega \in \Omega \mid \mu_F(\omega) \geq \alpha\} \quad (1)$$

F_α est appelé "coupe de niveau α " ou " α -coupe". F_α contient tous les éléments de Ω qui sont compatibles avec A au moins au niveau α . La famille $C(F) = \{F_\alpha \mid \alpha \in]0, 1]\}$ d' α -coupes est une famille monotone :

$$0 < \alpha \leq \beta \leq 1 \rightarrow F_\alpha \supseteq F_\beta \quad (2)$$

Elle constitue une représentation de F à l'aide d'ensembles classiques puisqu'on a (Zadeh, /83/) :

$$\forall \omega, \mu_F(\omega) = \sup \{ \alpha \mid \omega \in F_\alpha \} \quad (3)$$

Inversement étant donnée une famille finie et monotone d'ensembles $\{F_{\alpha_1} \dots F_{\alpha_m}\}$ affectés de poids α_i satisfaisant (2), elle forme l'ensemble des α -coupes d'un ensemble flou défini par (3). Dans le cas d'une famille infinie d'ensembles, la condition (2) ne suffit plus, il faut aussi que pour toute suite $(\alpha_n)_n$ croissante d'éléments de $]0,1]$

$$\lim_{n \rightarrow +\infty} \alpha_n = \alpha \quad \rightarrow \quad F_\alpha = \bigcap_{n=1}^{\infty} F_{\alpha_n} \quad (4)$$

Ce résultat est dû à Ralescu /63/. Quand la représentation ne contient que deux ensembles emboîtés, on obtient une notion d'ensemble flou suggérée indépendamment par Gentilhomme /21/.

On peut vouloir représenter un ensemble flou à l'aide, non seulement des α -coupes, mais aussi d'autres ensembles. Qualifions l'opportunité de représenter F par un ensemble classique A , par l'indice (Dubois et Prade, /15/)

$$r(A) = \inf_{\omega \in A} \mu_F(\omega) \quad (5)$$

il est facile de vérifier

$$A \subseteq B \rightarrow r(A) \geq r(B)$$

$$r(F_\alpha) = \alpha, \forall \alpha; \quad r(A) \geq \alpha \quad \forall A \subseteq F_\alpha$$

De plus, si on définit le support de F par

$$S(F) = \{\omega \mid \mu_F(\omega) > 0\} \quad (6)$$

alors

$$\forall A, S(F) \subset A \rightarrow r(A) = 0 \quad (7)$$

Par ailleurs si F est un ensemble usuel on souhaite que

$$\forall A \neq F, r(A) = 0 \quad (8)$$

Soit $\overset{\circ}{F} = \{\omega \mid \mu_F(\omega) = 1\}$ le noyau de F , alors on exclut de la représentation les sous-ensembles propres du noyau. On représente alors F par

$$\{(A, r(A)) \mid \overset{\circ}{F} \subseteq \underline{\underline{A}} \subseteq \overline{\overline{A}}\} = R(F)$$

On a encore :

$$\mu_F(\omega) = \sup \{r(A) \mid \omega \in A\} \quad (9)$$

Pour calculer $\mu_F(\omega)$ la représentation par α -coupe suffit. Ce ne sera pas toujours le cas dans d'autres circonstances.

Une autre représentation est implicite au chapitre 1 section 2. Un ensemble $E \subseteq \Omega$ peut être caractérisé par une mesure de possibilité Π_E telle que $\Pi_E(A) = 1$ si $E \cap A \neq \emptyset$ et 0 sinon. Une distribution de possibilité coïncide donc avec la définition d'un ensemble flou. Il est facile de voir que si une fonction de possibilité est définie par une probabilité de base, alors les éléments focaux forment la famille des α -coupes d'un ensemble flou F . Soient $\{A_1 \subseteq \dots \subseteq A_m\}$ ces éléments focaux on a

$$A_i = F_{\alpha_i} \quad \text{où} \quad \alpha_i = \sum_{j=1}^i m(A_j) \quad (10)$$

Ceci est une façon "probabiliste" de représenter un ensemble flou.

De même la fonction r définie par (5) est une "possibilité de base" puisque $r(\overset{\circ}{F}) = 1$ et que (9) exprime la possibilité supérieure de singleton $\{u\}$ au sens de l'équation (1.109) de la section 1.4.5. Néanmoins la fonction r n'est qu'un cas particulier de "possibilité de base".

La donnée d'un ensemble flou suffit à caractériser une fonction de possibilité, pourvu que cet ensemble flou soit normalisé, c'est-à-dire

$$\forall \omega, \mu_F(\omega) = 1$$

Fortet et Kambouzia /18/, Goodman /19/ ont fait remarquer qu'étant donné un ensemble aléatoire (ou dans le cas fini un couple de probabilités supérieure et inférieure) on peut définir un ensemble flou par

$$\mu_F(\omega) = P^*(\{\omega\}) \quad \forall u \quad (11)$$

(voir aussi Kampé de Fariet /34,35/)

Mais, sauf dans le cas des fonctions de possibilité, la connaissance de $P^*(\{\omega\})$ ne suffit pas à reconstituer toute l'information, ce qui limite la portée de cette remarque. De plus on peut très bien avoir $P^*(\{\omega\})=0 \forall \omega$, par exemple si P^* est une mesure de probabilité nonatomique sur Ω . Dans ce cas $F = \emptyset !$.

Quand on applique (11) à une mesure de probabilité sur un référentiel fini, cela revient à considérer des poids de probabilité comme des valeurs d'appartenance ; si on se souvient qu'une fonction de probabilité modélise l'idée d'un point de position inconnue et non d'un ensemble (car si $P(A) \in \{0,1\}$, alors P est une mesure de Dirac), cette assimilation peut sembler abusive.

Dans cette optique Höhle /32/ a montré qu'une fonction de crédibilité (ou probabilité inférieure) sur Ω peut être vue comme un filtre flou, c'est-à-dire un sous-ensemble flou de 2^Ω qui coïncide avec un filtre au sens de la topologie (Bourbaki /3/ ou Neveu /54/, p.9) quand $P_*(A) \in \{0,1\} \forall A \subseteq \Omega$. Un ultra-filtre flou est alors défini en ajoutant la propriété $1 - P_*(A) \geq P_*(\bar{A})$, et coïncide avec une mesure de probabilité. Dans le cas classique, un ultrafiltre correspond à un point, d'où l'interprétation d'une mesure de probabilité comme un "point flou". La relation de compatibilité (Cf § 1.4.2, 1.6.1) $P \leq \Pi$ s'interprète comme l'appartenance du "point flou" P à l'ensemble flou induit de Π .

Un dernier point de discussion concerne l'opportunité de normaliser ou non un ensemble flou. Tout dépend du caractère exhaustif du référentiel. Si on veut représenter l'ensemble des entiers très proches de 3.5, il est naturel de ne pas normaliser la fonction d'appartenance, car les nombres les plus proches de 3.5 sont hors de \mathbb{N} . En revanche si on utilise une fonction de possibilité, alors $\Pi(\Omega) = 1$, ce qui, sauf cas pathologique de peu d'intérêt pratique, implique que la distribution de possibilité soit normalisée. De même, une quantité floue définie sur $\mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ sera en général normalisée, le référentiel étant exhaustif. Ne pas normaliser une fonction de possibilité peut néanmoins s'interpréter comme un manque de confiance dans cette information (par exemple issue de deux sources conflictuelles) ou comme la possibilité de ne pas attribuer de valeur à la variable associée à cette fonction de possibilité (par exemple Ω est l'ensemble des modalités d'une action qu'on n'effectuera peut-être pas). Sur ces points voir Yager /79/

2.2 - DETERMINATION DES FONCTIONS D'APPARTENANCE

Une des questions que posent les personnes qui s'initient à la théorie des ensembles flous est : comment trouve-t-on les fonctions d'appartenance? Il convient de distinguer les cas où F représente l'idée que se fait un individu d'une catégorie vague de ceux où F est déduit de données statistiques.

2.2.1 - Catégorie vague au sens d'un individu

Il faut d'abord distinguer entre les catégories simples définies sur un référentiel-linéaire objectif ("grand") et les catégories complexes qui font intervenir simultanément plusieurs référentiels ("trapu") ou même dont les échelles de référence sont difficilement identifiables ("beau"). On s'intéresse ici aux catégories simples. Dans ce cas, l'évaluation d'une fonction d'appartenance est un exercice de psychométrie (par exemple Krantz et al. /46/). La fonction d'appartenance s'obtient par questionnaire. Fondamentalement une fonction d'appartenance sur Ω définit un ordre sur les éléments de Ω , et c'est cet ordre qui est important. $\omega_1 \geq \omega_2$ veut dire " ω_1 est plus F que ω_2 ". Norwich et Turksen /55/ ont fait le lien entre les théories classiques de psychométrie et l'évaluation d'une fonction d'appartenance : si on sait définir une relation d'ordre au sens large, partout définie, sur Ω , avec un plus grand et un plus petit élément, alors on peut représenter la catégorie F par un ensemble flou dont la fonction d'appartenance, à valeurs sur \mathbb{R} , est unique à une transformation strictement croissante près. Si on veut plus de précision il faut trouver une structure d'ordre plus riche, à savoir un ordre au sens large partout défini sur Ω^2 . Ce qu'on compare alors ce sont des paires (ω_1, ω'_1) et (ω_2, ω'_2) au sens de F , $(\omega_1, \omega'_1) \geq (\omega_2, \omega'_2)$ veut dire " ω_1 est plus F par rapport à ω'_1 que ω_2 ne l'est par rapport à ω'_2 ". Moyennant quelques axiomes de cohérence et de complétude (notamment Ω doit être un continuum), Norwich et Turksen /55/) montrent que la fonction d'appartenance est unique à une transformation affine croissante près, c'est-à-dire est unique sur $[0, 1]$, dès que le support et le noyau de F sont connus. En pratique il est facile d'obtenir une idée grossière de la forme de μ_F , qui suffira pour les applications. Si $\Omega = [a, b]$ on peut par exemple identifier le support $[s_1, s_2]$ et le noyau $[n_1, n_2] \subseteq [s_1, s_2]$. Pour identifier la forme de μ_F sur $[s_1, n_1]$ et $[n_2, s_2]$ on peut recueillir une information sur l'ordre entre

$$\left(\frac{s_i + n_i}{2}, s_i\right) \quad \text{et} \quad \left(n_i, \frac{s_i + n_i}{2}\right) \quad i = 1, 2$$

ce qui permet de discriminer parmi des représentations analytiques types de μ_F . On notera qu'une erreur sur s_i ou n_i aura moins d'importance dans le cas flou (faible évolution de μ_F autour de ces points) que dans le cas d'une représentation non floue (saut de μ_F au passage par ces points). Ce type d'approche simplifiée est décrite dans Kuzmin /47/ par exemple. Des approches plus sophistiquées sont celles de Saaty /66/), et Chu et al. /4/ (à partir des ratios $\mu_F(u)/\mu_F(u')$). Fung et Fu /20/ (comparaison de sous-ensembles représentant F) Mc Vicar Whelan /53/ (le point de transition entre "F" et "non F" comme variable aléatoire) Norwich et Turksen /55/ ($\mu_F(\omega)$ est une variable aléatoire pour tout $\omega \in \Omega$; voir Hirota /30/ pour une étude théorique de ce point de vue.)

Le cas d'une catégorie complexe relève du problème d'aggrégation de sous-ensembles flous (section suivante).

2.2.2 - Ensembles flous construits sur des données statistiques

On peut distinguer deux points de vue.

a) Succédané de probabilité

Dubois et Prade /12/ ont suggéré une approche canonique pour affaiblir une mesure de probabilité P sur U fini en une fonction de possibilité Π de façon à ce que (cf. ch.1, section 6.1)

$$\forall A \subseteq \Omega \quad \Pi(A) \geq P(A) \geq N(A) = 1 - \Pi(\bar{A}) \quad (12)$$

La distribution de possibilité associée est un ensemble flou tel que

$$\forall \omega, \mu_F(\omega) = \Pi(\{\omega\}) = \sum_{\omega' \in \Omega} \min(p(\omega), p(\omega')) \quad (13)$$

Dans le cas où $\Omega = \mathbb{R}$ et d'une densité de probabilité p, (13) devient

$$\forall \omega, \mu_F(\omega) = \int_{-\infty}^{+\infty} \min(p(\omega), p(s)) ds \quad (14)$$

Il est facile de voir que si $f_n(\omega)$ est la fréquence de l'événement ω après n observations, $\varphi_n(\omega)$ la quantité obtenue en substituant f_n à p dans (13),

alors

$$\lim_{n \rightarrow +\infty} f_n(\omega) = p(\omega) \rightarrow \lim_{n \rightarrow +\infty} \varphi_n(\omega) = \mu_F(\omega) \quad (15)$$

On peut donc construire une distribution de possibilité satisfaisant (14) à partir d'un histogramme, après un nombre suffisant d'observations. On notera que si après un nombre n_0 d'observations, l'histogramme induit toujours le même ordre sur Ω ($f_n(\omega_1) \geq f_n(\omega_2) \dots \geq f_n(\omega_{|\Omega|})$) il devient licite de l'utiliser au sens des possibilités parce que l'ordre induit sur les événements $A \subseteq \Omega$ par Π sera fixe $\forall n \geq n_0$. (ce qui est faux si on utilise des probabilités). Rappelons encore, au crédit de (13) qu'une simulation de l'ensemble flou F et un tirage aléatoire au sens de P fourniront les mêmes résultats.

b) Statistiques imprécises

On a vu au ch. 2 section 6.2 qu'un ensemble de données imprécises mais cohérentes généraient une fonction de possibilité (voir aussi Dubois & Prade /16/). Donc un ensemble flou pourra naturellement résumer l'information fournie par des instruments de mesure imprécis ; on peut aussi identifier une catégorie subjective simple au sens d'un groupe de personnes, chaque personne fournissant une représentation non floue de la catégorie. Le processus d'agrégation est décrit en 4.6.2.

REMARQUE : La question de l'évaluation de la fonction d'appartenance est liée au choix de l'ensemble des valeurs d'appartenance ; la discussion précédente suppose toujours $\mu_F(\omega) \in [0,1]$, ce qui implique que l'on puisse toujours ordonner linéairement (au sens large) les éléments de Ω du point de vue de l'adéquation à la catégorie F . Certains auteurs ont pris des hypothèses plus faibles : Goguen /22/ a remplacé $[0,1]$ par un treillis L ; Zadeh /86/ a suggéré pour exprimer l'incertitude sur $\mu_F(\omega)$ de considérer les valeurs d'appartenance comme des ensembles flous sur $[0,1]$. On obtient des ensembles flous dits de type 2. μ_F est alors à valeurs dans le treillis $[0,1]^{[0,1]}$. Ce processus peut être itéré à l'ordre $n \in \mathbb{N}$ (≥ 1) quelconque⁽¹⁾. Inversement si c'est l'élément ω de valeur d'appartenance $\mu_F(\omega)$ qui s'avère être imprécis,

(1) voir Mizumoto et Tanaka /50/.

Ω est en fait un ensemble de parties floues d'un autre univers V , ($\Omega = [0, 1]^V$) et F est un ensemble flou de niveau 2 (Zadeh /84/) sur V . Cette notion permet de modéliser des catégories vagues de plus en plus abstraites (par exemple, la couleur vue comme ensemble flou sur {noir, gris, rouge,...}) voir Goguen /24/ et Gottwald /26/ pour une étude systématique de cette notion qui se généralise à l'ordre p . Enfin, la notion d'ensembles flous de Gentilhomme /21/ comme paire d'ensembles emboîtés peut s'étendre à des paires d'ensembles flous de Zadeh (cf. chap.4).

2.3 - OPERATEURS ENSEMBLISTES FLOUS

Dans cette section on définit de façon axiomatique les extensions ponctuelles des opérateurs ensemblistes aux ensembles flous. Ces extensions ne sont pas uniques, même si on peut classer les opérateurs suivant les propriétés qu'ils préservent. De plus il y a des opérateurs ensemblistes flous qui ne correspondent à aucun opérateur ensembliste classique, bien qu'on connaisse, déjà certains d'entre eux par ailleurs (par exemple les moyennes).

2.3.1 - Opérateurs de complémentation

Un opérateur ponctuel de complémentation est une fonction C de $[0, 1]$ dans $[0, 1]$ telle que le complémentaire de l'ensemble flou F , dénoté \bar{F} soit défini par

$$\forall \omega \in \Omega, \mu_{\bar{F}}(\omega) = C(\mu_F(\omega)) \quad (16)$$

Les conditions suivantes doivent être imposées à C pour préserver l'intuition que l'on possède sur l'idée de complémentation :

- C1) $C(0) = 1$; $C(1) = 0$ (coïncidence avec le cas classique)
- C2) C est strictement décroissante (si on augmente le degré d'appartenance à F en passant de u à u' , on diminue par là celui relatif à \bar{F})
- C3) C est involutif (la double négation est une affirmation)

NB : En logique intuitioniste, cette dernière condition est omise.

Dès que C est continu, il existe un seuil unique $s \in]0, 1[$ tel que $s = C(s)$, qui dépend de C et qui fixe en quelque sorte le seuil d'appartenance à F connaissant C . Trillas /74/ a résolu les équations fonctionnelles définissant C

(voir Annexe), sous l'hypothèse de continuité, sous la forme

$$C(x) = \varphi^{-1}(1 - \varphi(x)) \quad (17)$$

où φ est une fonction croissante réelle telle que $\varphi(0) = 0$, $\varphi(1) = 1$ on voit facilement que $s = \varphi^{-1}(\frac{1}{2})$ est unique quand C est fixé. Si on ajoute la condition

$$\forall (\omega, \omega'), \mu_F(\omega) - \mu_F(\omega') = \mu_{\bar{F}}(\omega') - \mu_{\bar{F}}(\omega) \quad (18)$$

C'est-à-dire ω est par rapport à ω' autant F que ω' est \bar{F} par rapport à ω , alors (Bellman et Giertz /2/)

$$\mu_{\bar{F}}(\omega) = 1 - \mu_F(\omega) \quad (19)$$

ce qui est la plus simple définition possible de la négation ($\varphi(x)=x$) proposée par Zadeh.

Il y a d'autres classes plus générales de compléments, obtenues en relaxant certains des axiomes proposés ci-dessous : C2 affaibli en décroissance au sens large (Esteve et al./17/), rejet de l'involution (Higashi et Klir, /30/), étude spécifique des involutions (Ovchinnikov /58/), compléments dépendant de l'élément considéré (Lowen, /49/).

Notons qu'on n'a jamais $\bar{F}_\alpha = (\bar{F})_\alpha$ mais avec (19) on a

$$\forall \alpha \in]0, 1], (\bar{F})_\alpha = \bar{F}_{1-\alpha} \quad (20)$$

où F_α est la coupe stricte $\{\omega | \mu_F(\omega) > \alpha\}$.

REMARQUE : Autres opérations unaires : Certaines transformations de la fonction d'appartenance ont une interprétation en termes de modificateurs de catégories subjectives. Connaissant μ_F où F signifie par exemple "grand", on peut calculer $\mu_{F'}$ où F' signifie "assez grand", "très grand" etc... à l'aide d'une fonction croissante $f = [0, 1] \rightarrow [0, 1]$ telle que $f(0)=0$, $f(1)=1$, et d'une éventuelle translation sur l'échelle des tailles Ω de sorte que

$$\mu_{F'}(\omega) = f \circ \mu_F(\omega+a) \quad (21)$$

voir Zadeh /85/ ou Lakoff /48/ ..

Une autre transformation remarquable, duale de la négation est l'antonymie, définie sur un référentiel $\Omega = [a, b] \subset \mathbb{R}$ par

$$\forall \omega, \mu_{\text{ant}F}(\omega) = \mu_F(b - (\omega - a)) \quad (22)$$

si F signifie "grand", $\text{ant}F$ signifie "petit" (\neq "non grand")

2.3.2 - Opérateurs d'intersection et d'union : approche axiomatique

Les opérateurs d'union et d'intersection d'ensembles flous peuvent être ponctuellement définis à l'aide d'applications de $[0, 1]^2$ dans $[0, 1]$, notées u et i , telles que

$$\forall \omega \in \Omega, \mu_{F \cup G}(\omega) = u(\mu_F(\omega), \mu_G(\omega)) \quad (23)$$

$$\mu_{F \cap G}(\omega) = i(\mu_F(\omega), \mu_G(\omega)) \quad (24)$$

On va chercher à conserver le plus possible de propriétés des opérations usuelles d'union et d'intersection d'ensemble, c'est-à-dire rester aussi proche que possible de la structure de treillis Booléen de 2^Ω . Le résultat suivant (Dubois et Prade /10/) montre que $\tilde{\mathcal{F}}(\Omega) = [0, 1]^\Omega$ ne pourra jamais être un treillis Booléen quelque soit le choix de u et i .

Proposition 2.1 : Si la loi du tiers exclu $F \cup \bar{F} = \Omega$ (resp : de non-contradiction $F \cap \bar{F} = \emptyset$) est valide, alors l'union d'ensembles flous (resp : leur intersection) n'est plus idempotente i.e. $F \cup F \neq F$ (resp : $F \cap F \neq F$).

Dans la proposition ci-dessus les lois du tiers-exclu et de la non-contradiction s'écrivent au sens de n'importe quelle complémentation C satisfaisant C1-C3, sous la forme, $\forall x \in [0, 1], u(x, C(x)) = 1 ; i(x, C(x)) = 0$.

Par ailleurs on sait que l'idempotence de l'union et de l'intersection n'est qu'une conséquence de leur distributivité mutuelle.

En conséquence on doit choisir entre les deux options suivantes : définir i et u de façon à garder les lois du tiers exclu et de non contradiction (alors plus d'idempotence ni de distributivité mutuelle pour \cap et \cup), ou conserver l'idempotence et abandonner les lois du tiers-exclu et de non-contradiction.

Cette discussion préliminaire justifie d'imposer à i et u les axiomes suivants : ($x, y \in [0,1]$)

- Coïncidence avec l'intersection et l'union classique

$$U0) \quad u(0,1) = u(1,1) = u(1,0) = 1 ; u(0,0) = 0$$

$$I0) \quad i(0,1) = i(0,0) = i(1,0) = 0 ; i(1,1) = 1$$

- Commutativité

$$U1) \quad u(x,y) = u(y,x) ;$$

$$I1) \quad i(x,y) = i(y,x)$$

- Associativité

$$U2) \quad u(x,u(y,z)) = u(u(x,y),z)$$

$$I2) \quad i(x,i(y,z)) = i(i(x,y),z)$$

- Lois de De Morgan : il existe une complémentation C satisfaisant $C1-C3$.
telle que :

$$U3) \quad C(u(x,y)) = i(C(x),C(y))$$

$$I3) \quad C(i(x,y)) = u(C(x),C(y))$$

- Élément neutre

$$U3) \quad u(x,0) = x \quad (F \cup \emptyset = F)$$

$$I3) \quad i(x,1) = x \quad (F \cap \Omega = F)$$

- Monotonie

$U5-I5$ u et i sont non décroissants sur chaque argument

- Continuité

$U6-I6$ u et i sont continues.

$U1-U4$ et $I3-I4$ sont valides dans la théorie classique $U5$ et $I5$ sont naturelles parce que si ω appartient moins à F et à G que ω' alors ω' ne peut appartenir plus à $F \cup G$ ni à $F \cap G$ que ω . La continuité est une condition de nature technique quand Ω est infini.

Il est clair, au vu de l'annexe, que l'on ne peut choisir l'intersection et l'union d'ensembles flous que parmi les normes et co-normes triangulaires (continues) respectivement, dès qu'on adopte les axiomes ci-dessus. Ce fait a été mentionné simultanément par plusieurs chercheurs tels que Alsina et al /1/, Dubois /9/, Klement /39/, Prade /60/. Les principales classes d'intersection et d'union peuvent alors être caractérisées comme suit :

→ les opérateurs idempotents : $i(x,y) = \min(x,y)$ et $u(x,y) = \max(x,y)$.

Ils sont mutuellement distributifs; \min et \max sont les seuls opérateurs d'union et d'intersection vérifiant IO-I5 et UO-U5 qui soient idempotents (et mutuellement distributifs) résultat déjà prouvé il y a longtemps (Bellman et Giertz, /2/) \max et \min s'interchangent par dualité de Morgan à l'aide de n'importe quelle complémentation. Avec ces opérateurs, $\tilde{\mathcal{F}}(\Omega)$, l'ensemble des parties floues de Ω , est muni d'une structure de treillis distributif complet qui reflète celle de $([0,1], \max, \min)$. La complémentation n'est qu'une pseudo-complémentation puisqu'on a seulement

$$\forall \omega, \mu_{A \cap \bar{A}}(\omega) \leq 0,5 \quad ; \quad \mu_{A \cup \bar{A}}(\omega) \geq 0,5 \quad (25)$$

→ Les opérateurs archimédiens strictement monotones, c'est-à-dire tels que

$$\forall x \in]0,1[\quad i(x,x) < x, \quad u(x,x) > x$$

$$\forall y' > y \quad i(x,y') > i(x,y) \quad ; \quad u(x,y') > u(x,y)$$

ont été étudiés par Dombi /6/, Dubois /8/ et Singer /68/. Leur archétype est l'opérateur produit pour l'intersection et la somme probabiliste $(x+y-xy)$ pour l'union. Leur forme générale est (cf. Annexe)

$$i(x,y) = f^{-1}(f(x)+f(y))$$

où f est une bijection continue décroissante de $[0,1]$ dans $[0,+\infty)$, telle que $f(0) = +\infty$, $f(1) = 0$.

$$u(x,y) = \varphi^{-1}(\varphi(x)+\varphi(y))$$

où φ est une bijection continue croissante de $[0,1]$ dans $[0,+\infty)$ telle que $\varphi(0) = 0$, $\varphi(1) = +\infty$.

On les appellera intersection et unions strictes, du nom des normes et co-normes triangulaires qui les modélisent. $F = \text{Log}$ et $\varphi = \text{Log}(1-.)$ génèrent le produit et la somme probabiliste respectivement. On a les propriétés suivantes des opérations ensemblistes strictes :

- Elles ne sont jamais distributives l'une sur l'autre ni idempotentes et elles ne vérifient jamais les lois du tiers-exclu et de la non-contradiction
- Etant donné (f, φ) générant (i, u) , les lois de De Morgan ne sont pas forcément vérifiées par une complémentation involutive. En effet, si (i, u) vérifie I3-U3 pour C on a forcément $f = \varphi \circ C$ à une constante non-nulle multiplicative près. Puisque C est involutif, $\varphi^{-1} \circ f$ doit l'être aussi. Or en général, $\varphi^{-1} \circ f$ ne satisfait que C1 et C2.
- Etant donné i généré par f , il existe une union unique qui satisfait les lois de De Morgan au sens d'une complémentation C ; u est générée par $\varphi = f \circ C$. Si s est le seuil d'appartenance associé à $C((C(s)=s))$ alors, au sens de (i, u, C) avec $\varphi = f \circ C$ on a toujours (Dubois, /8/)

$$\forall \omega \quad \mu_F \wedge \bar{F}(\omega) \leq i(s, s) \quad (26)$$

$$\mu_F \cup \bar{F}(\omega) \geq u(s, s) = C(i(s, s)) \quad (27)$$

- $\forall n \in \mathbb{N}, (x_1 \dots x_n) \in]0, 1[^n$, alors $i(x_1 \dots x_n) > 0$ et $u(x_1 \dots x_n) < 1$. Si on note \bigcap^n et \bigcup^n l'intersection et l'union de F , n fois avec lui-même, on a

$$\lim_{n \rightarrow +\infty} \bigcap^n F = \overset{\circ}{F} \in 2^\Omega \quad \text{mais} \quad \forall n, S(\bigcap^n F) \supset \overset{\circ}{F}$$

$$\lim_{n \rightarrow +\infty} \bigcup^n F = S(F) \in 2^\Omega, \quad \text{mais} \quad \forall n, (\bigcup^n F) \subset S(F)$$

pour tout F non classique.

Des familles paramétrées d'opérateurs stricts ont été proposés dans la littérature. Hamacher /27/ a étudié les seules intersections, strictes, qui soit des fonctions rationnelles des arguments:

$$i(x, y) = \frac{ab}{\gamma + (1-\gamma)(a+b-ab)} \quad \gamma \geq 0$$

Frank /19/ a étudié les seuls opérateurs tels que

$$u(x,y) + i(x,y) = x + y :$$

$$i(x,y) = \text{Log}_s \left[1 + \frac{(s^x - 1)(s^y - 1)}{s - 1} \right] \quad s > 0$$

Pour d'autres familles paramétrées voir Dombi /6/.

→ les opérateurs nilpotents décrits par des normes et co-normes triangulaires nilpotentes (voir annexe) et dont les archétypes sont

$$\text{intersection} : i(x,y) = \max(0, x+y-1)$$

$$\text{union} : u(x,y) = \min(1, a+b). \quad (\text{somme bornée})$$

Ils ont été étudiés par Dubois /8/. Leur forme générale est

$$i(x,y) = f^*(f(x)+f(y))$$

où f est une bijection décroissante de $[0,1]$ dans $[0, f(0)]$ telle que

$$f(0) < +\infty, \quad f(1) = 0.$$

$$u(x,y) = \varphi^*(\varphi(x) + \varphi(y))$$

où φ est un générateur de complémentation. φ^* est la pseudo-inverse de φ (cf. Annexe).

On vérifie les propriétés suivantes des opérateurs nilpotents :

- . absence de mutuelle distributivité et d'idempotence.
- . tout opérateur de complémentation C génère une intersection et une union qui sont duales au sens de de Morgan, et satisfont les lois du tiers exclu et de la non-contradiction, pour cette complémentation.
- u et C sont générés par φ , i est généré par $f = \varphi(1) - \varphi$.
- . Etant donné un opérateur nilpotent d'union généré par φ et une négation quelconque C , $\varphi \circ C$ génère l'intersection duale de u , pour C , au sens de De Morgan. Si on veut de plus que les lois du tiers exclu et de non-contradiction soit satisfaites alors il faut que φ génère C (et donc $f = \varphi(1) - \varphi$).
- . $\forall F \in [0,1]^\Omega, \exists n \in \mathbb{N} . \bigwedge^n F = \overset{\circ}{F} ; \bigcup^n F = S(F)$. C'est la propriété de nilpotence.

Des familles paramétrées d'opérateurs nilpotents ont été proposés dans la littérature : l'intersection T_p de Schweizer et Sklar (voir annexe)

la famille de Yager /78/, respectivement générées par $f_p(x) = |1-x^{-p}|$ $p \in \mathbb{R}$ et, $f_q(x) = (1-x)^q$ $q > 0$. On notera que l'opérateur utilisé par Sugeno (cf chapitre 1) pour définir les mesures floues g_λ est la co-norme (l'union) nilpotente $\min(1, a+b+\lambda ab)$ $\lambda \geq -1$.

On peut trouver des opérateurs d'intersection et d'union qui ne soient ni archimédiens ni nilpotents telle la famille d'intersections proposées par Dubois et Prade /10/ :

$$i(x, y) = \frac{x \cdot y}{\max(x, y, \alpha)} \quad \alpha \in [0, 1] \quad (28)$$

qui est une somme ordinale du produit (voir annexe).

→ l'intersection et l'union fortes définies respectivement par les norme T_W et co-norme T_W^* (voir annexe). Ce sont les plus petite norme et plus grande co-norme triangulaires respectivement⁽¹⁾. Avec de tels opérateurs, les lois du tiers exclu et de non-contradiction sont toujours vérifiées au sens de n'im-
porte quelle complémentation vérifiant C1-C2, puisque

$$\forall (x, y) \in]0, 1]^2 \quad T_W^*(x, y) = 1,$$

et

$$\forall (x, y) \in [0, 1[^2 \quad T_W(x, y) = 0$$

REMARQUES : La propriété de distributivité peut être partiellement préservée en choisissant $i = \min$ (resp $u = \max$) et u (resp : i) quelconques puisque

$$u(x, \min(y, z)) = \min(u(x, y), u(y, z)) \quad (29)$$

$$i(x, \max(y, z)) = \max(i(x, y), i(x, z)) \quad (30)$$

sinon on ne garde éventuellement que des inégalités (cf. Prade /60/ Mizumoto /51/, /52/). On notera que $([0, 1], \min, u)$ et $([0, 1], \max, i)$ forment chacun une structure de dioïde pour toute co-norme u et toute t -norme i . (Voir Gondran et Minoux : Graphes et Algorithmes, Eyrolles, 1979, et le chapitre 7)

$$\cdot (A \cap B)_\alpha = A_\alpha \cap B_\alpha ; (A \cup B)_\alpha = A_\alpha \cup B_\alpha \text{ si et seulement si } \cap = \min \cup = \max$$

. Klement /45/ a étudié des opérateurs non-ponctuels dans la même optique que Lowen /49/. Voir Smets /72/ dans le cas de référentiels composites.

(1) Elles ne sont pas continues ; néanmoins cette qualité de bornes inférieures et supérieures fait qu'il faut les considérer.

2.3.3 - Opérateurs d'union et d'intersection : approches empiriques

En plus de l'approche axiomatique, il convenait de vérifier le bien fondé des opérateurs d'union et d'intersection pour traduire les conjonctions "et" et "ou" du langage naturel lorsqu'elles relient des catégories vagues. Des résultats d'expérimentations apparaissent par exemple, dans Hersh et Caramazza /28-29/ Zimmermann /88/, Thöle et al /73/, Zimmermann and Zysno /89/, Oden /57/.

On parvient à tirer les enseignements suivants : on peut effectivement, dans des conditions données, au sens d'une personne ou d'un groupe de personnes donnés, identifier un opérateur qui agrège des catégories simples. L'opérateur minimum semble être valide pour le "et" quand les deux catégories renvoient au même concept ("un homme grand, mais pas trop grand") ce que Zimmermann /88/ appelle le "et logique". Quand ce n'est pas le cas ("une voiture bon marché et rapide") alors non seulement d'autres opérateurs que le minimum peuvent être identifiés, mais encore on peut sortir du domaine des normes triangulaires, lorsque par exemple il y a compensation entre les deux concepts. On peut par exemple en conclure que la moyenne arithmétique ou la moyenne géométrique expriment le "et" de façon plus satisfaisante.

2.3.4 - Opérateurs ensemblistes spécifiques des ensembles flous

En théorie des ensembles classiques, il n'est que deux façons de combiner des ensembles de façon symétrique et telle que

$$\omega \in A, \omega \in B \rightarrow \omega \in A * B$$

$$\omega \in \bar{A}, \omega \in \bar{B} \rightarrow \omega \notin A * B$$

* doit être soit l'intersection, soit l'union. Ce n'est plus vrai dans le cadre des ensembles flous. On a vu que les opérateurs flous d'intersection (resp : union) étaient plus petits (resp : grands) que la fonction minimum (resp : maximum). On ne couvre donc qu'une partie des opérateurs d'agrégation possible. Dans cette section on étudie les opérateurs de moyenne (entre max et min) et les opérateurs auto-duaux au sens de De Morgan. Ces deux types d'opérations ensemblistes floues n'ont pas de contrepartie classique, mais sont parfois bien connus malgré tout. Ce qui est original ici c'est de les placer dans un contexte ensembliste.

2.3.4.1 - OPERATEURS DE MOYENNE

La moyenne de deux ensembles flous se définit à l'aide d'une application m de $[0,1]^2$ dans $[0,1]$ telle que

$$M1 \quad \min(x,y) \leq m(x,y) \leq \max(x,y) \quad \forall x,y \text{ mais } m \notin \{\min, \max\}$$

$$M2 \quad m(x,y) = m(y,x)$$

$$M3 \quad m \text{ est non décroissant en fonction de chaque argument}$$

On montre les propriétés remarquables suivantes

i) (Dubois et Prade /10/) si m est strictement croissant alors il ne peut pas être associatif

ii) (Dubois et Prade /13/) les seules moyennes associatives sont les médianes définies pour un seuil $\alpha \in]0,1[$ par

$$\begin{aligned} \text{med}(x,y,\alpha) &= y \quad \text{si } x \leq y \leq \alpha \\ &= \alpha \quad \text{si } x \leq \alpha \leq y \\ &= x \quad \text{si } \alpha \leq x \leq y \end{aligned}$$

De plus il est facile de voir que m est idempotent. Réciproquement, si m est idempotent alors $M3 \rightarrow M1$.

Donc l'associativité est une propriété incompatible avec la notion de moyenne. On la remplace par la bisymétrie qui s'écrit

$$M4 \quad \forall x,y,z,t \quad m[m(x,y), m(z,t)] = m[m(x,z), m(y,t)]$$

On notera que l'associativité implique la bisymétrie. Les équations fonctionnelles $M1-M4$, avec hypothèse de continuité et de croissance stricte ont déjà été résolues par le passé. La forme générale de la solution est :(cf. annexe).

$$m(x,y) = f^{-1} \left(\frac{f(x)+f(y)}{2} \right) \quad (31)$$

où f est continue et strictement monotone.

Par exemple $f(x) = x^\alpha$, $\alpha \in \mathbb{R}$ donne une famille m_α dont les cas particuliers sont

α	$m(x,y)$
$-\infty$	$\min(x,y)$
-1	$\frac{2xy}{x+y}$ (moyenne harmonique)
0	\sqrt{xy} (moyenne géométrique)
$+1$	$\frac{x+y}{2}$ (moyenne arithmétique)
$+\infty$	$\max(x,y)$

et qui est telle que $\alpha_1 < \alpha_2 \rightarrow m_{\alpha_1} < m_{\alpha_2}$

On notera que la médiane $\text{med}(x,y,\alpha)$ est bisymétrique sans être de la forme (31).

2.3.4.2 - OPERATEURS DE SOMME SYMETRIQUE

Silvert /68/ a suggéré l'étude d'une classe d'agrégation d'ensembles flous très intéressante sur le plan sémantique : les opérateurs auto-duaux au sens de De Morgan, c'est-à-dire des fonctions $\sigma : [0,1]^2 \rightarrow [0,1]$ qui satisfont

$$S1) \sigma(0,0) = 0 \quad \sigma(1,1) = 1$$

S2) σ est commutatif

S3) σ est non décroissant en fonction de chaque argument

S4) σ est continue

$$S5) 1 - \sigma(x,y) = \sigma(1-x, 1-y)$$

L'axiome S5 d'invariance par transformation de De Morgan peut être généralisé à l'aide d'une complémentation C en

$$S5') \exists C : C(\sigma(x,y)) = \sigma(C(x), C(y))$$

Toute somme symétrique au sens de S5, et vérifiant S1-S4 peut s'exprimer sous la forme (Silvert, /68/).

$$\sigma(x,y) = \frac{g(x,y)}{g(x,y) + g(1-x, 1-y)} \quad (32)$$

où g est une fonction quelconque, non décroissante, non négative, continue, telle que $g(0,0) = 0$

On peut montrer que (Silvert, /68/)

$$\rightarrow \sigma(x, 1-x) = \frac{1}{2} \quad \text{pour tout } x \in]0, 1[$$

$$\rightarrow \sigma(0, 1) \text{ n'est pas définie dès que } g(0, x) = 0 \quad \forall x; \text{ sinon } \sigma(0, 1) = \frac{1}{2}$$

\rightarrow La seule somme symétrique associative qui soit une moyenne au sens de M1 est $\text{med}(x, y, \frac{1}{2})$.

On peut choisir pour g une norme triangulaire ou une co-norme triangulaire par exemple. Avec $g(x, y) = x+y$ on retrouve la moyenne arithmétique. Avec $g = \min$ ou \max on trouve respectivement

$$\sigma_{\min}(x, y) = \frac{\min(x, y)}{1 - |x-y|} ; \quad \sigma_{\max}(x, y) = \frac{\max(x, y)}{1 + |x-y|} \quad (33)$$

Ces trois opérateurs sont des opérateurs de moyenne, mais σ_{\min} et σ_{\max} ne sont pas bisymétriques.

On peut généraliser le résultat de Silvert /67/ avec l'axiome S5' et montrer que la forme générale des sommes symétriques autoduales pour la négation C , de générateur φ (= somme C -symétrique) est :

$$\sigma(x, y) = \varphi^* \left(\frac{g(x, y)}{g(x, y) + g(C(x), C(y))} \right) \quad (34)$$

on a alors s unique, $s \in]0, 1[$, tel que

$$\sigma(x, C(x)) = s = C(s) \quad (35)$$

$\text{med}(x, y, s)$ est aussi la seule somme C -symétrique associative qui soit une moyenne.

Donc, pour les sommes symétriques on doit choisir une fois de plus entre l'associativité et l'idempotence.

Dombi /7/ utilisant les résultats classiques sur l'équation fonctionnelle d'associativité, a signalé que les sommes C -symétriques associatives strictement croissantes se mettaient sous la forme

$$\sigma(x, y) = f^{-1}(f(x) + f(y)) \quad (36)$$

où f est strictement monotone, $f(0)$ et $f(1)$ ne sont pas bornés et $f(s) = 0$, s étant donné par (35).

On peut alors montrer que ces sommes C-symétriques associatives sont telles que

$$\forall x, \sigma(x,s) = x \quad ; \quad \forall x \neq 0 \quad \sigma(x,1) = 1 \quad (37)$$

$$\forall x < s < y, \quad \sigma(x,x) < x \quad ; \quad \sigma(y,y) > y \quad ; \quad x < \sigma(x,y) < y \quad (38)$$

$$\forall x \neq 1 \quad \sigma(x,0) = 0 \quad (\text{d'où } g(x,0) = 0) \quad (39)$$

$$\text{Si } \max(x,y) < s, \quad \text{alors } \sigma(x,y) \leq \min(x,y) \quad (40)$$

$$\text{Si } \min(x,y) > s, \quad \text{alors } \sigma(x,y) \geq \max(x,y) \quad (41)$$

En particulier les sommes C-symétriques générées selon (34) avec une fonction g telle que $g(0,x) > 0$ pour quelque $x < 1$ (par exemple g est une co-norme) ne sont pas associatives. L'associativité doit être restreinte au domaine $]0,1[$, puisque pour les sommes symétriques associatives $\sigma(0,1)$ n'est pas défini.

Un exemple de somme symétrique associative est générée par $g(x,y) = x.y$.

$$\sigma_{\bullet}(x,y) = \frac{xy}{1-x-y+2xy} \quad (42)$$

En revanche $\sigma_{\dagger}(x,y)$, générée par la co-norme duale

$$\sigma_{\dagger}(x,y) = \frac{x+y-xy}{1-x-y+2xy} \quad (43)$$

n'est pas associative. $g(0,x) = 0$ n'est pas une condition suffisante d'associativité.

Considérons maintenant les sommes symétriques idempotentes. Etant non décroissantes, ce sont des opérateurs de moyenne au sens de M1-M3. Si on ajoute la propriété de bisymétrie (M4) et la croissance stricte alors on peut vérifier le résultat suivant.

Il n'y a qu'une seule somme C-symétrique idempotente, strictement croissante, et bisymétrique ; elle est de la forme :

$$\sigma(x,y) = \varphi^{-1}\left(\frac{\varphi(x)+\varphi(y)}{2}\right) \quad (44)$$

où φ génère $C^{(1)}$

En particulier la seule somme symétrique idempotente croissante et bisymétrique, de Silvert /68/ est la moyenne arithmétique. Dombi /7/ ; ne résoud pas le problème de la détermination des sommes symétriques strictement croissantes associatives, qui satisfont $S5'$ au sens d'une complémentation C . En fait, tout opérateur associatif de la forme (36) est C -symétrique pour et seulement pour la complémentation définie par⁽²⁾:

$$C(x) = f^{-1}(-f(x)) \quad (45)$$

En particulier les sommes symétriques associatives de Silvert ($C(x)=1-x$) sont générées par des fonctions telles que $f(1-x) = -f(x)$; par exemple la somme symétrique $\sigma(x,y)$ est générée par $f(x) = \text{Log}(\frac{1-x}{x})$. Etant donnée n'importe quelle fonction h monotone sur $[0, +\infty)$ et telle que $h(0)$ n'est pas borné,

1. Preuve on a $\sigma(x,y) = f^{-1}(\frac{f(x)+f(y)}{2})$ pour f strictement monotone.
et

$$\text{Cof}^{-1}(\frac{f(x)+f(y)}{2}) = f^{-1}(\frac{\text{foC}(x)+\text{foC}(y)}{2}) ;$$

Quand $x \stackrel{\Delta}{=} C(y)$ on a $\text{Cof}^{-1}(z) = f^{-1}(z)$ pour $z \stackrel{\Delta}{=} \frac{\text{foC}(x)+f(x)}{2}$, $\forall x$.

Si z dépend de x on en déduit que $f^{-1} = \text{Cof}^{-1}$ en au moins 2 points ce qui est impossible car f^{-1} et Cof^{-1} sont strictement monotones en sens inverse.

Donc

$$\frac{\text{foC}(x)+f(x)}{2} = f(s) = \frac{f(0)+f(1)}{2}$$

Donc

$$C(x) = f^{-1}(f(0)+f(1)-f(x))$$

Donc $\exists a, b$ tel que $f(x) = a\varphi + b$ ($b=f(0)$ $a=f(1)-f(0)$) et φ et f génèrent le même σ .

2. Preuve : $\sigma(x,y) = f^{-1}(f(x)+f(y)) = \text{Cof}^{-1}(\text{foC}(x)+\text{foC}(y))$

donc f et foC génèrent σ , et $\text{foC} = \lambda f$; $\lambda < 0$ puisque C décroît; donc

$C(x) = f^{-1}(\lambda f(x))$. On vérifie aisément que $C(1)=0$, $C(0) \leq 1$ et $C(s)=s$ pour s t.q. $f(s)=0$. La condition d'involution implique $\lambda^2=1$ donc $\lambda=-1$.

$f(x) = h(x) - h(1-x)$ est un générateur de sommes symétriques associatives autoduales pour la complémentation habituelle. Et seules les fonctions f de ce type génèrent les sommes symétriques associatives.

Enfin, notons les résultats suivants (Dubois et Prade /10/) qui précisent le domaine parcouru par les sommes symétriques de Silvert

. Si g génère σ par (32) et g est une application $[0,1]^2 \rightarrow [0,1]$ alors

$$\min(g(x,y), 1-g(1-x,1-y)) \leq \sigma(x,y) \leq \max(g(x,y), 1-g(1-x,1-y)) \quad (46)$$

Particulièrement si σ_i est généré par la norme triangulaire i et σ_u par la co-norme duale ($C(x)=1-x$) alors

$$i(x,y) \leq \sigma_i(x,y) \leq u(x,y) \quad (47)$$

$$i(x,y) \leq \sigma_u(x,y) \leq u(x,y) \quad (48)$$

C'est-à-dire que σ_i et σ_u sont des "moyennes" au sens de l'axiome M1 où on changerait "min" et "max" en i et u respectivement

Si i et u sont tels que $i(x,y) + u(x,y) = x+y$ (c'est la famille de Frank - voir annexe) - par exemple $i = \min$, produit, $\max(0, x+y-1)$... alors

$$x + y \leq 1 \rightarrow \sigma_i(x,y) \leq \frac{x+y}{2} \leq \sigma_u(x,y) \quad (49)$$

$$x + y \geq 1 \rightarrow \sigma_u(x,y) \leq \frac{x+y}{2} \leq \sigma_i(x,y) \quad (50)$$

$i \leq i'$, $x + y \leq 1 \rightarrow$

$$\sigma_i(x,y) \leq \sigma_{i'}(x,y) \leq \frac{x+y}{2} \leq \sigma_u(x,y) \leq \sigma_u(x,y) \quad (51)$$

Si $x + y \geq 1$

$$\sigma_i(x,y) \geq \sigma_{i'}(x,y) \geq \frac{x+y}{2} \geq \sigma_u(x,y) \geq \sigma_u(x,y) \quad (52)$$

Par exemple si $x + y \geq 1$

$$\frac{xy}{1-x-y+2xy} \leq \frac{\min(x,y)}{1-|x-y|} \leq \frac{x+y}{2} \leq \frac{\max(x,y)}{1+|x-y|} \leq \frac{x+y-xy}{1+x+y-2xy} \quad (53)$$

Enfin si un opérateur de moyenne (vérifiant M1-M3) est tel que

$$\min(x,y) \leq m(x,y) \leq \frac{x+y}{2} \quad \forall x, \forall y, \quad (m \neq \text{moyenne arithmétique})$$

Alors

$$\max(x,y) \geq 1 - m(1-x, 1-y) \geq \frac{x+y}{2} \quad (54)$$

Donc m n'est jamais une somme symétrique de Silvert. En revanche le dual de m est encore une moyenne. On peut donc considérer le dual de la moyenne géométrique, de la moyenne harmonique... Ces opérateurs peuvent générer des sommes symétriques par (32) qui seront encore des moyennes en vertu de (46).

2.3.5 - Opérateurs ensemblistes de différence et d'inclusion

Dans cette section, on considère l'extension floue des opérateurs ensemblistes de différence, d'inclusion et de différence symétrique. Ces opérateurs peuvent être construits, à partir de la connaissance de la complémentation et de l'union ou de l'intersection, par composition. Mais on peut aussi choisir pour point de départ une définition de l'inclusion et, connaissant l'union (resp : l'intersection) construire l'opérateur de différence et son dual.

2.3.5.1 - DIFFERENCE ENSEMBLISTE ET INCLUSION : 1ère construction

En théorie des ensembles classiques on définit :

$$B - A = B \cap \bar{A} \quad \forall A, B \in 2^\Omega \quad (55)$$

L'opérateur dual $\overline{B-A} = \bar{B} \cup A$ est directement lié à l'inclusion car

$$\bar{B} \cup A = \Omega \quad \leftrightarrow \quad B \subseteq A \quad (56)$$

Dans le cas d'ensembles flous, connaissant une complémentation C et une norme triangulaire i , on peut poser

$$\mu_{B-A}(\omega) = i(\mu_B(\omega), C(\mu_A(\omega))) \quad (57)$$

et $B \subseteq A$ si et seulement si

$$\mu_{\bar{B} \cup A}(\omega) = u(C(\mu_B(\omega)), \mu_A(\omega)) = 1 \quad \forall \omega \in \Omega \quad (58)$$

Si u et i sont duaux au sens de C on a $\bar{B} \cup A = \overline{B-A}$ comme dans le cas classique. On peut particulariser les opérateurs comme suit

C	i	$B-A$	$\bar{B} \cup A$
$1-x$	min	$\min(y, 1-x)$	$\max(1-y, x)$ (Dienes)
$1-x$	produit	$y(1-x)$	$1-y+xy$ (Reichenbach)
$1-x$	$\max(0, x+y-1)$	$\max(0, y-x)$	$\min(1, 1-y+x)$ (Lukasiewicz)
$1-x$	T_W	$T_W(y, 1-x)$	$T_W^*(1-y, x)$
$1-x$	norme stricte	$f^{-1}(f(y)+f(1-x))$	$1-f^{-1}(f(y)+f(1-x))$
$C(\varphi)$	$i(\varphi)$	$\max(0, \varphi^*(\varphi(y)-\varphi(x)))$	$\min(1, \varphi^*(\varphi(1)-\varphi(y)+\varphi(x)))$

(u est toujours C -dual de i , ci-dessus)

Dans le tableau ci-dessus on reconnaît plusieurs implications connues dans les logiques multivalentes (cf Rescher /65/).

La définition (58) de l'inclusion appelle les remarques suivantes :

. Si u est une co-norme nilpotente de même générateur φ que C alors

$$(58) \leftrightarrow \forall \omega, \mu_B(\omega) \leq \mu_A(\omega) \quad (59)$$

qui est la définition posée à l'origine par Zadeh /81/

. Si u est le maximum ou une co-norme stricte alors $\forall C$,

$$(58) \leftrightarrow (\mu_B(\omega) = 0 \text{ ou } \mu_A(\omega) = 1) \forall \omega$$

$$\leftrightarrow S(B) \subseteq \bar{A}$$

c'est-à-dire que le support de B est contenu dans le noyau de A , ce qui est une définition très stricte de l'inclusion. Ce résultat est encore valide pour tout u tel que $\max \leq u \leq \bar{u}$ où \bar{u} est une co-norme stricte.

On peut vouloir relaxer cette définition en choisissant un seuil α tel que

$$B \underset{\alpha}{\subseteq} A \leftrightarrow \mu_{\bar{B} \cup A} \geq \alpha \quad (60)$$

Compte tenu de l'absence de la loi du tiers exclu, et que C et u étant fixé on a toujours $\mu_{A \cup \bar{A}} \geq u(s,s)$ où $s = C(s)$ on est amené à choisir $\alpha \leq u(s,s)$. Néanmoins on perd alors la transitivité de l'inclusion généralement (par exemple pour $u =$ somme probabiliste, $C(x) = 1-x$, elle est perdue $\forall \alpha < 1$). Quand $u = \max$, $C(x) = 1-x$, on a montré (Dubois et Prade /11/ p.22) qu'en prenant $\alpha = \frac{1}{2}$ et en modifiant (58) en

$$B \prec A \leftrightarrow \forall \omega \in \Omega, \mu_B(\omega) \leq \frac{1}{2} \quad \text{or} \quad \mu_A(\omega) > \frac{1}{2} \quad (61)$$

on obtenait une inclusion faible transitive et cohérente avec (59) au sens ou

$$\mu_B \leq \mu_A \rightarrow B \prec A \quad (62)$$

2.3.5.2 - DIFFERENCES ENSEMBLISTES INDUITES PAR L'INCLUSION DE ZADEH

On admet maintenant (59) comme définition de l'inclusion (c.a.d. $\mu_A \leq \mu_B$). Dans la théorie classique l'opérateur B-A peut être défini à partir de l'union par

$$B - A = \bigcap \{X | A \cup X \supseteq B\} (= B \cap \bar{A}) \quad (63)$$

c'est-à-dire le plus petit ensemble dont l'union avec A recouvre B. De façon analogue

$$\overline{B - A} = \bigcup \{X | B \cap X \subseteq A\} (= \bar{B} \cup A) \quad (64)$$

Quand A et B sont flous, connaissant l'union et l'intersection sous forme de co-normes et de normes triangulaires, et l'inclusion de Zadeh on peut définir la différence d'ensemble flou B-A et son dual, noté $B \blacktriangleright A$ pour la circonstance, par

$$\forall \omega, \mu_{B-A}(\omega) = \inf\{x | u(x, \mu_A(\omega)) \geq \mu_B(\omega)\} \\ \triangleq \mu_A(\omega) \leftarrow \mu_B(\omega) \quad (65)$$

$$\mu_{B \blacktriangleright A}(\omega) = \sup\{x | i(x, \mu_B(\omega)) \leq \mu_A(\omega)\} \\ \triangleq \mu_B(\omega) \circ \mu_A(\omega) \quad (66)$$

(65) définit une application de $[0,1]^2$ dans $[0,1]$ notée \leftarrow et (66) une appli-

i	$x \sqsubset y$
min	$\begin{matrix} 1 & \text{si } x < y \\ y & \text{si } x > y \end{matrix}$ (Brouwer-Gödel) Sanchez 1976
produit	$\min(1, \frac{y}{x})$ (Goguen, 1971)
$\max(0, x+y-1)$	$\min(1, 1-x+y)$ (Lukasiewicz)
T_W	$\begin{matrix} 1 & \text{si } x < 1 \\ y & \text{si } x = 1 \end{matrix}$
norme stricte (f)	$\min(1, f^{-1}(f(y)-f(x)))$ (avec $f(x) < 0$ pour $x > 1$)
norme nilpotente ($\varphi(1)-\varphi$)	$\min(1, \varphi^*(\varphi(1)-\varphi(x)+\varphi(y)))$

u	$x \leftarrow y$
max	$\begin{matrix} 0 & \text{si } x \geq y \\ y & \text{si } x < y \end{matrix}$ (Sanchez 1976)
$x + y - xy$	$\max(0, \frac{y-x}{1-x})$
$\min(1, x+y)$	$\max(0, y-x)$ (Zadeh)
T_W^*	$\begin{matrix} 0 & \text{si } x > 0 \\ y & \text{si } x = 0 \end{matrix}$
co-norme stricte ou nilpotente (φ)	$\max(0, \varphi^*(\varphi(y)-\varphi(x)))$

(cf. aussi Prade /62/ et Valverde /75/ pour le tableau des opérateurs \sqsubset , Pedrycz /60/ pour \leftarrow).

cation notée $\overline{[x \sqcap y]}$ (resp $x \leftarrow y$) est le pseudo-complément de x par rapport à y dans un treillis de Brouwer (resp : Brouwerien dual) ; voir Sanchez /67/). Le tableau ci-contre donne pour chaque norme triangulaire (resp : co-norme triangulaire) les opérations \leftarrow et $\overline{[}$ générées. On reconnaît dans $\overline{[}$ d'autres implications multivalentes, sauf les implications type Lukaziewicz (et sa duale) déjà trouvées avec l'autre approche.

On peut chercher à retrouver une complémentation C telle que

$$\overline{B - A} = B \rightarrow A \quad (67)$$

Conformément au cas classique. On peut montrer facilement les résultats suivants⁽¹⁾:

si $B-A$ et $B \rightarrow A$ sont définis respectivement à l'aide de :

→ max et min : $\forall C, \overline{B - A} \neq B \rightarrow A$

→ une co-norme et une t-norme strictes C -duales : $\overline{B - A} \neq B \rightarrow A, \forall C$

→ une co-norme et une t-norme nilpotentes générées par φ

$$\overline{B - A} = B \rightarrow A \quad \text{au sens de } C \text{ générée par } \varphi$$

Les deux méthodes de construction sont équivalentes lorsqu'on part d'opérateurs nilpotents, ce qui est notable.

NB : connaissant $B-A$ et $B \rightarrow A$ on peut formellement reconstruire l'union et l'intersection en posant

$$\text{(intersection)} \quad B \sqcap A \stackrel{\Delta}{=} B - \bar{A} \quad (68)$$

$$\text{(union)} \quad B \sqcup A \stackrel{\Delta}{=} \bar{B} \rightarrow A \quad (69)$$

On définit alors une transformation T qui transforme l'union (resp : l'intersection) définie comme en 2.3.3 et intersection (resp : union) définie par (68) et (69), via (65) et (66).

(1) Preuve : max et min : évident. Pour des opérateurs stricts C -duaux :

(62) est équivalent à :

$$\exists C, C(\max(0, \varphi^{-1}(\varphi(y) - \varphi(x)))) = \min(1, C \circ \varphi^{-1}(\varphi \circ C(y) - \varphi \circ C(x))).$$

soit : $\forall y, \varphi(y) + \varphi \circ C(y) = \text{constante} = 2\varphi(s)$ par exemple.

tel que $s = C(s)$. Donc $C = \varphi^{-1}(2\varphi(s) - \varphi)$. Mais puisque $\varphi(1) = +\infty$

C n'est pas une complémentation ($C(0) \neq 1$)

Pour des opérateurs nilpotents le résultat est évident.

Si \leftarrow_i est g n r e par l'intersection i , on construit l'union $u_i \stackrel{\Delta}{=} T(i)$ telle que

$$u_i(x,y) = C(x) \leftarrow_i y$$

De m me

$$i_u = T(u) \quad \text{telle que} \quad i_u(x,y) = C(x) \overleftarrow_i y$$

on a par exemple des op rateurs d'"union" d finis par ($C(x) = 1-x$)

$$T(\text{min}) : u_{\text{min}}(x,y) = 1 \quad \text{si } x + y \geq 1 \\ = y \quad \text{sinon}$$

$$T(\text{produit}) : u_{\cdot}(x,y) = \min(1, \frac{y}{1-x})$$

et des op rateurs d'"intersection" d fini par ($C(x) = 1-x$)

$$T(\text{max}) : i_{\text{max}}(x,y) = 0 \quad \text{si } x + y \leq 1 \\ = y \quad \text{sinon}$$

$$T(x+y-xy) : i_{\ddagger}(x,y) = \max(0, \frac{x+y-1}{1-x})$$

On remarque que si i (resp : u) est strict ou  gal   min (resp : max) alors : $T(i)$ (resp $T(u)$) n'est pas commutatif, ce qui en fait de mauvais op rateurs ensemblistes. Si u et i sont nilpotents et C -duaux alors on a

$$T(i) = u \quad ; \quad T(u) = i$$

c'est- -dire que (68) et (69) ne g n rent pas de nouveaux op rateurs.

Les op rateurs nilpotents semblent donc les plus coh rents et les plus riches en structure pour d finir la diff rence ensembliste et l'inclusion pour des ensembles flous.

On v rifie que $(B-A) \cup A \subseteq A \cup B$ (au sens de Zadeh) o  $B-A$ est form  avec \cup . On a l' galit  pour $\cup = \text{max}$.

2.3.5.3 - DIFFERENCE SYMETRIQUE ET EQUIVALENCE

En th orie classique des ensembles la diff rence sym trique Δ s' crit indiff remment

$$A \Delta B = (A \cup B) \cap (\bar{A} \cup \bar{B}) = (A-B) \cup (B-A) = (A \cap \bar{B}) \cup (B \cap \bar{A}) \quad (70)$$

appel es respectivement formes 1, 2 et 3.

En théorie des ensembles flous ces trois formulations ne sont plus équivalentes. On suppose que l'on choisit des opérateurs ensemblistes \cup, \cap , complémentation satisfaisant les lois de De Morgan.

On vérifie alors l'égalité des formes 1 et 3 seulement pour min et max, quelque soit la complémentation. En effet la démonstration de cette égalité nécessite la distributivité mutuelle des opérateurs \cup et \cap . Les formes 2 et 3 sont égales pour des opérateurs nilpotents et la complémentation de même générateur.

En revanche les opérateurs stricts tel le produit et la somme probabiliste donnent naissance à trois types d'opérateurs différents.

Avec max et min on trouve (cf. Dubois et Prade /11/).

$$\begin{aligned} \text{forme 1 et 3 : } & \max(\min(1-x,y), \min(1-y,x)) \stackrel{\Delta}{=} x \delta y \\ \text{forme 2 : } & x \Delta_0 y = \max(x,y) \text{ si } x \neq y, \text{ et } 0 \text{ sinon} \end{aligned}$$

Avec $\max(0, x+y-1)$ et $\min(x+y, 1)$, $C(x) = 1 - x$ on trouve (/14/)

$$\begin{aligned} \text{forme 1 et 2 : } & x \Delta^- y = |x-y| \\ \text{forme 3 : } & x \Delta^+ y = \min(x+y, 2-x-y) \\ \text{et on a } & x \Delta^- y \leq x \delta y \leq x \Delta^+ y \end{aligned}$$

L'équivalence étant le complémentaire de la différence symétrique on peut construire des égalités d'ensembles flous en posant

$$A = B \leftrightarrow \mu_{A \Delta B} = 0 \quad (71)$$

On peut vérifier que pour $\Delta = \Delta^-$ ou Δ_0 , l'égalité des ensembles flous correspond à celle des fonctions d'appartenance.

Pour $\Delta = \delta$ ou Δ^+ (71) signifie $A = B \in 2^\Omega$ ce qui signifie que seuls des ensembles non flous peuvent être considérés comme égaux.

Si $\mu_A = \mu_B$ on vérifie $\mu_{A \Delta B} \leq 0,5$ pour $\Delta = \delta$ tandis que l'on peut avoir $\mu_{A \Delta A}(\omega) = 1$ pour $\Delta = \Delta^+$, par exemple si $\mu_A(\omega) = .5$ (doute total quant à $\omega \in A$).

On peut vouloir affaiblir l'égalité d'ensembles flous en posant par exemple $A \succ B \leftrightarrow A \prec B$ et $B \prec A$ au sens de (61) (cf. Dubois et Prade /11/).

2.4 - EVALUATIONS SCALAIRES ET COMPARAISON D'ENSEMBLES FLOUS

On considère ici divers points de vue pour évaluer un ensemble flou ou une paire d'ensembles flous soit pour en mesurer un attribut soit pour chiffrer une ressemblance. On envisage seulement la construction d'évaluations scalaires. On verra plus loin que l'on peut construire des évaluateurs flous dont ces évaluateurs scalaires pourront parfois être vus comme des moyennes ou des points particuliers.

2.4.1 - Evaluations scalaires d'un ensemble flou

Etant donné un ensemble flou F sur un référentiel Ω , on peut vouloir situer cet ensemble flou par rapport aux autres en répondant à des questions telles que : à quel point est-il flou ? quelle est sa précision ? combien possède-t-il d'éléments ? quelle importance occupe-t-il dans Ω ? à quel point est-il non-vidé ? etc... A chacune de ces questions correspond une évaluation scalaire de l'attribut qu'elle concerne : on peut ainsi évoquer l'entropie ou indice de flou, la spécificité, la cardinalité, les deux dernières questions correspondent respectivement à des évaluateurs universels et existentiels.

2.4.1.1 - ENTROPIE OU INDICE DE FLOU

Un ensemble flou F est d'autant plus éloigné d'un ensemble classique que les éléments pour lesquels l'appartenance ou non à F est incertaine sont nombreux. Si $\omega \in \Omega$, l'incertitude maximale à ce propos est atteinte quand $\mu_F(\omega) = .5$. Mathématiquement, un indice de flou sera une fonction $f : [0, 1]^\Omega$ dans $[0, +\infty)$, telle que $f(F)$ sera d'autant plus grand que l'ensemble sera flou, soit les axiomes suivants (De Luca et Termini /5/).

$$f(F) = 0 \leftrightarrow F \in 2^\Omega$$

$$f(F) \text{ est maximal si et seulement si } \forall \omega \in \Omega, \mu_F(\omega) = .5$$

$$f(F) \leq f(F') \leftrightarrow \forall \omega, \left| \mu_F(\omega) - \frac{1}{2} \right| \geq \left| \mu_{F'}(\omega) - \frac{1}{2} \right|$$

$$f(F) = f(\bar{F}) \quad (F \text{ est aussi flou que son complément})$$

Comme solutions particulières de cet ensemble d'axiomes, dans le cas fini, sous la forme $f(.) = \sum_{\omega \in \Omega} f_\omega(.)$, on peut citer un analogue de l'entropie de Shanon (De Luca et Termini /5/)

$$\forall \omega \quad f_{\omega}(x) = -x \ln(x) - (1-x) \ln(1-x) \quad (72)$$

et l'indice de flou de Kaufmann /37/ qui se fonde sur une distance de Hamming :

$$f(F) = \sum_{\omega \in \Omega} |\mu_F(\omega) - \mu_{F.5}(\omega)| \quad (73)$$

où $F.5$ est la .5-coupe de F .

Pour une bibliographie sur cette notion voir le livre de Dubois et Prade les thèses d'Oppencheim /58/ et de Prade /62/.

L'intérêt de cette notion est de fournir un critère de choix parmi les solutions d'un problème quand celles-ci sont générées sous la forme d'ensembles flous (par exemple les équations de relations floues, cf chap.3).

L'indice de flou permet de choisir la solution la plus tranchée, l'ensemble dont les limites sont les plus abruptes.

2.4.1.2 - PRECISION

On notera qu'un ensemble flou dont l'indice de flou est nul peut très bien n'être pas un singleton, c'est-à-dire être imprécis. C'est une différence notable entre entropie floue et entropie probabiliste. On peut caractériser la précision par une mesure de "spécificité" telle que celle proposée par Yager /78/

$$Sp(F) = \int_0^{\bar{\alpha}} \frac{1}{|F_{\alpha}|} d\alpha \quad (74)$$

où $\bar{\alpha} = \sup \{\alpha \mid A_{\alpha} \neq \emptyset\}$ et $|F_{\alpha}|$ est le cardinal de F_{α} (fini)

On vérifie que $Sp(F) = 1 \leftrightarrow \exists \omega \in \Omega, F = \{\omega\}$

Si F et F' sont normalisés alors $F \subseteq F' \rightarrow Sp(F) \geq Sp(F')$

Si $\mu_F(\omega_1) = \pi_1 \geq \mu_F(\omega_n) = \pi_2 \geq \mu_F(\omega_n) = \pi_n$, et $\pi_{n+1} = 0$, alors

$$Sp(F) = \sum_{i=1}^n \frac{\pi_i - \pi_{i+1}}{i} \quad (75)$$

C'est-à-dire $Sp(F)$ est la probabilité du (ou d'un des) singleton de possibilité maximale au sens de la transformation possibilité \leftrightarrow probabilité définie chap. 1 § 6. Une autre définition possible de la précision d'un ensemble flou pourrait

d'ailleurs être l'entropie de Shannon de la mesure de probabilité calculée à partir de F par l'intermédiaire de cette bijection.

2.4.1.3 - CARDINALITE

Une extension naturelle de la cardinalité aux ensembles flous a été suggérée par De Luca et Termini /5/ sous la forme :

$$|F| = \sum_{\omega \in \Omega} \mu_F(\omega) \quad (\Omega \text{ fini}) \quad (76)$$

on vérifie que si F est un ensemble classique $|F|$ est bien sa cardinalité, que

$$\mu_F \leq \mu_{F'} \rightarrow |F| \leq |F'|$$

et aussi

$$|F \cap F'| + |F \cup F'| = |F| + |F'| \quad (77)$$

seulement pour \cap et \cup définies par des normes et co-normes duales de la famille de Frank (voir annexe) c'est-à-dire entre autres (\min, \max) , (\cdot, \div) , $(\max(0, x+y-1), \min(x+y, 1)) \dots$

On peut étendre $|F|$ en $|F|_p$ tel que (Kaufmann /38/)

$$|F|_p = \sum_{\omega \in \Omega} \mu_F(\omega)^p \quad (78)$$

on a alors $|F|_0 = |S(F)|$

mais (77) ne tient plus guère que pour (\min, \max) quand $p > 1$. La cardinalité est un cas particulier de mesure d'énergie d'un ensemble flou (De Luca et Termini /5/, Oppencheim /58/) telle que

$$e(F) = 0 \leftrightarrow F = \emptyset \quad (79)$$

$$F \leq F' \quad (\mu_F \leq \mu_{F'}) \rightarrow e(F) \leq e(F') \quad (80)$$

2.4.1.4 - EVALUATEURS EXISTENTIELS ET UNIVERSELS

Dubois et Prade /14/ ont affaibli les axiomes des mesures d'énergie. Un évaluateur g d'un ensemble flou F est une fonction de $[0, 1]^\Omega$ dans $[0, 1]$ (échelle conventionnelle) telle que

$$g(\emptyset) = 0 \quad ; \quad g(\Omega) = 1 \quad (81)$$

$$\mu_F \leq \mu_{F'} \rightarrow g(F) \leq g(F') \quad (82)$$

C'est en quelque sorte une mesure monotone d'incertitude définie sur des ensembles flous, similaire aux mesures de Sugeno.

$$g \text{ est dit existentiel si } g(F) = 0 \leftrightarrow F = \emptyset \quad (83)$$

$$\text{universel si } g(F) = 1 \leftrightarrow F = \Omega \quad (84)$$

complet s'il est existentiel et universel.

Une mesure d'énergie est un évaluateur existentiel à un facteur de normalisation près ($e(\Omega)$).

Exemples :

$$g_{\pi}(F) = \sup_{\omega \in \Omega} \mu_F(\omega) \quad (\text{existentiel})$$

$$g_N(F) = \inf_{\omega \in \Omega} \mu_F(\omega) \quad (\text{universel})$$

$$g_{||}(F) = \frac{|F|}{|\Omega|} \quad (\text{existentiel et universel})$$

$$g_{\pi N}(F) = \frac{g_{\pi}(F) + g_N(F)}{2} \quad (\text{idem})$$

$$g_e(F) = \frac{e(F)}{e(\Omega)}$$

où

$$e(F) = \sum_{\omega \in \Omega} f_{\omega}(\mu_F(\omega))$$

$$f_{\omega}^1(0) = 0, \quad f_{\omega} \text{ est croissante sur } [0, +\infty)$$

On remarquera que pour toute complémentation C, et $\bar{g}(F) = \text{Cog}(\bar{F})$

$$g \text{ est universel} \leftrightarrow \bar{g} \text{ est existentiel}$$

$$\text{on a } g_{\pi} = \bar{g}_N$$

De plus en considérant $C_{\omega}(x) = f_{\omega}^{-1}(f_{\omega}(1) - f_{\omega}(x))$ comme complémentation (dépendant de ω) on a en choisissant $f_{\omega}(1) = 1, \forall \omega$

$$g_e(F) + g_e(\bar{F}) = 1 \quad \forall F \quad (85)$$

Cette propriété est aussi valide pour $\frac{g_{\pi} + g_N}{2}$. Elle rappelle les probabilités.

Si on se donne un opérateur de moyenne m et un opérateur existentiel g alors

$$g_m = m(g, \bar{g}) \quad \text{est complet.}$$

Si m possède un générateur f que l'on peut toujours prendre tel que $f(0)=0$, $f(1)=1$, (cf. Annexe) alors

$$\forall F, \text{fog}_m(F) + \text{fog}_m(\bar{F}) = 1 \quad (86)$$

où

$$\mu_{\bar{F}} = \text{Co}\mu_F, C = f^{-1}(1-f)$$

On pourrait chercher à généraliser (77) aux évaluateurs complets qui ont des propriétés probabilistes, tandis que les évaluateurs purement existentiels ont un comportement possibiliste car mesurant à quel point un ensemble flou est vide ou non.

2.4.2 - Indices scalaires de comparaison d'ensembles flous

A l'aide des évaluateurs scalaires d'ensembles flous, on peut construire des indices de comparaison d'ensembles flous au sens de la ressemblance partielle, de l'inclusion, ou de la similarité. On résume ici les points saillants de l'article de Dubois et Prade /14/.

La ressemblance partielle mesure à quel point deux ensembles flous F et G sur Ω ont des éléments communs. Un indice de ressemblance partielle PM sera une application de $[0,1]^\Omega \times [0,1]^\Omega$ dans $[0,1]$, notée PM et telle que

$$i) \quad PM(F,G) = 0 \leftrightarrow F \cap G = \emptyset$$

$$ii) \quad PM(F,G) = PM(G,F)$$

$$iii) \quad F \leq G \text{ ou } G \leq F \rightarrow PM(G,F) = 1 \text{ ou } \overset{\circ}{G} \text{ et } \overset{\circ}{F} \text{ sont les noyaux de } G \text{ et } F.$$

Il est normal de construire PM à partir d'une mesure de $F \cap G$ par l'intermédiaire d'un évaluateur existentiel g , c'est-à-dire

$$PM(F,G) = \frac{g(F \cap G)}{g(F) * g(G)} \quad (87)$$

La condition iii) amène à poser $a * b = \min(a,b)$.

De même un indice d'inclusion I de F dans G sera défini par les axiomes suivants :

- i) $I(F,G) = 1 \leftrightarrow \overline{F} \cup G = \Omega$ (ou $F-G = \emptyset$)
- ii) $F \cap G = \emptyset \rightarrow I(F,G) = 0$

Il est naturel de définir I fonction d'un évaluateur universel de $\overline{F} \cup G$, soit $g(\overline{F} \cup G)$. On montre alors que la forme la plus naturelle de I est

$$I(F,G) = \frac{g(\overline{F} \cup G) - g(\overline{F})}{1 - g(\overline{F})} \quad (88)$$

qui n'est pas sans rappeler la règle de Dempster (cf. Ch. 1).

Enfin les indices de similarité S peuvent être caractérisés par les conditions

- i) $S(F,G) = 1 \leftrightarrow F \Delta G = \emptyset$
- ii) $F \cap G = \emptyset \rightarrow S(F,G) = 0$
- iii) $S(F,G) = S(G,F)$

S est alors naturellement défini à l'aide

- soit d'un évaluateur universel de $\overline{F \Delta G}$ on a alors *

$$S(F,G) = \frac{g(\overline{F \Delta G}) - g(\overline{F \cup G})}{1 - g(\overline{F \cup G})} \quad (89)$$

- soit par une combinaison symétrique d'indices d'inclusion, par une opération telle que $a \star b = 1 \leftrightarrow a = b = 1$; $0 \star 0 = 0$.

Une telle approche des indices de comparaison a permis de retrouver de nombreux évaluateurs proposés ça et là dans la littérature de la taxonomie et classification automatique (cf. Dubois et Prade /14/). Il suffit de particulariser g et de choisir un opérateur ensembliste flou. Par ailleurs Yager /77/ a défini un indice de flou en calculant la ressemblance partielle entre F et son complémentaire \overline{F} ; un ensemble a une frontière d'autant mieux définie qu'on parvient à le distinguer de son complémentaire.

En posant $g(F) = |F|/|\Omega|$, où |F| est le cardinal de F (cf.(76)), les formules (87), (88) et (89) fournissent des indices de comparaison de type probabiliste étudiés par Willmott /76/. g_{π} et $g_{\mathbb{N}}$ fournissent des indices possibilistes. ⁽¹⁾

L'étude de propriétés supplémentaires intéressantes pour PM,I,S (par exemple définir une transitivité pour I,S) est encore à faire, même si on connaît des résultats pour des cas particuliers.

(1) étudiés plus en détail au chapitre 4.

2.4.3 - Mesures d'incertitude d'événements flous

Un événement flou (= mal défini) peut être décrit par un ensemble flou. On peut donc essayer d'étendre les mesures d'incertitude décrites au chap. 1 à l'évaluation de la connaissance que l'on possède sur l'occurrence d'un événement flou. Ces mesures d'incertitude peuvent être supposées fournir des résultats scalaires ou des résultats imprécis. C'est le premier cas qu'on envisage ici. La seconde possibilité est considérée au chapitre suivant. Les mesures scalaires d'incertitude d'événements flous sont très liées aux évaluateurs décrits en 2.4.1.

2.4.3.1 - PROBABILITE D'UN EVENEMENT FLOU

Si (Ω, \mathcal{A}, P) est un espace de probabilités, \mathcal{A} étant une σ -algèbre sur Ω et P une mesure de probabilité, un événement flou F est décrit par une fonction d'appartenance μ_F Borel-mesurable ($\forall \alpha, F_\alpha \in \mathcal{A}$) la probabilité de F est définie, selon Zadeh /82/, par

$$P(F) = \int_{\Omega} \mu_F(x) dP(x) \quad (90)$$

C'est l'espérance mathématique de la fonction d'appartenance. (90) généralise la cardinalité d'un ensemble flou (76).

On peut vérifier que si $\mu_{\bar{F}} = 1 - \mu_F$, $P(\bar{F}) = 1 - P(F)$. De plus si

$\mu_F \wedge \mu_G + \mu_F \vee \mu_G = \mu_F + \mu_G$ (\wedge et \vee sont des opérateurs duaux construits sur les normes triangulaires et co-normes de Frank - voir Annexe) alors

$$P(F \wedge G) + P(F \vee G) = P(F) + P(G) \quad (91)$$

par exemple les trois paires (\min, \max) , $(xy, x+y-xy)$, $(\max(0, x+y-1), \min(1, x+y))$ d'opérateurs préservent (91). Les récents travaux de Klement (/41 - 44/) étudient systématiquement le concept de σ -algèbre floue (au sens d'une co-norme) et établissent dans quelles conditions des fonctions P satisfaisant (91) se mettent sous la forme (90).

La notion de conditionnement par événement flou a été définie par Zadeh ($P(F/G) = P(F \wedge G) / P(G)$ ou \wedge est définie par le produit) et a reçu un début d'étude par Smets /71/

2.4.3.2 - POSSIBILITE D'UN EVENEMENT FLOU

Si on reste fidèle à la sémantique du concept de possibilité, on est amené, comme Zadeh /87/ , à utiliser pour mesurer la possibilité d'un événement flou F un évaluateur de ressemblance partielle entre F et l'ensemble flou F_π que définit une distribution de possibilité π .

$$\Pi_i(F) = \sup_{\omega \in \Omega} i(\mu_F(\omega), \pi(\omega)) \quad (91)$$

Si F est normalisé, (91) est un cas particulier de (87). L'opérateur minimum peut bien sûr être utilisé dans (91). On note $\Pi(F) = \Pi_{\min}(F)$. Si $i =$ produit on obtient ce que Kaufmann /39/ appelle admissibilité.

On vérifie les propriétés suivantes :

$$\forall F \quad \Pi_i(F) \leq \Pi(F)$$

. Si l'union est au sens du maximum alors, pour toute norme triangulaire i

$$\forall F, G \in [0, 1]^\Omega \quad \Pi_i(F \cup G) = \max(\Pi_i(F), \Pi_i(G))$$

$$\text{. Si } F \subseteq G \text{ (} \mu_F \leq \mu_G \text{) alors } \Pi_i(F) \leq \Pi_i(G)$$

$$\text{. } \Pi_i(F) = 0 \iff F \cap F_\pi = \emptyset \quad \text{où } \cap \text{ est définie par } i$$

$$\text{. } F \subseteq F_\pi^{\circ} \rightarrow \Pi_i(F) = \sup \mu_F ; \quad F_\pi \subseteq F^{\circ} \rightarrow \Pi_i(F) = 1$$

De même, la nécessité d'un événement flou correspond à l'inclusion de F_π dans F, et peut donc s'exprimer à l'aide d'un indice d'inclusion.

$$N_{u,C}(F) = \inf_{\omega \in \Omega} u(\mu_F(\omega), C\pi(\omega)) \quad (92)$$

pour une co-norme u et une complémentation C .

Si i est C -dual de u alors

$$\Pi_i(F) = C \left[N_i(\bar{F}) \right] \quad \text{où} \quad \mu_{\bar{F}} = C\mu_F$$

Quand $C(x) = 1-x$ on pose $N_u = N_{u,C}$ et quand $u = \max$ on pose $N = N_u$.

On vérifie les propriétés suivantes :

. $\forall u, N_u(F) \geq N(F)$ pour tout F

. Si l'intersection est définie par le minimum alors

$$\forall u, C, N_{u,C}(F \cap G) = \min(N_{u,C}(F), N_{u,C}(G))$$

. $N_{u,C}(F) = 1 \leftrightarrow \bar{F} \cap F = \Omega$

. $F \cap F = \emptyset \rightarrow N_{u,C}(F) = 0$ (π est normalisé)

. Si $i = 1 - u(1-, 1-)$, alors $\forall F, N_u(F) = 1 - \Pi_i(\bar{F}) \leq \Pi_i(F)$ (1)

. $\max(\Pi(F), \Pi(\bar{F})) \geq \frac{1}{2}$; $\min(N(F), N(\bar{F})) \leq \frac{1}{2}$

Donc la notion de possibilité d'un événement flou au sens de Zadeh (1978) conserve l'essentiel des propriétés des mesures de possibilité. Les inégalités ci-dessus, qui sont plus faibles que les égalités $\max(\Pi(A), \Pi(\bar{A})) = 1$ $\min(N(A), N(\bar{A})) = 0$ pour des événements non flous, sont dues à l'absence de tiers exclu et de loi de contradiction pour max et min. Néanmoins on peut remarquer que pour $u =$ somme bornée et sa duale i

$$\min(\Pi_i(F) + \Pi_i(\bar{F}), 1) = 1 \quad \forall i$$

$$\max(0, N_u(F) + N_u(\bar{F}) - 1) = 0$$

NB : on peut bien sûr définir le degré de nécessité d'un événement flou à partir des opérateurs d'inclusion définis en 2.3.5.2. Les possibilités duales ainsi obtenues seront dissymétriques en π et μ_F , ce qui dans ce contexte n'est pas choquant.

2.4.3.3 - PROBABILITES SUPERIEURES ET INFERIEURES D'EVENEMENTS FLOUS

Smets /70/ a tiré parti des concepts d'espérance mathématique supérieures et inférieures (cf. Ch.1) pour étendre (90), c'est-à-dire la probabilité d'un événement flou au sens de Zadeh à des probabilités supé-

(1) Preuve : $\Pi_i(F) + \Pi_i(\bar{F}) = \sup i(\mu_A, \pi) + \sup i(1 - \mu_A, \pi)$
 $\geq \sup [i(\mu_A, \pi) + i(1 - \mu_A, \pi)]$
 et si $\pi(\omega) = 1$ $\geq i(\mu_A(\omega), 1) + i(1 - \mu_A(\omega), 1) = 1$

rieures - on a alors, si Ω est fini et m une probabilité de base

$$P^*(F) = E^*(\mu_F) = \sum_{A \subseteq \Omega} m(A) \max_{\omega \in A} \mu_F(\omega) \quad (93)$$

$$P_*(F) = E_*(\mu_F) = \sum_{A \subseteq \Omega} m(A) \min_{\omega \in A} \mu_F(\omega) \quad (94)$$

Si m définit une fonction de possibilité, on obtient par (93) et (94) de nouvelles définitions de la possibilité et de la nécessité d'un événement flou qui ne coïncident pas avec (91) et (92) : en particulier

$$P^*(F \cup G) \neq \max(P^*(F), P^*(G))$$

$$P_*(F \cap G) \neq \min(P_*(F), P_*(G))$$

Donc ce ne sont plus des mesures de possibilité ou de nécessité au sens strict. Néanmoins Smets /70/ a montré que les probabilités supérieures et inférieures d'événement flous satisfaisaient encore les propriétés de monotonie d'ordre infini, avec l'intersection et l'union définies par min et max respectivement. Elles restent donc des mesures de plausibilité et de crédibilité au sens de Shafer

2.4.3.4 - MESURE D'INCERTAIN DECOMPOSABLE D'UN EVENEMENT FLOU

Si on considère une mesure de pseudo-probabilité (cf. Ch. 1) g telle que $\hat{\psi}_0 g$ est une mesure de probabilité sur Ω fini on peut poser, conformément à (90),

$$g(F) = \hat{\psi}^{-1} \left(\sum_{\omega_i \in \Omega} \mu_{F_i}(\omega_i) \hat{\psi}(g_i) \right) \quad (95)$$

où g_i est la densité associée à g .

Si $F \cap G = \emptyset$ on voit que, si $*$ est la co-norme de générateur $\hat{\psi}$

$$g(F \cup G) = g(F) * g(G) \quad (96)$$

est encore valide ($F \cup G$ est alors indépendant de la co-norme définissant \cup).

Plus généralement si \cup et \cap sont définis par des normes et co-normes duales de la famille de Frank ($u(x,y)+i(x,y) = x+y$) - par exemple $i = \min$, produit, $\max(0, x+y-1)$ - on a encore

$$g(F) \star g(G) = g(F \cap G) \star g(F \cup G) \quad (97)$$

Donc (95) semble une bonne définition.

Dans le cas de pseudo-possibilité, basée sur une co-norme stricte une définition peut être suggérée à partir d'une mesure basée sur la co-norme $x+y-xy$, conformément à (90) et (95).

$$g(A) = 1 - \prod_{\omega_i \in A} (1-g_i) \quad (98)$$

d'où, en notant la forme du générateur $\varphi(x) = -\text{Log}(1-x)$.

$$g(F) = 1 - \prod_{\omega_i \in \Omega} (1-g_i)^{\mu_F(\omega_i)} \quad (99)$$

(Avec la convention $0^0 = 1$). Plus généralement si h est une fonction croissante continue, $h(0) = 0$, $h(1) = 1$, alors $1-h^{-1}(h(1-.) \circ h(1-.)$) est une co-norme stricte et (99) s'étend en

$$g(F) = 1 - h^{-1}\left(\prod_{\omega_i \in \Omega} h(1-g_i)^{\mu_F(\omega_i)}\right) \quad (100)$$

Néanmoins une définition conforme à la possibilité d'un événement flou au sens de Zadeh pourrait être envisagée pour n'importe quelle pseudo-possibilité (par forcément basée sur une co-norme stricte) :

Soit Fg le sous-ensemble flou défini par $\mu_{Fg}(\omega_i) = g_i$ (Fg est normalisé).

On peut poser

$$g(F) = \star_{\omega_k \in \Omega} \mu_{Fg} \wedge^F(\omega_i) = \star_{\omega_k \in \Omega} i(g_k, \mu_F(\omega_k)) \quad (101)$$

où i est un opérateur d'intersection; (101) généralise (91) et est un opérateur de ressemblance partielle.

On omet ici de fournir les expressions duales de (98)-(101) qui se construisent facilement.

NB : Une alternative possible à (95) est

$$g(F) = \hat{\varphi}^{-1}\left(\sum_{\omega_i \in \Omega} \hat{\varphi}(\mu_F(\omega_i)g_i)\right)$$

Elle vérifie (96) mais jamais (97) sauf pour $\cup = \max$ et $\cap = \min$. On constate les mêmes propriétés pour la troisième alternative possible suggérée par Schwyhla (cf. chap.1 § 3.5) $g(F) = \hat{\varphi}^{-1}\left(\sum_{\omega_i \in \Omega} [\hat{\varphi} \circ \mu_F(\omega_i)] \cdot \hat{\varphi}(g_i)\right)$.

2.4.3.5 - MESURES GENERALES D'INCERTITUDE D'UN EVENEMENT FLOU

Sugeno (1974) a mesuré l'incertitude d'un événement flou au sens d'une fonction d'ensemble monotone quelconque g , par l'intermédiaire de ce qu'il appelle "l'intégrale floue".

$$g(F) = \int_{\Omega} \mu_F \circ g(.) = \sup_{\alpha} \min(\alpha, g(F_{\alpha})) \quad (102)$$

en supposant que pour tout α , F_{α} est évaluable par g .

Bien que (102) ait quelques analogies de propriétés avec l'intégrale de Lebesgue, elle ne correspond pas à l'idée d'espérance mathématique (= de valeur moyenne) mais à celle de médiane (cf Kandel /38/, Dubois et Prade /11/ p.134) puisque dans le cas fini, si $\{\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_p\} = \mu_F(\Omega)$, $g(F)$ est le terme médian de l'ensemble à $2p+1$ éléments $\{\alpha_i, i=1, p\} \cup \{g(F_{\alpha_i}) | i=2, p\}$. D'ailleurs $g(F)$ ne généralise pas la probabilité d'un événement flou au sens de Zadeh (1968) (cf 90) ni la probabilité supérieure d'un événement flou au sens de Smets. En revanche, si $g = \Pi$ est une mesure de possibilité alors, (Dubois et Prade /11/).

$$g(F) = \Pi(F) \quad (103)$$

C'est-à-dire que l'intégrale floue généralise la possibilité d'un événement flou au sens de Zadeh (1978) (cf. (91)).

Néanmoins, comme l'a montré Höhle /32/ on peut étendre la notion d'espérance mathématique à des fonctions d'ensemble monotones quelconques. En effet (93) s'écrit encore (Smets, /70/) :

$$P^*(F) = \sum_{i=1}^p \alpha_i (P^*(F_{\alpha_i}) - P^*(F_{\alpha_{i+1}})) \quad (104)$$

avec $F_{\alpha_{p+1}} = \emptyset$, et $\mu_F(\Omega) = \{\alpha_1 < \alpha_2 < \dots < \alpha_p\}$

(104) ne dépend pas de l'existence d'une probabilité de base m , et peut donc être prise comme définition à la place de (102), pour généraliser la probabilité d'un événement flou au sens de Zadeh /82/ .

De même, (94) peut s'écrire

$$P_{\star}(F) = \sum_{i=1}^p \alpha_i (P_{\star}(F_{\alpha_i}) - P_{\star}(F_{\alpha_{i+1}})) \quad (105)$$

et on vérifie

$$P_{\star}(F) = E_{\star}(\mu_F) = 1 - E_{\star}(1 - \mu_F) = P_{\star}(\bar{F})$$

Si g est une mesure de Sugeno quelconque, (104) ou (105) s'exprime donc

$$g(F) = \sum_{i=1}^p \alpha_i (g(F_{\alpha_i}) - g(F_{\alpha_{i+1}})) \quad (106)$$

On peut considérer comme légitime de proposer, simultanément à (106), la quantité $\bar{g}(F)$, où $\forall A \subseteq \Omega$, $\bar{g}(A) = 1 - g(\bar{A})$. Néanmoins, dans le cas général on ne peut rien dire sur la valeur relative de $g(F)$ et $\bar{g}(F)$, bien que la relation $\bar{g}(F) = 1 - g(\bar{F})$ soit encore valide.

Il est intéressant de considérer la formule duale de (102), à savoir $\int_{\Omega} \mu_F \circ \bar{g}(\cdot)$; on notera que la quantité $1 - g(\bar{F})$ vaut

$$\begin{aligned} 1 - \int_{\Omega} (1 - \mu_F) \circ g(\cdot) &= \inf_{\alpha} \max(1 - \alpha, g(\bar{F})_{\alpha}) \\ &= \inf_{\beta} \max(\beta, \bar{g}(F_{\bar{\beta}})) \end{aligned} \quad (107)$$

où

$$F_{\bar{\beta}} = \{\omega \mid \mu_F(\omega) > \beta\}$$

On montre que généralement

$$\int_{\Omega} \mu_F \circ g(\cdot) = \inf_{\alpha} \max(\alpha, g(F_{\alpha})) \quad (108)$$

De sorte que

$$\int_{\Omega} \mu_F \circ \bar{g}(\cdot) = 1 - \int_{\Omega} \mu_{\bar{F}} \circ g(\cdot)$$

est généralement valide.

En particulier

$$N(F) \stackrel{\Delta}{=} 1 - \prod(\bar{F}) = \sup_{\alpha} \min(\alpha, N(F_{\alpha}))$$

REFERENCES

- /1/ ALSINA C., TRILLAS E., VALVERDE L. (1980)
On non-distributive logical connectives for fuzzy sets theory.
BUSEFAL n°3, Eté 1980, L.S.I. Univ. P. Sabatier, Toulouse,
 pp.18-29
- /2/ BELLMAN R.E., GIERTZ M. (1973)
On the analytic formalism of the theory of fuzzy sets. Information Sciences vol.5, pp.149-156.
- /3/ BOURBAKI N. (1965)
Topologie Générale. Chap. I Structures topologiques -
 Hermann, Paris
- /4/ CHU A.T.W., KALABA R.E., SPINGARN K. (1979)
*A comparison of two methods for determining the weights of
 belonging to fuzzy sets.* J. of Optimization Theory and Appl.
 Vol.27, n°4, pp.531-538.
- /5/ DE LUCA A., TERMINI S. (1972b)
*A définition of a non-probabilistic entropy in the setting
 of fuzzy sets theory.* Information and Control Vol.20, pp.301-312.
- /6/ DOMBI J. (1982)
*A general class of fuzzy-operators, the De Morgan class of
 fuzzy-operators and fuzziness measures induced by fuzzy
 operators.* Fuzzy Sets and Systems 8, 149-163.
- /7/ DOMBI J. (1982)
*Basic concepts for a theory of evaluation : the aggregative
 operator.* European J. Operations Research 10, 282-293
- /8/ DUBOIS D. (1980)
Triangular norms for fuzzy sets. Proc. 2nd Int. Seminar on
 Fuzzy Set Theory, Johannes Kepler Univ. Linz, Austria, Sept.
 15-20, 1980 (E.P. Klement, ed.) pp.39-68.

- /9/ DUBOIS D. (1980)
Ensembles flous et conception assistée par ordinateur.
 Rapport de recherche N°199, IMAG, Grenoble.
- /10/ DUBOIS D., PRADE H. (1979a)
 New results about properties and semantics of fuzzy-set-theoretic operators. Proc. 1st Int. Symp. on Policy Analysis & Information Systems. Durham, North-Carolina, June 28-30, 1979, pp.167-174. Version révisée in "Fuzzy Sets : Theory and Applications to Policy Analysis and Information Systems (Paul P. Wang, S.K. Chang, eds). Plenum Publ. Corp. New-York, 1980, pp.59-75.
- /11/ DUBOIS D., PRADE H. (1980)
Fuzzy Sets and Systems : theory and applications.
 Academic Press, New-York.
- /12/ DUBOIS D., PRADE H. (1981)
 Unfair coins and Necessity Measures. Proc. 3d Int. Seminar on fuzzy set theory (E.P. Klement, ed.) - J. Kepler Universität Linz, 223-231. Version révisée dans Fuzzy Sets and Systems vol.10, n°3, 1983.
- /13/ DUBOIS D., PRADE H. (1982b)
 Criteria aggregation and ranking of alternatives in the framework of fuzzy set theory. - A paraître dans "Fuzzy Sets and Decision Analysis" (B.R. Gaines, H. J. Zimmermann, L.A. Zadeh, eds). TIMS studies in Management Sciences, North Holland. Résumé par D. Dubois sous le titre "Quelques classes d'opérateurs remarquables pour combiner des ensembles flous" BUSEFAL n°1, Hiver 1979-80; L.S.I., Univ. P. Sabatier, Toulouse pp.29-35.
- /14/ DUBOIS D., PRADE H. (1982)
 A unifying view of comparison indices in a fuzzy set theoretic framework in "Fuzzy Set and Possibility Theory : Recent Developments, (R.Yager Ed.) Pergamon Press, Oxford, 3-13 .
- /15/ DUBOIS D., PRADE H. (1982)
 Towards fuzzy differential calculus part 2 : integration on fuzzy intervals, Fuzzy Sets and Systems 8, 105-116.

- /16/ DUBOIS D., PRADE H. (1982)
Fuzzy Sets and Statistical data. Présenté à EURO V, Lausanne, Suisse.
- /17/ ESTEVA F., TRILLAS E., DOMINGO X. (1981)
Weak and strong negation functions for fuzzy set theory. Proc. 11th IEEE Int. Symp. on Multiple-Valued Logic, Oklahoma City, pp.23-26.
- /18/ FORTET R., KAMBOUZIA M. (1976)
Ensembles aléatoires et ensembles flous. Publications Econométriques Vol.IX, Fasc.1.
- /19/ FRANK M.J. (1979)
On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$. A equat. Math., 19, 194-226.
- /20/ FUNG L.W., FU K.S. (1974)
The k -th Optimal policy for decision-making in Fuzzy Environments. In : "Identification and Systems Parameter identification". (P. Eykhoff, ed.) North Holland 1052-1059.
- /21/ GENTILHOMME Y. (1968)
Les ensembles flous en linguistique. Cahiers de Linguistique Théorique et Appliquée (Bucarest) t.V, pp.47-63.
- /22/ GOGUEN J.A. (1967)
L-fuzzy sets. J. of Math. Anal. & Appl., Vol. 18, n°1, pp.145-174
- /23/ GOGUEN J.A. (1971)
The logic of inexact concepts, Synthèse 19, 325-373.
- /24/ GOGUEN J.A. (1974)
Concept representation in natural and artificial languages : Axioms, extensions and applications for fuzzy sets. Int. J. Man-Machine Studies, Vol. 6, pp.513-561.
- /25/ GOODMAN I.R. (1981)
Fuzzy sets as equivalence classes of random sets. Dans "Fuzzy Sets and Possibility Theory : Recent Developments" (R.R. Yager, ed.) Pergamon Press.
- /26/ GOTTWALD S. (1979)
Set theory for fuzzy sets of higher levels. Fuzzy Sets and Systems, Vol.1, pp.125-151.

- /27/ HAMACHER H. (1975)
Über logische Verknüpfungen Unscharfer Aussagen und deren Zugehörige Bewertungsfunktionen. In : Progress in Cybernetics and Systems Research Vol.3 (R. Trappl, G.J. Klir, L. Ricciardi, eds.). Hemisphere Publ. Corp., New York, pp.276-287, 1978.
- /28/ HERSH H.M., CARAMAZZA A. (1976)
A fuzzy set approach to modifiers and vagueness in natural languages. J. of Exp. Psychol. : General vol.105, pp.254-276.
- /29/ HERSH H.M., CARAMAZZA A., BROWNELL H.H. (1979)
Effects of context on fuzzy membership functions. Dans "Advances in Fuzzy Set Theory and Applications (M.M Gupta R.K. Ragade, R.R. Yager, eds.), pp.389-408. North-Holland.
- /30/ HIGASHI M., KLIR G.J. (1982)
On measures of fuzziness and fuzzy complements. Int. J. of General Systems, vol. 8 n°3.
- /31/ HIROTA K. (1981)
Concepts of probabilistic sets, Fuzzy Sets and Systems, Vol.5 n°1, 31-46
- /32/ HÖHLE H. (1982)
Integration with respect to fuzzy measures, IFAC Symposium on Theory and Application of Digital Control, New-Delhi.
- /33/ HÖHLE H. (1982)
Fuzzy filters as a basis for the definition of fuzzy subsets and fuzzy points, 4th Int. Seminar on fuzzy set theory, J. Kepler Universität Linz, Autriche.
- /34/ KAMPE DE FERIET J. (1980)
Une interprétation des mesures de plausibilité et de crédibilité au sens de G. Shafer et de la fonction d'appartenance définissant un ensemble flou de L. Zadeh. Pub. IRMA - Lille (France), Vol.2, fasc. 6, n°2, pp. II-01-II-22.
- /35/ KAMPE DE FERIET J. (1982)
Interprétation of membership functions of fuzzy sets in terms of plausibility and belief, in Fuzzy information and Decision processes (M.M. Gupta, E. Sanchez, Eds) North Holland 93-98

- /36/ KANDEL A., BYATT W.J. (1978)
Fuzzy sets, fuzzy algebra and fuzzy statistics, Proceedings of the IEEE, 68, 1619-1639.
- /37/ KAUFMANN A. (1973)
*Introduction à la Théorie des Sous-Ensembles Flous.
 1 Eléments Théoriques de Base. Masson, Paris.*
- /38/ KAUFMANN A. (1977)
*Introduction à la Théorie des Sous-Ensembles Flous. Vol.4
 Compléments et Nouvelles Applications. Masson, Paris.*
- /39/ KAUFMANN A. (1978)
Le calcul des admissibilités. Colloque International sur la théorie et les applications des ensembles flous, Marseille.
- /40/ KLEMENT E.P. (1980)
Fuzzy σ -algebras and fuzzy measures with respect to t-norms, table ronde CNRS, Université de Lyon.
- /41/ KLEMENT E.P. (1980a)
Fuzzy σ -algebras and fuzzy measurable functions. Fuzzy Sets and Systems, vol.4, pp.83-93.
- /42/ KLEMENT E.P. (1980b)
Characterization of finite fuzzy measures using Markoff-kernels. J. Math. Anal. & Appl. Vol.75, pp.330-339.
- /43/ KLEMENT E.P. (1980c)
Characterization of fuzzy measures by classical measures. Dans Fuzzy Sets : Theory and Applications to Policy Analysis and Information Systems (P.P. Wang, S.K. Chang, eds.) Plenum Press, pp.25-33.
- /44/ KLEMENT E.P. (1982)
Construction of fuzzy σ -algebras using triangular norms, J. Math. Anal. & Applications 85, 543-566,
- /45/ KLEMENT E.P. (1982)
Opérations on fuzzy sets : an axiomatic approach, to appear in Information Sciences.
- /46/ KRANTZ D.H., LUCE R., SUPPES P., TVERSKI A. (1971)
Foundations of Measurement. Vol.1, Academic Press, New York.

- /47/ KUZ'MIN V.B. (1981)
*A parametric approach to description of linguistic values of variables and hedges. Fuzzy Sets and Systems, Vol.6 pp.27-41.
 Errata p.205.*
- /48/ LAKOFF G. (1973)
Hedges : A study in meaning criteria and the logic of fuzzy concepts. J. of Philos. Logic. Vol.2, pp.458-508.
- /49/ LOWEN R. (1978)
On fuzzy complements. Information Sciences 14, 107-113.
- /50/ MIZUMOTO M., TANAKA K. (1976)
Some properties of fuzzy sets of type 2. Information and Control vol. 31, pp.312-340
- /51/ MIZUMOTO M. (1981)
Fuzzy sets and their operations I. Information and Control 48, 30-48.
- /52/ MIZUMOTO M. (1982)
Fuzzy sets and their operations II. Information and Control 50, 160-174
- /53/ Mc VICAR WHELAN P.J. (1978)
*Fuzzy sets, the concept of height and the hedge very
IEEE Trans. on Systems Man, and Cybernetics, 8, 507511.*
- /54/ NEVEU J. (1970)
Calcul des Probabilités. Masson, Paris.
- /55/ NORWICH A.M., TURKSEN I.B. (1982)
*The fundamental measurement of fuzziness, in "Fuzzy Set and Possibility Theory : Recent Advances". (R. Yager, ed.)
 Pergamon Press, 49-60.*
- /56/ NORWICH A.M., TURKSEN I.B. (1982)
Stochastic Fuzziness in "Approximate Reasoning in Decision Analysis" (M.M. Gupta, E. Sanchez Eds) North Holland, 13-22.
- /57/ ODEN C.G., LOPES L.L. (1982)
*On the internal structure of fuzzy subjective categories in
"Fuzzy Set and Possibility Theory : Recent advances (R.Yager ed.) Pergamon Press, 75-89.*

- /58/ OPPENCHAIM S. (1980)
Classification et distances énergétiques entre sous-ensembles flous, Thèse 3ème Cycle Université Tech. de Compiègne.
- /59/ OVCHINNIKOV S. (1980)
 Involutions in fuzzy set theory. Stochastica, VI, N°3, 227-231.
- /60/ PEDRYCZ W. (1982)
Fuzzy Control and Fuzzy Systems, Dept. Mathematics, Delft. University of Technology, Holland.
- /61/ PRADE H. (1980)
 Unions et intersections d'ensembles flous, BUSEFAL n°3, 58-62, Univ. P. Sabatier, Toulouse.
- /62/ PRADE H. (1982)
Modèles mathématiques de l'Imprécis et de l'Incertain en vue d'applications au raisonnement naturel, Thèse d'Etat Univ. Paul Sabatier - Toulouse
- /63/ RALESCU D. (1979)
 A survey of the representation of fuzzy concepts and its applications, IN : Advances in Fuzzy Set Theory and Applications (M.M Gupta, R.K. Ragade, R.R. Yager Eds) North Holland '77-91.
- /64/ RALESCU D., ADAMS G. (1980)
 The fuzzy integral. J. Math. Anal. Appl. 75, 562-570
- /65/ RESCHER N. (1969)
Many-valued logic. Mc Graw Hill
- /66/ SAATY T.L. (1974)
 Measuring the fuzziness of sets. J. Cybernetics, Vol. 4 n°4 53-61
- /67/ SANCHEZ E. (1976)
 Resolution of composite fuzzy relation equations, Information and Control, 30, 38-48
- /68/ SILVERT W. (1979)
 Symmetric Summation : a class of operations on fuzzy sets. IEEE Trans. on Systems, Man, and Cybernetics 9, 657-659

- /69/ SINGER D. (1980)
Etude des demi-groupes abeliens totalement ordonnés, en vue d'une application à la théorie des ensembles flous. Mémoire de DEA, Dept. de Mathématiques, Université de Lyon.
- /70/ SMETS P. (1981)
The degree of belief in a fuzzy event. Information Sciences 25, 1-19.
- /71/ SMETS P. (1982)
Probability of a fuzzy event : an axiomatic approach, Fuzzy Sets and Systems, 7, 153-164.
- /72/ SMETS P. (1982)
Elementary Semantic Operators in "Fuzzy set and possibility theory : Recent Developments" (R.R. Yager Ed.) Pergamon Press, 247-257.
- /73/ THÖLE U., ZIMMERMANN H.J., ZYSNO P. (1979)
On the suitability of minimum and product operators for the intersection of fuzzy sets. Fuzzy Sets and Systems, vol. 2, pp. 167-180.
- /74/ TRILLAS E. (1979)
Sobre funciones de negacion en la teoría de conjuntos difusos. Stochastica, vol. III, n°1, pp. 47-59.
- /75/ VALVERDE L. (1982)
Indistinguishability relations with implicative part 4th Int. Seminar on fuzzy set theory, J. Kepler Universität Linz, Autriche.
- /76/ WILLMOTT T. (1981)
Mean Measures of containment and equality between fuzzy sets Proc 11th IEEE Int. Symp. on Multiple-valued logic, Oklahoma City, 183-190.
- /77/ YAGER R.R. (1979)
On the measure of fuzziness and negation Part 1 : membership in the unit interval, Int. J. General Systems 5, 221-229.
Part 2 : Lattices : Information and Control 44, 236-260, 1980.
- /78/ YAGER R.R. (1979)
Measuring tranquility and anxiety in decision-making : an application of fuzzy sets.

- /79/ YAGER R.R. (1980)
On a general class of fuzzy connectives, Fuzzy Sets and Systems
 4, 235-242.
- /80/ YAGER R.R. (1982)
Some questions related to linguistic variables, BUSEFAL n°10
 54-65, LSI, Université Paul Sabatier, Toulouse.
- /81/ ZADEH L.A. (1965)
Fuzzy sets. Information & Control. Vol. 8, pp.338-353.
- /82/ ZADEH L.A. (1968)
Probability measures of fuzzy events. J. Math. Anal. & Appl.
 Vol.23, pp.421-427.
- /83/ ZADEH L.A. (1971)
Similarity relations and fuzzy orderings. Information Sciences
 Vol.3, pp.177-200.
- /84/ ZADEH L.A. (1971)
Quantitative fuzzy semantics. Information Sciences, Vol.3,
 pp.159-176.
- /85/ ZADEH L.A. (1972)
A fuzzy set-theoretic interpretation of linguistic hedges.
J. of Cybernetics, Vol.2 n°3, pp.4-34.
- /86/ ZADEH L.A. (1975)
The concept of a linguistic variable and its applications to
approximate reasoning. Information Sciences. Part. 1, Vol. 8,
pp.199-249 ; Part 2, Vol. 8, pp.301-357 ; Part 3, Vol.9, pp.
 43-80.
- /87/ ZADEH L.A. (1978)
Fuzzy Sets as a basis for a theory of possibility, Fuzzy sets
and Systems, Vol.1, 3-28.
- /88/ ZIMMERMANN H.J. (1978)
Results of empirical studies in fuzzy set theory. dans
"Applied General Systems Research" (G. Klir, ed.) pp.303-312,
Plenum Press.
- /89/ ZIMMERMANN H.J. ZYSNO P. (1980)
Latent connectives in human decision making. Fuzzy Sets and
Systems. Vol.4 n°1, pp.37-51.

CHAPITRE 3

PRINCIPES D'EXTENSION

Dans ce chapitre, on étudie comment un ensemble flou sur un domaine peut être induit par une fonction à valeurs sur ce domaine à partir d'autres ensembles flous qui limitent le ou les arguments de cette fonction. Cela correspond dans l'optique probabiliste à l'étude des fonctions de variables aléatoires. Un équivalent possibiliste de la notion de probabilité jointe va naturellement intervenir : c'est la relation floue dont on donnera un aperçu des propriétés et problèmes qu'elle pose. Ce chapitre contient en fait l'outil de base pour les deux chapitres suivants.

3.1 - TRANSPORT D'UNE MESURE D'INCERTITUDE PAR UNE APPLICATION SIMPLE

Soit Ω un ensemble muni d'une mesure d'incertitude g monotone. Soit \mathcal{A} l'ensemble des parties de Ω sur lequel g est défini. On se donne une application $f: \Omega \rightarrow U$; alors si un ensemble $A \subseteq U$ est tel que $f^{-1}(A) \triangleq \{\omega: f(\omega) \in A\} \in \mathcal{A}$, on peut calculer la mesure d'incertitude g_f de A déduite de g par f :

$$g_f(A) = g(f^{-1}(A)) \quad (1)$$

L'espace d'incertitude (Ω, \mathcal{A}, g) induit donc un autre espace d'incertitude (U, \mathcal{B}, g_f) où $\mathcal{B} = \{A/f^{-1}(A) \in \mathcal{A}\}$.

g_f est bien une mesure monotone puisque

$$A \subseteq B \rightarrow f^{-1}(A) \subseteq f^{-1}(B) \rightarrow g_f(A) \leq g_f(B)$$

On a par exemple (c'est intéressant dans le cas fini)

$$\mathcal{A} = 2^\Omega \rightarrow \mathcal{B} = 2^U$$

On peut, pour définir (1), partir de deux espaces topologiques (Ω, \mathcal{A}) et (U, \mathcal{B}) et d'une fonction continue f ; (1) définit une mesure d'incertitude g_f sur (Ω, \mathcal{B}) .

On peut aussi choisir pour \mathcal{A} et \mathcal{B} des σ -algèbres, et pour f une fonction mesurable ($f^{-1}(A) \in \mathcal{A} \forall A \in \mathcal{B}$).

On peut choisir pour g une probabilité supérieure (on sait alors d'après le chapitre 1 que g_f en sera une, et $1-g_f(\bar{A})$ définira une probabilité inférieure).

On peut aussi transporter une fonction d'ensemble décomposable. Quand Ω est fini, et g est basé sur une co-norme \star alors

$$\forall A \subseteq U, \quad g_f(A) = g(f^{-1}(A)) = \sum_{\omega_i \in f^{-1}(A)} \star g_i \quad (2)$$

où $\{g_1 \dots g_n\}$ est la densité associée à g .

Il est facile de vérifier que g_f est encore une mesure d'incertitude basée sur \star et que

$$\begin{aligned} g_f(A \cup B) &= g(f^{-1}(A \cup B)) = g(f^{-1}(A) \cup f^{-1}(B)) \\ &= g_f(A) \star g_f(B) \end{aligned}$$

la densité associée à g_f est définie par

$$\forall u \in U \quad g_f(\{u\}) = g_f(f^{-1}(u))$$

il est facile de voir que le support de g_f ($=\{u/g_f(\{u\}) > 0\}$) est fini et c'est $f(\Omega)$.

Comme $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$, les mesures d'incertitude basées sur une t-norme se transportent aussi, avec conservation de la nature de g_f , de Ω à U .

Un cas particulier de (2) déjà connu est celui où g est une mesure de probabilité. Si $\Omega=U=R$, g est une mesure de probabilité sur R (muni de la σ -algèbre borelienne), construite à partir d'une variable aléatoire X , f est dite "fonction d'une variable aléatoire". Connaissant la fonction de répartition F_X associée à g , ($F_X(x) = \text{Prob}(X \leq x)$) le calcul de la fonction de répartition F_Y où $Y = f(X)$, définie par

$$F_Y(y) = \text{Prob} \{ f(X) \leq y \} \quad (3)$$

est abordé dans la plupart des ouvrages sur la théorie des probabilités.

L'extension de ces notions connues à des pseudo-probabilités (isomorphes des probabilités) telles qu'elles sont définies au chapitre 1 devrait pouvoir se faire sans problème. Le cas des pseudo-possibilités (voir ch. 1) est plus délicat à étendre à des ensembles infinis. (Si \star est une co-norme stricte g est isomorphe à une mesure non finie)

3.2 - CAS D'UNE MESURE DE POSSIBILITE

Si Π est une mesure de possibilité définie par une distribution $\pi = \mu_F$ où F est un ensemble flou normalisé de Ω , (1) correspond à ce que Zadeh /36/ a appelé "principe d'extension". On a

$$\begin{aligned}\Pi_f(A) &= \Pi(f^{-1}(A)) \\ &= \text{Sup} \{ \pi(\omega) / \omega \in \Omega, f(\omega) \in A \} \\ &= 0 \quad \text{si } f^{-1}(A) = \emptyset\end{aligned}\tag{4}$$

La quantité

$$N_f(A) = N(f^{-1}(A)) = 1 - \overline{\Pi(f^{-1}(A))}\tag{5}$$

est égale à $1 - \Pi_f(\bar{A})$ seulement si f est définie sur tout Ω (car alors $\overline{f^{-1}(\bar{A})} = f^{-1}(A)$), c'est-à-dire est une application. Dans ce cas Π_f est une mesure de possibilité sur \mathcal{U}

NB : plus généralement si une mesure d'incertitude g est transportée par une fonction telle que $f^{-1}(U) \not\subseteq \Omega$, alors g_f ne sera plus normalisée nécessairement puisque $g_f(U) < g(\Omega)$ est possible. On est alors tenu de distinguer entre $N_f(A)$ et $1 - \Pi_f(\bar{A})$: "Il est nécessaire que ω ait son image dans A " n'est plus équivalent à : "il est impossible que ω ait son image dans \bar{A} " (car ω peut n'avoir aucune image).

Quand f est une application, Π_f et N_f correspondent à une même distribution de possibilité π_f définie par (Zadeh /36/)

$$\begin{aligned}\pi_f(u) &= \Pi(f^{-1}(u)) = \text{sup} \{ \pi(\omega) / \omega : f(\omega)=u \} \\ &= 0 \quad \text{si } f^{-1}(u) = \emptyset\end{aligned}\tag{6}$$

Si $f^{-1}(U) \neq \Omega$ alors Π_f se construit à l'aide de π_f défini ci-dessus mais N_f se construit à l'aide de

$$\begin{aligned} \pi_f^*(u) &= \overline{\Pi(f^{-1}(U - \{u\}))} = \Pi(f^{-1}(u) \cup (\Omega - f^{-1}(U))) \\ &\geq \pi_f(u) \end{aligned}$$

On remarquera que π_f^* est toujours normalisé.

Interprétant π comme un ensemble flou F , on peut noter $\pi_f \stackrel{\Delta}{=} \mu_{F(F)}$ puisque si π est binaire, c'est-à-dire est la fonction caractéristique d'un ensemble $E \subseteq \Omega$, alors π_f est la fonction caractéristique de $f(E) = \{f(u)/u \in E\}$. Si F et G sont ensembles flous sur Ω alors, au sens de l'inclusion de Zadeh

$$f(F \cup G) \subseteq f(F) \cup f(G) \quad (7)$$

$$f(F \cap G) \subseteq f(F) \cap f(G) \quad (8)$$

où \cup et \cap sont définis par une co-norme et une t-norme continue.

Si l'union est au sens du maximum alors

$$f(F \cup G) = f(F) \cup f(G) \quad (9)$$

Si f est surjective, et pour toute complémentation C on a (Prade /21/) :

$$\overline{f(\bar{F})} \subseteq f(F) \quad (10)$$

on notera que

$$\begin{aligned} \mu_{f(\bar{F})}(u) &= \inf\{\pi(\omega)/\omega : f(\omega) = u\} \\ &= 1 \text{ si } f^{-1}(u) = \emptyset \end{aligned} \quad (11)$$

ce qui contraste avec (6)

Si f est injective, alors (6) s'écrit

$$\begin{aligned} \pi_f(u) &= \mu_{f(F)}(u) = \pi(f^{-1}(u)) \quad \text{si } f^{-1}(u) \neq \emptyset \\ &= 0 \text{ sinon} \end{aligned}$$

Dans ce cas pour tout opérateur d'union et d'intersection on a l'égalité

(1) on utilise la propriété $\sup\{\mu_F(\omega) * \mu_G(\omega) / \omega \in D\} \leq \sup\{\mu_F(\omega) * \mu_G(\omega') / (\omega, \omega') \in D^2\}$
 $= \sup\{\mu_F(\omega) / \omega \in D\} * \sup\{\mu_G(\omega) / \omega \in D\}$, $\forall D \subseteq \Omega$, $*$ continue non décroissante

dans (7) et (8). Si f est bijective alors (11) et (6) coïncident et (10) tient avec l'égalité.

Comme le transport d'une mesure de possibilité par une application généralise le concept d'image d'un ensemble par une application, on peut se demander si π_F peut se construire à partir des α -coupes de F tel que $\pi = \mu_F$. Cette question a été résolue par Nguyen /17/. Il a montré

$$\forall F \in [0, 1]^\Omega, \mu_{f(F)}(u) = \sup \{ \alpha / u \in f(F_\alpha) \} \quad (12)$$

$$f(F_\alpha) \subseteq [f(F)]_\alpha \quad \text{en général ;} \quad (13)$$

$$f(F_\alpha) = [f(F)]_\alpha \quad \text{avec } F_\alpha = \{u | \mu_F(u) > \alpha\} \quad (\alpha\text{-coupe stricte)}$$

condition de commutativité : $f(F_\alpha) = [f(F)]_\alpha$

si et seulement si $\forall u \in U$, le supremum dans (6) est atteint pour un certain ω (par exemple si f est injective).

3.3 - RELATIONS FLOUES ET PRODUIT CARTESIEN FLOU. NON-INTERACTION

On appelle relation floue un ensemble flou (ou une distribution de possibilité) sur un produit cartésien de référentiels. Le pendant probabiliste de la relation floue est la densité de probabilité jointe.

Soit Ω_1 et Ω_2 deux référentiels. La relation floue R a pour fonction d'appartenance $\mu_R = \pi$ dont les arguments sont $\omega_1 \in \Omega_1$ et $\omega_2 \in \Omega_2$

La projection de R sur Ω_1 ou encore distribution de possibilité marginale est π_1 telle que (Zadeh, 1975)

$$\pi_1(\omega_1) = \Pi(\{\omega_1\} \times \Omega_2) = \sup_{\omega_2} \pi(\omega_1, \omega_2) \quad (14)$$

C'est le pendant d'une probabilité marginale

Si F_1 est une donnée, sur Ω_1 , alors on peut étendre μ_{F_1} à $\Omega_1 \times \Omega_2$ par

$$\mu_{C_2(F_1)}(\omega_1, \omega_2) = \mu_{F_1}(\omega_1) \quad \forall \omega_1, \omega_2$$

$C_2(F_1)$ est alors l'extension cylindrique de F_1 (Zadeh /32/) à Ω_2 .

Soient F_1 et F_2 deux ensembles flous sur Ω_1 et Ω_2 respectivement. On peut étendre la notion de produit et de co-produit cartésien à F_1 et F_2 .

Dans le cas classique on a

$$\text{Produit cartésien} = A_1 \times A_2 = \{(\omega_1, \omega_2) / \omega_1 \in A_1, \omega_2 \in A_2\}$$

$$\text{Co-produit cartésien} = A_1 + A_2 = \{(\omega_1, \omega_2) / \omega_1 \in A_1\} \cup \{(\omega_1, \omega_2) / \omega_2 \in A_2\}$$

C'est-à-dire

$$A_1 \times A_2 = C_2(A_1) \cap C_1(A_2)$$

$$A_1 + A_2 = C_2(A_1) \cup C_1(A_2) (= \overline{A_1 \times A_2})$$

en termes d'extensions cylindriques. Il est clair que le produit (resp : co-produit) cartésien d'ensembles flous peut se définir comme l'intersection (resp : l'union) de leurs extensions cylindriques, soit (Dubois et Prade /8/) :

$$\mu_{F_1 \times F_2} = i(\mu_{F_1}, \mu_{F_2}) \quad \text{où } i \text{ est une norme triangulaire}$$

$$\mu_{F_1 + F_2} = u(\mu_{F_1}, \mu_{F_2}) \quad \text{où } u \text{ est une co-norme triangulaire.}$$

On vérifie les propriétés suivantes

. Le noyau (resp : support) de $F_1 \times F_2$ (resp : $F_1 + F_2$) est le produit (resp : co-produit) cartésien des noyaux (resp : supports).

$$\forall \alpha \in]0, 1], (F_1 \times F_2)_\alpha = F_{1\alpha} \times F_{2\alpha} \quad \leftrightarrow \quad i = \text{minimum}$$

(même chose en remplaçant \times par $+$, i par u , \min par \max)

. Si $\mathcal{P}_i(R)$ est la projection de R sur Ω_i on a toujours

$$R \subseteq \mathcal{P}_1(R) \times \mathcal{P}_2(R) \tag{15}$$

où \times est défini par le minimum, et \subseteq au sens de Zadeh

. Réciproquement si F_1 et F_2 sont deux ensembles flous normalisés sur Ω_1 et Ω_2 respectivement, alors la plus grande relation floue R telle que $F_i = \mathcal{P}_i(R)$ est $F_1 \times F_2$ où \times est défini par le minimum. Une relation floue R sera dite séparable si et seulement si

$$R = \mathcal{P}_1(R) \times \mathcal{P}_2(R) \quad (\text{au sens du minimum}) \tag{16}$$

Les variables x_1 et x_2 telles que le domaine de variation de (x_1, x_2) est

délimité par R seront alors dites non-interactives. Si F_i est l'ensemble flou des valeurs possibles (sur Ω_i) de x_i , x_1 et x_2 sont non interactives si la distribution de possibilité jointe de (x_1, x_2) est

$$\pi_{x_1, x_2} = \min(\mu_{F_1}, \mu_{F_2}) \quad (17)$$

On notera une certaine analogie entre indépendance stochastique et non-interaction possibiliste. Si P_1 et P_2 sont deux mesures de probabilité marginales, et P_{12} une mesure de probabilité sur $\Omega_1 \times \Omega_2$ alors, si pour $A_i \subseteq \Omega_i$, $i = 1, 2$

$$P_{12}(A_1, A_2) = \min(P_1(A_1), P_2(A_2)) \quad (18)$$

On en déduira que A_1 implique A_2 (si $P_1(A_1) \leq P_2(A_2)$), ou avec le point de vue fréquentiel A_1 ne s'est produit que lors d'épreuves où A_2 se produisait. Donc (18) est loin de signifier l'indépendance de A_1 et A_2 .

En revanche (17) définit la relation floue la plus grande (au sens de l'inclusion) ou de façon équivalente la fonction de possibilité la moins contrainte, qui ait pour projections F_1 et F_2 . Autrement dit si F_1 et F_2 sont les seules contraintes agissant sur (x_1, x_2) , (17) est la façon la plus licite de définir π_{x_1, x_2} . (17) n'est plus valide dès que x_1 et x_2 sont liées par quelque autre relation (par exemple si $\exists f : x_2 = f(x_1)$). On voit donc que (17) et (18) s'interprètent de façon radicalement différente. On sait que l'indépendance des événements A_1 et A_2 s'écrit

$$P_{12}(A_1, A_2) = P(A_1) \cdot P(A_2) \quad (19)$$

qui exprime, avec le point de vue fréquentiel, que A_1 se produit avec la même fréquence parmi les épreuves où A_2 s'est manifesté que parmi celles où A_2 n'a pas eu lieu (et inversement). Au contraire, (17) veut dire que le domaine de variation de x_1 est indépendant de celui de x_2 et inversement.

On peut affaiblir (17) comme suit. Une relation floue R normalisée est dite quasi-séparable (Dubois et Prade, /8/) si et seulement si il existe une norme triangulaire \perp telle que

$$\mu_R = \mu_{\mathcal{P}_1(R)} \perp \mu_{\mathcal{P}_2(R)} \quad (20)$$

Autrement dit R est encore un produit cartésien, mais R_α n'en est plus forcément un, quand $\alpha \neq 1$. Si $\mu_R = \pi_1 \perp \pi_2$ où π_1 et π_2 sont normalisés alors il est facile de vérifier $\pi_1 = \mathcal{P}_1(R)$, $\pi_2 = \mathcal{P}_2(R)$ et,

$$R_\alpha = \bigcup_{\beta, \gamma \mid \beta \perp \gamma \geq \alpha} [\mathcal{P}_1(R)]_\beta \times [\mathcal{P}_2(R)]_\gamma \quad (21)$$

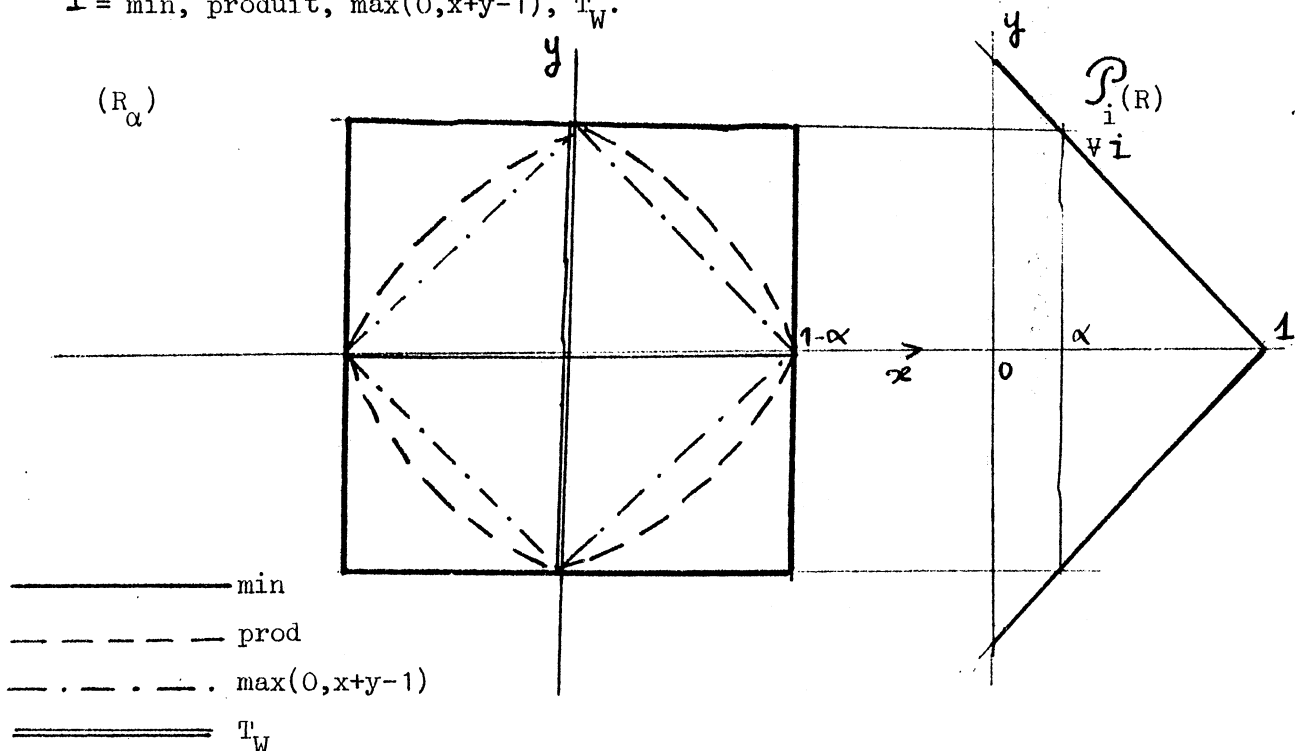
. Si $\perp =$ produit, (20) est semblable à (19) mais exprime que l'on pénalise le fait que x_1 et x_2 s'éloignent simultanément des noyaux $\mathcal{P}_1(R)$ et $\mathcal{P}_2(R)$ mais la propriété (valable pour toute norme stricte) $S(R) = S(\mathcal{P}_1(R)) \times S(\mathcal{P}_2(R))$ est conservée.

. Si $\perp = \max(0, x+y-1)$, alors on interdit à x_1 et x_2 de s'éloigner simultanément trop loin du noyau R . En particulier

$$S(R) \subsetneq S(\mathcal{P}_1(R)) \times S(\mathcal{P}_2(R)) \quad (22)$$

. Si $\perp = T_W$ seule l'une des deux variables est admise à sortir du noyau de la projection correspondante de R .

Exemple : Quatre situations sont décrites par le schéma ci-dessous, relatif à la relation floue sur \mathbb{R}^2 : $\mu_R(x,y) = \max(0, (1-|x|) \perp (1-|y|))$: $\perp = \min, \text{produit}, \max(0, x+y-1), T_W$.



On notera que pour des probabilités on a toujours

$$P_{12}(A_1, A_2) \geq \max(0, P(A_1) + P(A_2) - 1) \quad (23)$$

ce qui du point de vue fréquentiel veut dire qu'au cours d'épreuves simultanées sur $\Omega_1 \times \Omega_2$:

. A_1 et A_2 ne peuvent être vus comme événements exclusifs que s'il ne se sont jamais produits ensemble ($\rightarrow P(A_1) + P(A_2) \leq 1$)

. Si $P(A_1) + P(A_2) \geq 1$ A_1 et A_2 se sont produits simultanément au moins $[N(P(A_1) + P(A_2)) - N]$ fois sur N épreuves (et pas plus si (23) est une égalité).

Cette interprétation est cohérente avec (20), $\perp = \max(0, x+y-1)$ car on interdit alors à certains événements individuellement possibles de se produire simultanément. On voit aussi clairement que $P_{12}(A_1, A_2) = 0$ (événements mutuellement exclusifs) est équivalent à $P_{12}(A_1, A_2) = T_W(P(A_1), P(A_2)) (= 0$ sauf si l'un des deux événements se produit toujours auquel cas A_1 et A_2 ne sont certainement pas exclusifs - sauf trivialement).

3.4 - FONCTION DE POSSIBILITE DEFINIE SUR UN PRODUIT CARTESIEN

Soit Π_{12} une fonction de possibilité définie à partir d'une distribution π_{12} sur $\Omega_1 \times \Omega_2$. Notons que si π_{12} est séparable

$$\forall A_1 \times A_2 \subseteq \Omega_1 \times \Omega_2 \quad \Pi_{12}(A_1 \times A_2) = \min(\Pi_1(A_1), \Pi_2(A_2))$$

Si Π_{12} est quasi-séparable au sens de la t-norme \perp continue

$$\Pi_{12}(A_1 \times A_2) = \Pi_1(A_1) \perp \Pi_2(A_2) \quad (24)$$

où Π_1 et Π_2 sont les fonctions de possibilités marginales déduite de π_{12} . De plus $\forall \pi_{12}$ (séparable ou non)

$$\begin{aligned} \Pi_{12}(A_1 + A_2) &= \sup \{ \pi_{12}(\omega_1, \omega_2) / (\omega_1, \omega_2) \in (A_1 \times \Omega_1) \cup (A_2 \times \Omega_2) \} \\ &= \max(\Pi_1(A_1), \Pi_2(A_2)) \end{aligned} \quad (25)$$

d'où

$$N_{12}(A_1 \times A_2) = \min(N_1(A_1), N_2(A_2)) \quad (26)$$

où N_{12} est la fonction de nécessité bâtie sur π_{12} , N_i est la fonction de nécessité bâtie sur la possibilité marginale π_i de π_{12} , $i = 1, 2$.

Si F_1 et F_2 sont des événements flous, la possibilité de l'événement flou $F_1 \times F_2$ au sens de Zadeh /38/ s'exprime simplement si le produit cartésien $F_1 \times F_2$, l'intersection entre $F_1 \times F_2$ et la relation floue de fonction d'appartenance π_{12} s'expriment à l'aide de la même t-norme i telle que $\pi_{12} = i(\pi_1, \pi_2)$ on a alors

$$\Pi_{12}^{(i)}(F_1 \times F_2) = i(\Pi_1^{(i)}(F_1), \Pi_2^{(i)}(F_2)) \quad (27)$$

où (i) spécifie la t-norme utilisée pour $\Pi^{(i)}(F)$. Dans (27) on utilise la commutativité, l'associativité et la continuité de i .

Si $F_1 + F_2$ est le co-produit séparable ($+ = \text{maximum}$) alors puisque toute norme triangulaire i est distributive sur \max :

$$\Pi_{12}^{(i)}(F_1 + F_2) = \max(\Pi_1^{(i)}(F_1), \Pi_2^{(i)}(F_2)) \quad (28)$$

où i est continue.

Dualement on a

$$N_{12}^{(u)}(F_1 \times F_2) = \min(N_1^{(u)}(F_1), N_2^{(u)}(F_2)) \quad (29)$$

si x est le produit cartésien séparable ($x = \min$) et $u = 1 - i(1 - \cdot, 1 - \cdot)$ et en remarquant que $N_{12}^{(u)}(F_1 \times F_2) = 1 - \prod_{12}^{(i)}(\bar{F}_1 + \bar{F}_2)$.

Si $+$ est le co-produit cartésien au sens de u , et avec l'hypothèse de continuité on a encore la duale de (27) :

$$N_{12}^{(u)}(F_1 \times F_2) = u(N_1^{(u)}(F_1), N_2^{(u)}(F_2)) \quad (30)$$

(27)-(30) généralisent des résultats décrits par Prade (/22/) quand $i = \min$, $u = \max$.

3.5 - TRANSPORT D'UNE FONCTION DE POSSIBILITE DEFINIE SUR UN PRODUIT CARTESIEN

Soit f une fonction de deux variables : $\Omega_1 \times \Omega_2 \rightarrow U$. Si les deux arguments de f sont définis par une relation floue ou possibilité jointe π_{12} , celle-ci engendre une fonction de possibilité sur U telle que

$$\forall A \subseteq U, \Pi_f(A) = \Pi_{12}(f^{-1}(A)) \quad (31)$$

où $f^{-1}(A)$ est un sous-ensemble de $\Omega_1 \times \Omega_2$.

On obtient l'équivalent possibiliste des fonctions de 2 variables aléatoires.

Si P_{12} est la mesure de probabilité jointe, P_f est définie par

$$\forall A \subseteq U, P_f(A) = P_{12}(f^{-1}(A))$$

Π_f correspond à la distribution π_f telle que

$$\begin{aligned} \pi_f(u) &= \sup \{ \pi_{12}(\omega_1, \omega_2) / f(\omega_1, \omega_2) = u \} \\ &= 0 \text{ si } f^{-1}(u) = \emptyset \end{aligned} \quad (32)$$

Si π_{12} est séparable on trouve la forme la plus répandue du principe d'extension : Zadeh /36/

$$\begin{aligned} \pi_f(u) &= \sup \{ \min(\pi_1(\omega_1), \pi_2(\omega_2)) / f(\omega_1, \omega_2) = u \} \\ &= 0 \text{ si } f^{-1}(u) = \emptyset \end{aligned} \quad (33)$$

Si π_i est la distribution de possibilité de la variable x_i équivalente à un ensemble flou F_i , $i = 1, 2$, (33) donne la distribution de possibilité définissant le domaine de variation $f(F_1, F_2)$ de $z = f(x_1, x_2)$, quand x_1 et x_2 sont non-interactives.

Si x_1 et x_2 sont interactives, on peut avoir l'une des situations suivantes :

. π_{12} est quasi-séparable, c'est-à-dire qu'on contrôle l'évolution simultanée de x_1 et x_2 hors de leurs noyaux; on substitue à "min" une t-norme dans (33) (Dubois et Prade, /8/).

. Il existe une relation entre x_1 et x_2 qui restreint le domaine de (x_1, x_2) à un sous-ensemble D de $\Omega_1 \times \Omega_2$. Alors (31) devient

$$\Pi_f(A) = \Pi_{12}(f^{-1}(A) \cap D) \quad (34)$$

. Il existe une relation floue comme entre x_1 et x_2 , soit R , alors Π_f est induite par la distribution jointe

$$\pi_{12} = i(\mu_R, \min(\mu_{F_1}, \mu_{F_2})) \quad (35)$$

pour une intersection i .

Des propriétés analogues à (7)-(13) existent. On a notamment la condition de commutativité avec des α -coupes :

$$f(F_\alpha, G_\alpha) = f(F, G)_\alpha \leftrightarrow \text{le supremum est atteint dans (33), } \forall u$$

et encore

$$f(F_\alpha^-, G_\alpha^-) = (f(F, G))_\alpha^-, \quad \forall \alpha \in [0, 1] \quad (\alpha\text{-coupes strictes})$$

On peut bien sûr considérer l'expression duale de (33) en remplaçant "min" par "max" et "sup" par inf, qui représente $f(\bar{F}, \bar{G}) \leq f(F, G)$ dès que $f^{-1}(U) = \Omega_1 \times \Omega_2$

3.6 - EXEMPLES D'UTILISATION DU PRINCIPE D'EXTENSION

Le principe d'extension, tel qu'il est défini pour des fonctions de possibilité en (4) ou (33), est très utile pour donner un prolongement canonique de concepts mathématiques déjà connus, quand on veut les appliquer à des quantités ou objets imprécis, décrits par des ensembles flous. On peut alors exprimer l'imprécision d'un objet mal défini sur une grandeur qui le caractérise. On donne ici deux exemples :

- la cardinalité floue d'un ensemble flou (et son extension directe à la probabilité d'un événement flou)
- la compatibilité d'un ensemble flou avec un autre.

En fait tout le chapitre suivant est une application du principe d'extension aux opérations sur les réels, et constitue à lui seul un autre exemple particulièrement fécond.

3.6.1 - Cardinalité floue d'un ensemble flou

Soit F un ensemble flou sur Ω , supposé fini. On peut voir F comme un ensemble flou \mathcal{F} d'ensembles classiques au sens usuel, qui sont ses représentants, au sens du chapitre 2, équation (9).
Soit: $\forall A \in 2^\Omega$,

$$\begin{aligned} \mu_{\mathcal{F}}(A) = r(A) &= \min_{\omega \in A} \mu_F(\omega) \quad \text{si } A \supseteq \overset{\circ}{F} \\ &= 0 \quad \text{si } A \not\supseteq \overset{\circ}{F} \end{aligned} \quad (36)$$

La cardinalité d'un ensemble classique est une fonction $2^\Omega \rightarrow \mathbb{N}$ telle que

$$|A| = \sum_{\omega \in \Omega} \mu_A(\omega) \quad (37)$$

qui a été étendue aux ensembles flous au chapitre 2 (§ 24.1.3) en un indice scalaire, simplement en conservant (37).

On définit à partir de \mathcal{F} et de l'application $|\cdot|$ un ensemble flou sur \mathbb{N} par (Dubois, /6/).

$$\forall n \in \mathbb{N} \quad \pi(n) = \sup \{ \mu_{\mathcal{F}}(A) \mid |A| = n, A \in 2^\Omega \} \quad (38)$$

on note $\|\mathcal{F}\|$ l'ensemble flou ainsi obtenu. On a les propriétés suivantes : (Dubois /6/, Dubois & Prade, /10/).

. Pour toute α -coupe F_α , $\pi(|F_\alpha|) \geq \alpha$; en particulier si F n'est pas normalisé : $F = \emptyset$ et $\pi(0) = 1$, c'est-à-dire $\|\mathcal{F}\|$ est toujours normalisé.

. $S(\|\mathcal{F}\|) = \{n \in \mathbb{N} : |F| \leq n \leq |S(F)|\}$; $\max \|\mathcal{F}\|_\alpha = |F_\alpha|$

. Si A est un ensemble classique : $\|A\| = |A|$

. La fonction d'appartenance de $\|\mathcal{F}\|$ est non croissante sur son support $\forall n, n' \in S(\|\mathcal{F}\|) \quad n \geq n' \rightarrow \pi(n) \leq \pi(n')$

. Si $F \subseteq G$, alors $\|\mathcal{F}\|$ est "à droite" de $\|\mathcal{G}\|$, c'est-à-dire $\forall \alpha$ $\max \|\mathcal{F}\|_\alpha \leq \max \|\mathcal{G}\|_\alpha$. Cette propriété généralise le fait que l'inclusion d'un ensemble dans un autre entraîne une inégalité sur leurs cardinaux. (mais $F \subseteq G \not\Rightarrow \|F\| \subseteq \|G\|$ ce qui n'aurait d'ailleurs aucun sens).

. Si on généralise l'addition d'entiers par le principe d'extension alors la cardinalité floue reste additive.

Si M et N sont des ensembles flous de \mathbb{N} , alors $M \oplus N$ est défini par (33) :

$$\mu_{M \oplus N}(p) = \max \{ \min(\mu_M(m), \mu_N(n)) \mid m+n=p \}$$

On vérifie que $\forall F, G \in [0, 1]^\Omega$, et $\cap = \min \cup = \max$:

$$\|\mathcal{F}\| \oplus \|\mathcal{G}\| = \|\mathcal{F} \cap \mathcal{G}\| \oplus \|\mathcal{F} \cup \mathcal{G}\| \quad (39)$$

Un résultat similaire, mais plus faible existe pour le produit cartésien.

. Le cardinal scalaire d'un ensemble flou $|F|$ (37) est l'espérance mathématique supérieure du cardinal flou $\|F\|$, au sens de la mesure de possibilité construite sur la distribution π (cf. chap.1). L'espérance mathématique inférieure de $\|F\|$ est le cardinal du noyau de F . Ce résultat indique que cardinalités scalaires et floue ne sont pas des concepts divergents mais qu'ils se complètent.

NB : Le concept de cardinal flou a été introduit par Zadeh (/40/) avec la définition

$$\forall n \in \mathbb{N}, \pi'(n) = \sup \{ \alpha / |F_\alpha| = n \} \quad (40)$$

au lieu de (38). L'inconvénient de cette définition réside dans le fait que π' n'est plus monotone sur le segment d'entiers $\llbracket |F|, |S(F)| \rrbracket$, et l'additivité (39) ne tient plus (Dubois /5/). Plusieurs chercheurs dont Blanchard (/2/), Klement (/14/) ont suggéré de substituer à (40) l'expression

$$\pi''(n) = \sup \{ \alpha / |F_\alpha| \leq n \}$$

On montre alors que π'' égale π sur $S(\|F\|)$ mais $\pi''(n) = 1 \forall n \leq |F|$. C'est-à-dire que π'' coïncide avec le concept de cardinal tel qu'on le définit en mathématiques pures, c'est-à-dire l'ensemble des entiers entre 0 et $|A|$. Néanmoins, d'un point de vue sémantique $\pi''(n)$ est le degré de possibilité pour que F contienne au plus n éléments, et non exactement n éléments, comme l'exprime $\pi(n)$ dans (38). (40) a la même sémantique que (38) mais la représentation de F par des ensembles classiques est incomplète.

L'intérêt du concept de cardinal flou est de pouvoir modéliser des quantificateurs en environnement imprécis (c'est-à-dire répondre à des questions telles que "combien y a-t-il d'objets tels que 'Predicat flou'").

3.6.2 - Probabilité floue d'un événement flou

Au chapitre 2 (§ 4.3.1) on a signalé l'existence d'une définition scalaire de la probabilité d'un événement flou. En fait on peut trouver naturel qu'un événement incomplètement défini ait une probabilité imprécise

Yager (/31/) a proposé un analogue de (40). Si F est un événement flou sur (Ω, \mathcal{A}, P) alors la probabilité de F est un ensemble flou $\tilde{P}(F)$ de $[0,1]$ tel que

$$\forall a \in [0,1] \quad \mu_{\tilde{P}(F)}^{\alpha}(a) = \sup \{ \alpha / P(F_{\alpha}) = a \}$$

où μ_F est mesurable, c'est-à-dire $F_{\alpha} \in \mathcal{A}$, σ -algèbre de Ω . Cette quantité présente les mêmes inconvénients que son homologue sur la cardinalité. On peut utiliser l'analogie de (38)

$$\forall a \in [0,1] \quad \mu_{\tilde{P}(F)}^{\alpha}(a) = \sup \{ \mu_{\mathcal{F}}(A) / P(A) = a, A \in \mathcal{A} \}$$

Avec cette nouvelle définition, on peut montrer, dans le cas fini, que les propriétés des cardinaux flous s'étendent aux probabilités d'événements flous (voir Dubois et Prade /10/)

. l'additivité $\tilde{P}(F \cap G) \oplus \tilde{P}(F \cup G) = \tilde{P}(F) \oplus \tilde{P}(G)$ pour l'union et l'intersection définis par max et min respectivement.

. la probabilité scalaire d'un événement flou telle qu'elle fut définie par Zadeh en 1968, est l'espérance mathématique supérieure au sens des fonctions de Dempster-Shafer, de $\tilde{P}(F)$.

L'espérance mathématique inférieure est $P(\overset{\circ}{F})$.

La notion de probabilité d'un événement flou semble donc avoir une définition canonique. Il serait intéressant d'obtenir les résultats ci-dessus pour des espaces probabilisés quelconques, et plus seulement finis.

3.6.3 - Compatibilité de deux ensembles flous

Une façon de comparer deux ensembles flous qui ne recourt pas à l'utilisation d'un indice scalaire de comparaison (voir chap.2) a été proposée par Zadeh (/39/). L'inconvénient d'un indice scalaire est d'être très partiel : on ne considère qu'un seul attribut. Une autre approche consiste à considérer l'un des ensembles flous comme un élément de référence imprécisément

défini. La compatibilité d'un ensemble flou avec cette référence est alors obtenue en appliquant le principe d'extension à l'une des fonctions d'appartenance, ce qui fournit un degré d'appartenance flou. Plus précisément soit F et G deux ensembles flous sur Ω . G est considéré comme la référence. La compatibilité de F par rapport à G est un ensemble flou $\tau(F/G)$ de $[0,1]$ défini par (cf. (6)):

$$\begin{aligned}\mu_{\tau(F/G)}(a) &= \sup \{ \mu_G(\omega) / \omega \in \Omega, a = \mu_F(\omega) \} \\ &= 0 \text{ si } \mu_F^{-1}(a) = \emptyset\end{aligned}\quad (41)$$

on peut noter $\tau(F/G) = \mu_F(G)$, c'est-à-dire qu'on obtient bien un degré d'appartenance flou. La quantité duale (cf. (11)) $\bar{\tau}(F/\bar{G})$ a été considérée par Sanchez (/27/):

$$\begin{aligned}\mu_{\bar{\tau}(F/\bar{G})}(a) &= \inf \{ \mu_G(\omega) / \omega \in \Omega, a = \mu_F(\omega) \} \\ &= 1 \text{ si } \mu_F^{-1}(a) = \emptyset\end{aligned}\quad (42)$$

D'un point de vue pratique (41) revient à placer dans le carré de côté unité les couples $(x,y) = (\mu_F(\omega), \mu_G(\omega))$ pour tout $\omega \in \Omega$ en supprimant les (x,y) tels que $\exists (x',y')$ avec $x = x'$ et $y' > y$. On considère alors le graphe obtenu comme une quantité floue.

Cette quantité s'avère très utile pour représenter le degré de vérité d'une proposition logique floue par rapport à une autre (Zadeh /39/). $\tau(F/G)$ peut être vu comme le degré de vérité de l'affirmation "X est F" quand on sait par ailleurs que "X est G".

Pour la comparaison de F et G on peut en fait considérer $\mu_F(G)$ et $\mu_G(F)$ qui fournissent des informations différentes. $\tau(F/G)$ est une quantité floue dont les propriétés seront signalées dans le chapitre suivant consacré à ces quantités.

3.7 - TRANSPORT D'UNE MESURE DE POSSIBILITE PAR UNE FONCTION MULTIVOQUE

3.7.1 - Cas général

Ce type de transformation a déjà été considéré au chapitre 1.

Si Γ est une fonction multivoque $\Omega \rightarrow 2^U$ elle induit sur U , à partir d'une fonction de possibilité Π sur Ω , une fonction de possibilité supérieure Π_{Γ}^* sur U qui est une fonction de possibilité, et une fonction de possibilité inférieure Π_{Γ}^* qui n'en est pas une. On peut vérifier que la distribution de possibilité associée à Π_{Γ}^* est π_{Γ} définie par une généralisation de (6)

$$\begin{aligned} \pi_{\Gamma}(u) &= \sup \{ \pi(\omega) / u \in \Gamma(\omega) \} \\ &= 0 \text{ si } u \notin \Gamma(\Omega) \end{aligned} \quad (43)$$

Si Γ est défini pour tout ω alors

$$N_{\Gamma}(A) = 1 - \Pi_{\Gamma}^*(\bar{A}) \quad \forall A \subseteq U \quad (44)$$

On peut noter π_{Γ} sous forme ensembliste : $\Gamma(F)$ où $\pi = \mu_F$; $\pi_{\Gamma} = \mu_{\Gamma(F)}$
 $\Gamma(F)$ est une généralisation de

$$\forall E \subseteq \Omega, \Gamma(E) = \bigcup_{\omega \in E} \Gamma(\omega) \quad (45)$$

On a comme dans le cas d'une application simple (cf. (7) et (8))

$$\Gamma(F \cup G) \subseteq \Gamma(F) \cup \Gamma(G) \quad (46)$$

$$\Gamma(F \cap G) \subseteq \Gamma(F) \cap \Gamma(G) \quad (47)$$

L'égalité dans (46) tient pour \cup défini par le maximum.

Si $\Gamma(\Omega) = U$ alors pour toute complémentation

$$\overline{\Gamma(\bar{F})} \subseteq \Gamma(F) \quad (48)$$

puisque $\Gamma(\bar{F})$ est défini dualement par rapport à (43) :

$$\begin{aligned} \mu_{\Gamma(\bar{F})}(u) &= \inf \{ \pi(\omega) / u \in \Gamma(\omega) \} \\ &= 1 \text{ si } u \notin \Gamma(\Omega) \end{aligned} \quad (49)$$

Si $\forall \omega \neq \omega', \Gamma(\omega) \cap \Gamma(\omega') = \emptyset$ (U est alors un raffinement de Ω),

alors $\pi_{\Gamma}(u) = \pi(\omega)$ si ω est unique tel que $u \in \Gamma(\omega)$

= 0 si ω n'existe pas

et l'égalité tient toujours dans (46) et (47).

Si $\Gamma(\Omega) = U$ et U raffine Ω alors (49) et (43) coïncident et (48) est une égalité.

Les trois résultats sur les α -coupes, notamment la condition de commutativité restent valides pour des applications multivoques (cf. Dubois et Prade, /9/) (voir 3.2, équ.(12),(13)).

3.7.2 - Inverse d'une application

Un exemple simple de transport par une fonction multivoque est celui où il existe une application simple, $\varphi : U \rightarrow \Omega$ et où on veut transporter un ensemble flou F de Ω vers U . On a alors

$$\begin{aligned} \Gamma &= \varphi^{-1} : \Omega \rightarrow 2^U \\ \Gamma^{-1*}(u) &= \{\varphi(u)\} ; \Gamma^{-1*}(A) = \varphi(A) \\ \Gamma_{*}^{-1}(u) &= \{\varphi(u)\} \quad \text{si } \varphi(u) = \varphi(u') \rightarrow u = u' \\ &= \emptyset \quad \text{sinon} \end{aligned}$$

U est un raffinement de Ω car $\omega \neq \omega' \rightarrow \varphi^{-1}(\omega) \cap \varphi^{-1}(\omega') = \emptyset$
on a alors $\Pi_{\Gamma}(u) = \pi(\varphi(u)) \forall u$,
et $N_{*\Gamma}(A) = 1 - \Pi_{\Gamma}^{*}(A)$ si φ est surjective.
Les propriétés suivantes sont toujours valides

$$\begin{aligned} \varphi^{-1}(F \cup G) &= \varphi^{-1}(F) \cup \varphi^{-1}(G) \\ \varphi^{-1}(F \cap G) &= \varphi^{-1}(F) \cap \varphi^{-1}(G) \\ \varphi^{-1}(\bar{F}) &= \overline{\varphi^{-1}(F)} \quad \text{si } \forall u, \varphi(u) \neq \emptyset \end{aligned}$$

$\varphi^{-1}(F)$ est appelé "ensemble mal connu" par Prade (1982).

Exemple : Seuil flou

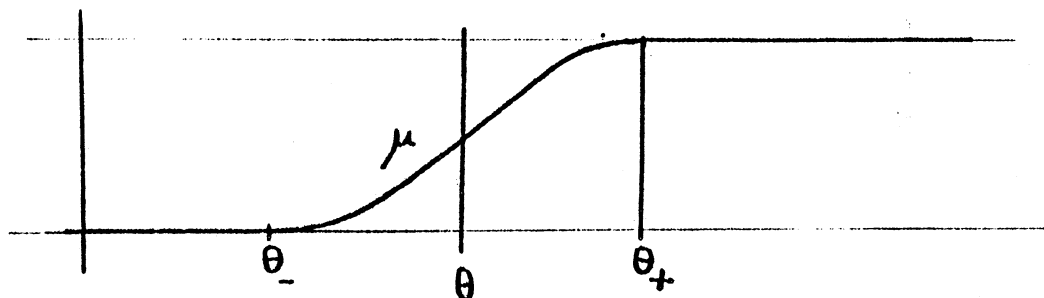
Soit U un ensemble fini et φ une fonction numérique $U \rightarrow \mathbb{R} = \Omega$ qui a chaque objet $u \in U$ en associe une évaluation $\varphi(u)$. Dans de nombreux problèmes, on rencontre la nécessité de sélectionner un sous-ensemble S d'objets dont l'évaluation dépasse un certain seuil θ . Mais il arrive souvent que le choix de ce seuil soit laissé à l'appréciation d'un utilisateur et que le résultat d'une procédure dépende beaucoup de la pertinence de ce choix. Dans ce cas l'utilisation d'un seuil précis peut avoir certains

inconvénients tels que : sélectivité trop grande ($S = \emptyset$) ou trop faible ($S = U$); forte sensibilité du résultat aux très faibles variations de θ (par exemple rejet abusif si u est tel que $\varphi(u) = \theta - \varepsilon$ où ε est très petit).

Une façon de prévenir ces inconvénients consiste à choisir un seuil flou sous la forme d'un ensemble flou de \mathbb{R} de fonction d'appartenance μ . On remarque que S est l'image inverse de l'intervalle $[\theta, +\infty)$ par f . On a donc la fonction caractéristique de S définie par

$$\mu_S(u) = \mu_{[\theta, +\infty)}[\varphi(u)] \quad , \quad \forall u \in U$$

Si on a un seuil flou, on se donne une fonction d'appartenance μ croissante sur \mathbb{R} pour le représenter mathématiquement.



Dans ce cas S est un ensemble flou défini par

$$\mu_S(u) = \mu \circ \varphi(u) \quad \forall u$$

L'information contenue dans μ est plus complète que celle contenue dans θ et permet d'éviter les inconvénients inhérents aux seuils précis.

Notons

$$\theta_+ = \inf \{ \theta / \mu(\theta) = 1 \}$$

$$\theta_- = \sup \{ \theta / \mu(\theta) = 0 \}$$

θ_+ est le seuil au-delà duquel tout objet u tel que $\varphi(u) \geq \theta_+$ doit être sélectionné. θ_- le seuil en deçà duquel tout objet u tel que $\varphi(u) \leq \theta_-$ doit être rejeté. Le tableau ci-dessous précise clairement l'intérêt d'un seuil flou :

Situation	Conséquence (seuil précis)	Conséquence (seuil flou)
$S = \emptyset$	baisser le seuil (de combien ?) et recommencer la sélection	Fin de la procédure. Il n'y a aucun objet pertinent dans U
$S = U$	augmenter le seuil (de combien ?) et recommencer la sélection	Fin de la procédure. Tous les objets de U sont pertinents
Modification très légère du seuil	$\theta \rightarrow \theta + \epsilon$ ou $\theta \rightarrow \theta - \epsilon$ $\mu_S(u)$ peut passer de 0 à 1 ou inversement $\forall \epsilon \dots$	$\theta^+ \rightarrow \theta^+ \pm \epsilon$: $\mu_S(u)$ a une faible variation pour certains u. (évolution lente).

En fait il est clair que le concept seuil flou revient à modifier la fonction d'évaluation en une nouvelle fonction, de façon à accentuer un contraste entre des objets pertinents et des objets moins pertinents. Ce genre de modification peut être faite sans référence aucune aux ensembles flous. Néanmoins, si on se replace dans le contexte du principe d'extension on peut interpréter toute modification d'une fonction d'évaluation en terme de sélection avec un seuil imprécis et exhiber la valeur de ce seuil. Dans certaines situations il paraît judicieux de prendre comme donnée la valeur floue du seuil et d'en déduire la nouvelle fonction d'évaluation, dans la mesure où ce seuil flou aura une interprétation facile à dégager dans la situation considérée.

On peut par exemple étendre la notion d' α -coupe d'un ensemble flou à des α -coupes floues (Zadeh /37/)... .

3.7.3 - Images inverses d'un ensemble flou par une fonction multivoque

On peut enfin transporter π de Ω sur U quand on ne connaît qu'une fonction multivoque $T : U \rightarrow \Omega$. On est alors amené à définir les inverses supérieurs et inférieurs d'ensembles flous. Ces inverses sont définis par leur fonction d'appartenance (Prade, 1982).

$$\begin{aligned} \mu_{T^{-1*}(F)}(u) &= \sup \{ \pi(\omega) / \omega \in T(u) \} \\ &= 0 : T(u) = \emptyset \end{aligned} \quad (50)$$

où $\pi = \mu_F$; si F est non flou, on retrouve $T^{-1*}(F) = \{u/T(u) \mid F \neq \emptyset\}$.

De même

$$\begin{aligned} \mu_{T_*^{-1}(F)}(u) &= \inf \{ \pi(\omega) / \omega \in T(u) \} \\ &= 0 \quad \text{sinon} \end{aligned} \quad (51)$$

Quand F est non flou on retrouve $T_*^{-1}(F) = \{u/\emptyset \neq T(u) \subseteq F\}$. Il est clair qu'au sens de l'inclusion de Zadeh

$$T_*^{-1}(F) \subseteq T^{-1*}(F) \quad (52)$$

$T^{-1*}(F)$ (resp : $T_*^{-1}(F)$) est l'ensemble flou des éléments de U dont l'image par T est possiblement (resp : nécessairement) dans F .

On vérifie

$$T^{-1*}(F \cup G) \subseteq T^{-1*}(F) \cup T^{-1*}(G) \quad (53)$$

$$T^{-1*}(F \cap G) \subseteq T^{-1*}(F) \cap T^{-1*}(G) \quad (54)$$

$$T_*^{-1}(F \cup G) \supseteq T_*^{-1}(F) \cup T_*^{-1}(G) \quad (55)$$

$$T_*^{-1}(F \cap G) \supseteq T_*^{-1}(F) \cap T_*^{-1}(G) \quad (56)$$

$$\text{Si } T^{-1}(\Omega) = \bigcup T_*^{-1}(F) = \overline{T^{-1*}(\bar{F})} \quad (57)$$

L'égalité dans (53) et (56) est valide quand \cap et \cup sont définis par min et max (Prade, /22/).

Prade (/22/) appelle "mal connus" les ensembles du type $T_*^{-1}(F)$ et $T^{-1*}(F)$ quand F est flou. En effet le fait que F est un ensemble incomplètement spécifié (mal défini) entraîne une mauvaise connaissance de son image par une application (qu'elle soit simple, multivoque, multivoque inverse, ...).

Quand Ω est un produit cartésien, et F une relation floue séparable, on peut étendre les résultats obtenus au 3.4 pour des applications multivoque, multivoque inverse (e.g. Prade (/22/)).

(1) cette inclusion ne tient que dans la mesure où dans la définition de l'inverse inférieur on exclut les éléments u tels que $T(u) = \emptyset$.

NB : notons que (50) et (51) correspondent à deux fonctions de possibilité

$$\begin{aligned}\Pi_{T^{-1}\star}(A) &= \sup \{ \pi(\omega) / \omega \in T(A) \} \\ &= \Pi(T(A))\end{aligned}$$

$$1 - \Pi'(T(A)) = 1 - \sup \{ 1 - \pi(\omega) / \omega \in T(A) \}$$

où Π' a pour distribution (non normalisée) $1 - \pi$

3.8 - TRANSPORT D'UNE FONCTION DE POSSIBILITE PAR UNE RELATION FLOUE

3.8.1 - Définitions

Soit R une relation floue sur $\Omega \times U$, que l'on peut voir comme une fonction de Ω dans les parties floues de U , soit $\tilde{\Gamma}$ telle que

$$\forall \omega \in \Omega, \mu_{\tilde{\Gamma}(\omega)} = \mu_R(\omega, \cdot) \quad (58)$$

L'image d'un ensemble $E \subseteq \Omega$ par $\tilde{\Gamma}$ est un ensemble flou $\tilde{\Gamma}(E)$ défini par

$$\mu_{\tilde{\Gamma}(E)}(u) = \sup_{\omega \in E} \mu_R(\omega, u) \quad (59)$$

qui est le degré de possibilité de l'événement E au sens de la distribution $\mu_R(\omega, \cdot)$. Si E est un ensemble flou F , on étend naturellement (59) en

$$\forall \omega \in \Omega \quad \mu_{\tilde{\Gamma}(F)}(u) = \sup_{\omega} i(\mu_R(\omega, u), \mu_F(\omega)) \quad (60)$$

où i est une norme triangulaire. (60) s'écrit aussi

$$\tilde{\Gamma}(F) = F \circ_i R \quad (61)$$

On exprime donc ici la composition sup- i d'un ensemble flou par une relation floue, qui généralise (11) et (36) ; on note $\circ = \circ_{\min}$.

On vérifie que :

. si l'union $F \cup F'$ est définie par le maximum alors

$$(F \cup F') \circ_i R = (F \circ_i R) \cup (F' \circ_i R)$$

. si l'intersection $F \cap F'$ est définie par le minimum alors

$$(F \cap F') \circ_i R \subseteq (F \circ_i R) \cap (F' \circ_i R)$$

où \subseteq exprime l'inégalité des fonction d'appartenance

$$F' \subseteq F \rightarrow F' \circ_i R \subseteq F \circ_i R$$

L'ensemble flou $\overline{\check{F}}(\bar{F})$ vaut au sens d'une complémentation C continue:

$$\begin{aligned} \forall \omega \in \Omega \quad \overline{\mu_{\check{F}}^{\check{F}}(\omega)}(u) &= C \left[\sup_{\omega} i(\mu_R(\omega, u), \text{Co}\mu_F(\omega)) \right] \\ &= \inf_{\omega} u(\text{Co}\mu_R(\omega, u), \mu_F(\omega)) \end{aligned}$$

où u est le C-dual de i.

Cette relation généralise (11). Si $\exists \omega, \mu_R(\omega, u) = 1$ et $u = 1 - i(1-, 1-)$, alors $\mu_{\check{F}}^{\check{F}}(\omega)(u) \leq \mu_{\check{F}}^{\check{F}}(F)(u)$, d'après 2.4.3.2. D'où si \check{F} est surjective ($\leftrightarrow \forall u, \exists \omega, \mu_R(\omega, u) = 1$)

$$\overline{\check{F}}(\bar{F}) \subseteq \check{F}(F)$$

3.8.2 - Intérêt de la composition des relations floues

La composition des relations floues s'est avérée un outil de base pour construire des modèles utilisant la théorie de Zadeh. On peut signaler

- la théorie des systèmes flous (Zadeh /32/, Negoita-Ralescu /16/) qui, d'abord construction abstraite, a pris tout son sens avec l'avènement de l'approche linguistique des processus complexes ou mal définis (Zadeh /34/) qui a eu des applications en automatique (les "contrôleurs" flous de Mandani /15/, voir bibliographie dans Dubois et Prade /7/ ou Willaëys /30/, ou en dynamique des systèmes (Wenstop /29/, Adamo et Karsky /12/).

- Les concepts de similitude, comme généralisation de la relation d'équivalence. La composition de relations floues permet de définir des notions affaiblies de transitivité d'une relation floue sur $\Omega \times \Omega$. (voir Zadeh /33/, Kaufmann /13/). Pour des applications de la relation de similitude à la classification automatique on se réfèrera au livre de Bezdek /1/

- Les concepts d'ordre et de préordre peuvent être nuancés à l'aide de la relation floue (là encore la transitivité floue est utile) voir Zadeh /33/, Orłowski /18/, Dubois et Prade /7/ pour une bibliographie, Roy /23/ pour une utilisation des relations d'ordre floues dans une procédure d'aide à la décision. (cf. Aussi le chapitre 6 § 3).

- La relation floue est commode pour représenter la notion d'inférence dans des modèles de raisonnement approché. La composition de relation floues est un moyen simple de représentation du passage de prémisses imprécises à des conclusions, ou plus généralement d'une chaîne d'inférences. Sur ce thème on se reportera aux travaux de Sanchez (/24/25/26//28/), au langage PRUF de Zadeh /39/ qui suggère un "modus ponens généralisé", et à la thèse de Prade /22/.

Sanchez /24/, à cette occasion, a défini et entrepris la résolution d'équations matricielles dites "équations de relations floues" qui généralisent les équations booléennes et prennent la forme :

trouver X tel que $R = XoS$

où R , X , S sont des relations floues sur des univers $\Omega \times \Omega''$, $\Omega \times \Omega'$, $\Omega' \times \Omega''$ respectivement, et o est une composition sup-min. Pedrycz /19/, a récemment étendu les résultats de Sanchez à la composition sup-norme triangulaire

Sanchez /26/ considère aussi une équation qui est similaire à celle qui définit les vecteurs propres d'une matrice en algèbre linéaire, à savoir :

trouver X tel que $X = XoR$

où X est un ensemble flou sur Ω et R une relation floue sur $\Omega \times \Omega$.

Dans ce mémoire on ne s'étendra pas plus longuement sur ce thème (voir néanmoins Prade, /22/) qui mériterait à lui seul une étude

détaillée. On se contentera de signaler le lien entre ces équations de relations floues et les "algèbres min-max" étudiées indépendamment par Cuninghame-Greene [4], [5] d'une part et la structure de dioïde systématisée par Gondran et Minoux ([1], Chapitre 3/). Il suffit de rappeler que $([0,1], \max, t\text{-norme})$ est un dioïde, qu'un dioïde est une structure générale suffisante pour faire de l'algèbre linéaire, et que les algèbres de Cuninghame-Greene sont de type $(\max,+)$. Une synthèse entre ces deux théories et celle initiée par Sanchez serait très utile pour progresser dans les divers domaines évoqués ci-dessus (systèmes flous, classification, analyse des préférences, modèles de raisonnement approché), en évitant de redéfricher un domaine déjà balisé par d'autres.

REFERENCES

- /1/ BEZDEK J. (1981)
Pattern Recognition with fuzzy objective function algorithms,
Plenum Press, N.Y.
- /2/ BLANCHARD N. (1982)
Theories cardinales et ordinales des ensembles flous.
Thèse 3ème cycle, Dept. de Mathematiques - Univ. de Lyon 1.
- /3/ CUNINGHAME-GREEN R.A. (1976)
Projections in minimax algebra. Math. Programming 10, 111-123
- /4/ CUNINGHAME-GREEN R.A. (1979)
*Minimax algebra Lecture notes in Economical and Mathematical
Systems, Springer Verlag.*
- /5/ DUBOIS D. (1981)
Propriétés de la cardinalité floue d'un ensemble flou fini.
BUSEFAL N°5, 11-12
- /6/ DUBOIS D. (1982)
A new definition of the cardinality of finite fuzzy sets.
BUSEFAL n°8, 65-67
- /7/ DUBOIS D., PRADE H. (1980)
Fuzzy sets and systems : theory and applications
Academic Press, N.Y.
- /8/ DUBOIS D., PRADE H. (1981)
Additions of interactive fuzzy numbers.
IEEE Trans. on Automatic Control, 26, 926-936.
- /9/ DUBOIS D., PRADE H. (1982)
*The use of fuzzy numbers in decision analysis, in "Fuzzy
Information and Decision Processes"*
(M.M. GUPTA, E. SANCHEZ eds). North Holland - 309-321.
- /10/ DUBOIS D., PRADE H. (1982)
Upper and lower possibilistic expectations and their applications.
4th Int. Seminar on Fuzzy Set Theory, LINZ, Autriche.

- /11/ GONDRAN M., MINOUX M. (1979)
Graphes et Algorithmes. Eyrolles.
- /12/ KARSKY M., ADAMO J.M. (1977)
Application de la dynamique des systèmes et de la logique floue à la modélisation d'un problème de postes en raffinerie.
Congrès AFCET Modélisation et Maitrise des Systèmes, Versailles.
- /13/ KAUFMANN A. (1973)
Introduction à la théorie des sous-ensembles flous.
Vol. 1, Masson, Paris.
- /14/ KLEMENT E.P. (1980)
On the cardinality of fuzzy sets, European Meeting on Cybernetics and Systems Research, Vienne, Autriche.
- /15/ MANDANI E. (1976)
Advances in the linguistic synthesis of fuzzy controllers.
Int. J. Man-Machine Studies 8, 669-678
- /16/ NEGOITA C.V., RALESCU D. (1975)
Applications of fuzzy set theory to systems analysis,
Birkhäuser Verlag
- /17/ NGUYEN H. (1978)
A note on the extension principle for fuzzy sets.
J. Math. Anal. Appl. 64, 369-380.
- /18/ ORLOWSKI S. (1978)
Decision-making with a fuzzy preference relation.
Fuzzy Sets & Syst. 1, 155-168.
- /19/ PEDRYCZ W. (1982)
Fuzzy relational equations with triangular norms and their resolutions. - BUSEFAL n°11.
- /20/ PRADE H. (1980)
Possibilité - nécessité - principe d'extension.
BUSEFAL n°4 - 60-63.
- /21/ PRADE H. (1980)
Compatibilité Qualification Niveau de précision.
BUSEFAL n°4, 71-78

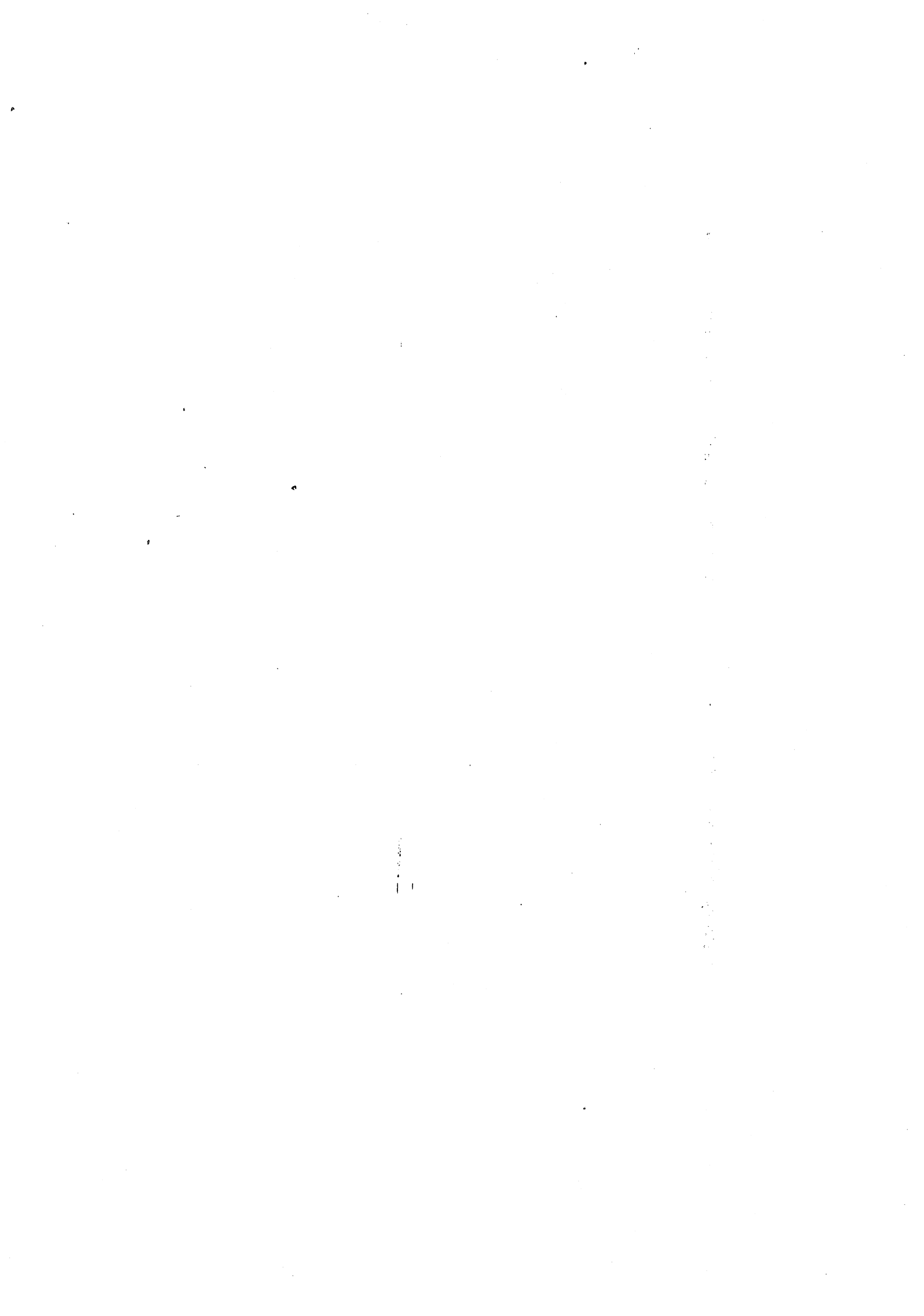
- /22/ PRADE H. (1982)
Modèles mathématiques de l'imprécis et de l'incertain, en vue d'applications au raisonnement approché. - Thèse d'Etat Univ. P. Sabatier TOULOUSE.
- /23/ ROY B. (1981)
 ELECTRE III. Un algorithme de classements fondé sur une représentation floue des préférences, en présence de critères multiples. - Cahiers du C.E.R.O. 20, 3-24.
- /24/ SANCHEZ E. (1976)
 Resolution of composite fuzzy relation equations. Information and Control, Vol. 30, pp 38-48.
- /25/ SANCHEZ E. (1977)
 Solutions in composite fuzzy relation equations : Application to medical diagnosis in Brouwerian Logic. dans "Fuzzy Automata and Decision Processes" (M.M. Gupta, G.N. Saridis, B.R. Gaines, eds) North-Holland, pp 221-234.
- /26/ SANCHEZ E. (1978)
 Resolution of eigen fuzzy sets equations. Fuzzy Sets and Systems, vol.1, pp 69-74.
- /27/ SANCHEZ E. (1980)
 Mesures de possibilité, qualifications de vérités et classification de formes linguistiques en médecine. Table Ronde CNRS "Quelques Applications Concrètes Utilisant les Derniers Perfectionnements de la Théorie du Flou, Lyon, 23-25 Juin 1980, (11 p.).
- /28/ SANCHEZ E. (1981)
 Eigen fuzzy sets and fuzzy relations. J. Math. Anal. and Appl. Vol.81 n°2, pp 399-421.
- /29/ WENSTOP F. (1976)
 Fuzzy set simulation models in a systems dynamics perspective, Kybernetes 6, 209-218.
- /30/ WILLAEYS D. (1980)
Contribution à la théorie des sous-ensembles flous en vue de son application à l'automatique. Thèse d'Etat, Univ. de Valenciennes.

- /31/ YAGER R. (1979)
A note on probabilities of fuzzy events,
Information Sciences 18, 113-129.
- /32/ ZADEH L.A. (1971)
*Toward a theory of fuzzy systems. In "Aspects of network and
System theory (R.G. Kalman, N. de Clavis, eds) 469-490.*
- /33/ ZADEH L.A. (1971)
Similarity measures and fuzzy orderings. Inf. Sci. 3, 177-200.
- /34/ ZADEH L.A. (1973)
*Outline of a new approach to the analysis of complex systems
and decision processes.*
IEEE Trans. Systems Man and Cybernetics, 3, 28-44.
- /35/ ZADEH, L.A. (1975a)
Calculus of fuzzy restrictions. Dans :
Fuzzy Sets and Their Applications to Cognitive and Decision
Processes (L.A. Zadeh, K.S. Fu, K. Tanaka, M. Shimura, eds) -
pp.1-39, Academic Press.
- /36/ ZADEH, L.A. (1975b)
*The concept of a linguistic variable and its application to
approximate reasoning.*
Information Sciences. Part.1, Vol.8, pp 199-249 ; Part 2,
Vol.8, pp 301-357 ; Part 3, Vol.9, pp 43-80.
- /37/ ZADEH L.A. (1976)
Fuzzy sets and their applications to pattern classification
and cluster analysis. Memo UCB/ERL M607, University of Cali-
fornia, Berkeley.
- /38/ ZADEH L.A. (1978a)
Fuzzy sets as a basis for a theory of possibility.
Fuzzy Sets and Systems - 1, 3-27.
- /39/ ZADEH L.A. (1978b)
PRUF - A meaning representation language for natural languages.
Int. J. Man-Machine Studies Vol.10, n°4, pp 395-460.
- /40/ ZADEH L.A. (1979)
A theory of approximate reasoning,
Machine Intelligence, Vol.9, Elsevier, 149-194.



CHAPITRE 4

LE CALCUL DES INTERVALLES FLOUS



Ce chapitre expose ce qu'on peut appeler "arithmétique floue"; il propose une approche canonique - dans le cadre de la théorie des possibilités - pour la représentation mathématique, la manipulation algébrique, et la comparaison de quantités imprécises.

Cette théorie est en total accord avec ce qu'on appelle couramment le calcul d'incertitudes ou calcul d'erreurs dont elle est une extension au cas d'intervalles valués. Une grosse part des travaux effectués par l'auteur concerne cette théorie. Son intérêt sera illustré dans la troisième partie de ce mémoire sur des exemples en théorie de la décision et en recherche opérationnelle. Essentiellement elle propose un raffinement de l'analyse de sensibilité qui peut ainsi devenir nuancée, et ce sans augmentation significative de la masse de calculs nécessaires. Le calcul des nombres flous peut prendre le relai de l'analyse probabiliste quand celle-ci est trop difficile à mener à bien, bien sûr au prix d'une certaine perte d'information, plus ou moins grande selon les types de problèmes abordés.

4.1 - QUANTITES, INTERVALLES, et NOMBRES FLOUS

4.1.1 - Définitions

Une quantité floue Q est un ensemble flou sur les réels, c'est-à-dire une application μ_Q de \mathbb{R} dans $[0,1]$. μ_Q sera naturellement vue comme une distribution de possibilité sur les valeurs que peut prendre une variable x , ce qu'on note

$$\pi_x \stackrel{\Delta}{=} \mu_Q$$

Conformément à la discussion relative à la normalisation des ensembles flous et à la nature "possibiliste", par définition, de la quantité floue Q on supposera sauf précision contraire que μ_Q est normalisée, c'est-à-dire

$$\exists m, \quad \mu_Q(m) = 1 \tag{1}$$

de façon à ce que la mesure de possibilité Π associée soit normalisée :

$$\Pi(\mathbb{R}) = \Pi(S(Q)) = 1$$

(c'est-à-dire x est un nombre réel) - m est une valeur modale de Q .

NB : on peut, dans certains cas ne pas admettre une représentation normalisée de Q . On interprète alors $\sup \mu_Q = \Pi(\mathbb{R})$ comme un degré de fiabilité globale de l'information exprimée par Q (cf. Dubois et Prade, /14/).

On peut définir un type de quantité floue qui généralise la notion d'intervalle :

Un intervalle flou est une quantité floue convexe c'est-à-dire dont la fonction d'appartenance est quasi-concave : (Zadeh /49/)

$$\forall u, v', \forall w \in [u, v], \mu_Q(w) \geq \min(\mu_Q(u), \mu_Q(v)) \quad (2)$$

(NB : cette définition de la convexité d'un ensemble flou s'étend sur \mathbb{R}^n).

Une quantité floue est convexe si et seulement si ses α -coupes sont convexes, c'est-à-dire sont des intervalles (bornés ou non).

Les intervalles fermés sont généralisés par des intervalles flous dont la fonction d'appartenance est semi-continue supérieurement (s.c.s) c'est-à-dire, par définition, dont les α -coupes sont des intervalles fermés.

Les sous-ensembles compacts de \mathbb{R} (fermés et bornés) sont généralisés par les quantités floues s.c.s à support compact.

On appellera nombre flou un intervalle flou s.c.s. à support compact et de valeur modale unique. Si M est un nombre flou de valeur modale m , M est une représentation possible de "environ m " (Dubois et Prade, 1978). Dans le cas d'un intervalle flou, l'ensemble des valeurs modales est un intervalle.

NB : Un autre concept de nombre flou a été proposé par Höhle /28/ sous la forme d'une fonction de répartition probabiliste. L'idée est que \mathbb{R} est isomorphe à un sous-ensemble des fonctions de répartition probabilistes par la transformation

$$r \mapsto f \quad \text{telle que} \quad \begin{aligned} f(u) &= 0 & \text{si } u < r \\ &= 1 & \text{si } u \geq r \end{aligned}$$

Dans l'optique de cet isomorphisme, une fonction de répartition est un

modèle naturel de nombre flou (on introduit une transition graduelle) qui généralise bien l'idée de nombre. En revanche les intervalles flous que l'on considère ici ne généralisent pas l'idée de nombre (avec les structures de groupe sous-jacentes). Ce qu'on appelle nombre flou est encore la généralisation d'un intervalle, est un nombre imprécis plutôt qu'un nombre mal localisé ou fluctuant.

Une quantité floue Q sera dite multimodale s'il existe un ensemble fini d'ensembles flous convexes $\{M_i / i \in I\}$ tels que Q est l'union des M_i au sens de max :

$$\mu_Q = \max_{i \in I} \mu_{M_i} \quad (3)$$

4.1.2 - Interprétations

Un intervalle flou est une représentation commode de quantités imprécises, plus riche en informations qu'un intervalle précis. En effet on rencontre souvent des situations où l'on doit estimer la précision d'un paramètre, ou fournir un souhait quant à la valeur d'un attribut, mais où l'intervalle est une représentation peu satisfaisante. En effet doit-on fixer les bornes de cet intervalle en étant pessimiste (l'intervalle est alors large et les calculs qu'on effectuera en suite ne signifieront plus rien, noyés qu'ils seront dans l'imprécision) ou en étant optimiste (il y aura alors un risque pour que la valeur ainsi cernée sorte du domaine assigné et cela rend illusoire les résultats plus précis obtenus) ? L'intervalle flou permet d'avoir une représentation à la fois pessimiste et optimiste : on choisira le support de façon à être sûr que la quantité cernée n'en sortira pas, et le noyau de l'intervalle flou contiendra les valeurs les plus plausibles. On rencontre ce type d'imprécision quand une grandeur numérique est fournie par un individu sous forme linguistique. L'intervalle flou est alors une représentation moins arbitraire que l'intervalle, et a fortiori, le nombre.

On notera qu'on peut, à l'aide de questionnaires simples, obtenir des renseignements qui aillent au-delà de la simple spécification du noyau et du support d'une quantité floue (cf. Dubois /9/, /11/, Kuz'min /29/), pour atteindre une information sur la forme de la fonction d'appartenance (transition soit linéaire, soit abrupte (ou progressive) aux limites du support (ou du noyau))).

On peut aussi donner, dans un autre contexte, une interprétation statistique à une quantité floue, en accord avec le chapitre 1 (§ 6) soit en transformant une mesure de probabilité sur \mathbb{R} en nombre flou (§ 6.1) soit en considérant la quantité floue comme résultant d'un processus de recueil de données imprécises (e.g. intervalles d'erreurs, § 6.2).

Dans les deux cas le degré de possibilité $\Pi(A)$ (resp : nécessité $N(A)$) d'un événement A , vu comme un intervalle de \mathbb{R} est une borne supérieure (resp : inférieure) de la valeur $P(A)$ de probabilité pour qu'une variable aléatoire prenne sa valeur dans A .

4.1.3 - Valeur moyenne d'un intervalle flou (Dubois et Prade /22/)

Dans cette optique un intervalle flou peut être vu comme la donnée d'une densité de probabilité p sur l'intervalle $[0,1]$, et d'une multi application $\Gamma = [0,1] \rightarrow \mathbb{R}$ telle que $\forall \alpha \in [0,1]$ $\Gamma(\alpha)$ soit un intervalle fermé et $\alpha_1 \leq \alpha_2 \rightarrow \Gamma(\alpha_2) \subseteq \Gamma(\alpha_1)$.

En notant $\Gamma(\alpha) = [\gamma_+(\alpha), \gamma_-(\alpha)]$, les images inverses d'un ensemble A de la forme $(-\infty, u]$ par Γ sont

$$A^* = \{\alpha \in [0,1] / \Gamma(\alpha) \cap A \neq \emptyset\} = \{\alpha / \gamma_-(\alpha) \leq u\} = \gamma_-^{-1}(A)$$

$$A_* = \{\alpha \in [0,1] / \Gamma(\alpha) \subseteq A\} = \{\alpha / \gamma_+(\alpha) \leq u\} = \gamma_+^{-1}(A)$$

Soit X une variable aléatoire réelle, telle que la probabilité $P([\alpha_1, \alpha_2]) = \int_{\alpha_1}^{\alpha_2} p(\alpha) d\alpha$ corresponde au nombre de fois que le résultat des expériences statistiques sur X a été un intervalle contenant $\Gamma(\alpha_1)$ et contenu dans $\Gamma(\alpha_2)$. Clairement, la probabilité $P(X \leq u)$, définissant la fonction de répartition de X n'est connue qu'imprécisément et on a

$$P(A_*) \leq P(X \leq u) \leq P(A^*) \quad (4)$$

donc $P(A_*)$ est la fonction de répartition inférieure de X , soit $F_*(u)$ et $P(A^*)$, la fonction de répartition supérieure $F^*(u)$. $P(A_*)$ et $P(A^*)$ sont des probabilités supérieures et inférieures, et (4) peut encore se lire

$$P_*((-\infty, u]) = F_*(u) \leq F(u) \leq P^*((-\infty, u]) = F^*(u) \quad (5)$$

On montre les résultats suivants (Dubois et Prade /22/)

$$\exists u_1, \forall u \geq u_1 \quad F^*(u) = 1 \quad (6)$$

$$\exists u_2 \geq u_1 : \forall u \leq u_2 \quad F_*(u) = 0 \quad (7)$$

$$\forall a, \forall b, \quad P^*([a,b]) = \sup_{u \in [a,b]} P^*(\{u\}) \quad (8)$$

$$P_*([a,b]) = \inf_{u \notin [a,b]} \{1 - P^*(\{u\})\} \quad (9)$$

. La quantité floue Q définie par $\mu_Q(u) = P^*(\{u\})$ est un intervalle flou. On a bien affaire à une distribution de possibilité de par (8) et (9).

On notera que

$$F^*(u) = \sup_{u' \leq u} \mu_Q(u') \quad (10)$$

$$F_*(u) = 1 - \sup_{u' > u} \mu_Q(u') \quad (11)$$

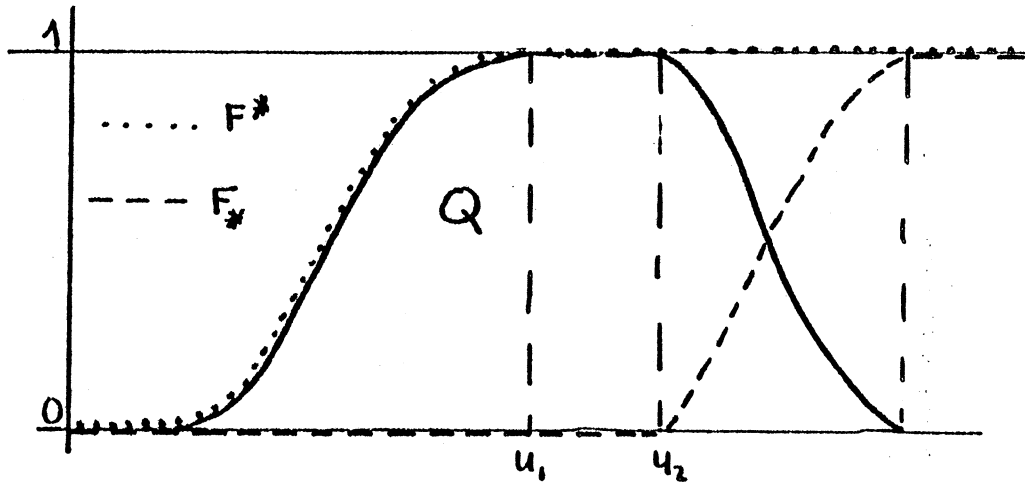


FIGURE 4.1

La fonction d'appartenance μ_Q s'exprime donc en fonction de p , γ_- et γ_+ .

NB : Un cas particulier de telle représentation est le cas où p = densité uniforme; $\Gamma(\alpha) = Q_\alpha$ (α -coupe) .

Un intervalle flou définit donc un ensemble de mesures de probabilité P dont les fonctions de répartition F_P sont bornées par F^* et F_* satisfaisant (6) et (7) ; F^* épouse la partie croissante de μ_Q , F_* est l'image inversée de la partie décroissante de μ_Q .

Dans cette optique on peut facilement donner un sens à la valeur moyenne d'un intervalle flou : ce sera un intervalle $[E_{\star}(Q), E^{\star}(Q)]$ tel que

$$E_{\star}(Q) = \inf \left\{ \int_{-\infty}^{+\infty} u dF_P(u) / F_{\star}^{\geq F_P \geq F_{\star}} \right\} \quad (12)$$

$$E^{\star}(Q) = \sup \left\{ \int_{-\infty}^{+\infty} u dF_P(u) / F_{\star}^{\geq F_P \geq F_{\star}} \right\} \quad (13)$$

La valeur moyenne d'un intervalle flou Q sera donc l'ensemble de toutes les valeurs moyennes des variables aléatoires compatibles (= vérifiant (5)) avec Q . On voit facilement que

$$E_{\star}(Q) = \int_{-\infty}^{+\infty} u dF_{\star}^{\star}(u) \quad (14)$$

$$E^{\star}(Q) = \int_{-\infty}^{+\infty} u dF_{\star}^{\star}(u) \quad (15)$$

Il paraît naturel que la valeur moyenne d'un intervalle flou soit un intervalle. Si $Q = [a, b]$ on vérifie bien $E_{\star}(Q) = a$, $E^{\star}(Q) = b$. Soit Q et Q' deux intervalles flous et $Q+Q'$ l'ensemble de toutes les mesures de probabilité obtenues en convolant une probabilité P compatible avec Q et une probabilité compatible avec Q' . Alors on vérifie

$$E_{\star}(Q) + E_{\star}(Q') \leq E_{\star}(Q+Q') \leq E_{\star}(Q) + E^{\star}(Q') \quad (16)$$

$$E_{\star}(Q') + E^{\star}(Q) \leq E^{\star}(Q+Q') \leq E^{\star}(Q) + E^{\star}(Q') \quad (17)$$

C'est-à-dire que l'additivité des valeurs moyennes ne tient plus dans ce cadre. En fait on retrouvera cette additivité si on additionne Q et Q' au sens des possibilités, c'est-à-dire par une convolution sup-min.

4.2 - ARITHMETIQUE FLOUE NON-INTERACTIVE PESSIMISTE

Dans cette section on se pose le problème suivant : étant donné plusieurs variables possibilistes non interactives $x, y, z \dots$ définie chacune par une quantité floue qui en limite le domaine, comment calcule-t-on la quantité floue qui borne le domaine d'une variable $f(x, y, z, \dots)$ où f est une fonction réelle donnée.

4.2.1 - Résultat fondamental

Soient M et N deux quantités floues qui limitent le domaine de variation de x et de y , et f une fonction de \mathbb{R}^2 dans \mathbb{R} . On définit l'ensemble flou $f(M,N)$ qui limite le domaine de x et y conformément au principe d'extension de la théorie des possibilités : (Zadeh /51/)

$$\begin{aligned} \forall w \in \mathbb{R}, \quad \mu_{f(M,N)}(w) &= \Pi(f^{-1}(w)) \\ &= \sup \{ \min(\mu_M(u), \mu_N(v)) : f(u,v) = w \} \\ &= 0 \text{ si } f^{-1}(w) = \emptyset \end{aligned} \quad (18)$$

On suppose ici que x et y sont non-interactives, d'où l'opérateur minimum dans l'expression du principe d'extension.

L'investigation de (18) a été initialisée indépendamment par plusieurs chercheurs dont l'auteur (cf. Nahmias /36-37/, Dubois et Prade /13/, Mizumoto et Tanaka /32-33/). On résume ci-après les principaux résultats obtenus.

Un premier résultat concerne la possibilité de représenter $f(M,N)$ à l'aide des α -coupes de M et N . Il est basique pour des questions d'interprétation et de calcul pratique.

Proposition 4.1 : Si μ_M et μ_N sont pseudo-concaves, normalisées, et s.c.s (i.e. sont des intervalles flous) ; si $\forall \alpha > 0$ M_α et $N_\alpha \neq \mathbb{R}$.

Si f est continue, définie sur \mathbb{R}^2 , $\sup f = +\infty$ $\inf f = -\infty$.

Si f est isotone, c'est-à-dire :

$$u \geq u' \quad \text{et} \quad v \geq v' \quad \rightarrow \quad f(u,v) \geq f(u',v')$$

Alors

$$f(M,N)_\alpha = f(M_\alpha, N_\alpha) \quad \forall \alpha > 0 \quad (19)$$

$\mu_{f(M,N)}$ est s.c.s.

$f(M,N)$ est un intervalle flou

Preuve : On sait déjà que $f(M_\alpha, N_\alpha) \subseteq f(M,N)_\alpha \quad \forall \alpha$

Il reste à montrer l'inclusion inverse. (19) est prouvé par Nguyen /38/

si M et N sont, de plus, à support compact, et par Dubois et Prade (/14/), si μ_M et μ_N sont continues. La proposition ci-dessus est plus générale puisque μ_M et μ_N peuvent être fonctions caractéristiques d'intervalles fermés bornés ou non. La démonstration, nouvelle, est donnée ci-dessous.

Supposons M_α et N_α bornés dans un premier temps.

1. $\forall \alpha > 0 \exists a, b, c, d$ (dépendant de α) : $M_\alpha = [a, b]$ $N_\alpha = [c, d]$
2. $f(M_\alpha, N_\alpha) = [f(a, c), f(b, d)]$
3. $\forall u < a \mu_M(u) < \alpha, \forall u > b \mu_M(u) < \alpha, \text{ etc...}$
4. $\forall u, u', v, v'$ si $f(u, v) > f(u', v')$ et $u' > u$ alors $v' \leq v$
5. si m et n sont valeurs modales de M et N alors $\mu(f(m, n)) = 1$ en notant μ la fonction d'appartenance de $f(M, N)$.
6. $\mu(f(a, c)) \geq \alpha \quad \mu(f(b, d)) \geq \alpha$ (de 2. et (19))
7. $\forall w < f(a, c), \mu(w) \leq \alpha$
 car : si $\mu(w) > \alpha$, soit u, v tel que $f(u, v) = w$
 et $\min(\mu_M(u), \mu_N(v)) > \alpha$ alors $u \in M_\alpha$ et $v \in N_\alpha$
 et donc $f(u, v) \geq f(a, c) > w$; contradiction.
8. $\forall w > f(b, d), \mu(w) \leq \alpha$
9. $\forall w < f(a, c), \mu(w) \neq \alpha$
 car : $f^{-1}(w) \subseteq \overline{[a, +\infty) \times [c, +\infty)}$. Donc si
 $D_a(w) = \{(u, v) : f(u, v) = w ; u \geq a\}$
 $D_c(w) = \{(u, v) : f(u, v) = w, v \geq c\}$
 $D_o(w) = \{(u, v) : f(u, v) = w, u < a, v < c\}$

alors

$$\sup\{v / (u, v) \in D_a(w)\} = c' < c \quad (1)$$

donc

$$\sup\{\min(\mu_M(u), \mu_N(v)) / (u, v) \in D_a(w)\} < \mu_N(c') < \alpha \quad (\text{de 3}).$$

(1) car : si $c' \geq c$ alors $\exists u \geq a, \forall \epsilon > 0, \exists v > c - \epsilon$ tel que $f(u, v) = w$. Mais $u \geq a \rightarrow \lim_{\epsilon \rightarrow 0} f(u, c - \epsilon) \geq f(a, c) > w$. Donc $\exists \epsilon, f(u, c - \epsilon) > w$. Par ailleurs :
 $f(u, c - \epsilon) \leq f(a, v) = w$. Contradiction.

La même chose est valable sur $D_c(w)$.

Donc

$$\mu(w) = \sup \{ \min(\mu_M(u), \mu_N(v)) \mid (u,v) \in D_0(w) \}$$

Si $\mu(w) = \alpha$ alors la proposition P suivante est fautive

$$P : \exists \varepsilon > 0, \forall (u,v) \in D_0(w), u < a - \varepsilon \text{ ou } v < c - \varepsilon$$

car alors : $\mu(w) \leq \max(\mu_M(a - \varepsilon), \mu_N(c - \varepsilon)) < \alpha$

donc

$$\neg P : \forall \varepsilon > 0, \exists (u,v) \in D_0(w), u \geq a - \varepsilon \text{ et } v \geq c - \varepsilon \text{ est vrai.}$$

alors

$$\forall \varepsilon > 0 \exists (u,v) \quad a - \varepsilon \leq u < a, \quad c - \varepsilon \leq v < c$$

et

$$f(u,v) = w.$$

Donc

$$\lim_{\varepsilon \rightarrow 0} f(a - \varepsilon, c - \varepsilon) = w < f(a, c)$$

donc f n'est pas continue. Contradiction.

10. $\forall w > f(b, d), \mu(w) \neq \alpha$

11. de 7.8.9.10. on déduit $w \notin f(M_\alpha, N_\alpha) \rightarrow w \notin f(M, N)_\alpha$
donc $f(M, N)_\alpha \subseteq f(M_\alpha, N_\alpha)$ d'où (19)

12. Si $M_\alpha = [a, +\infty)$, $N_\alpha = [c, +\infty)$, la "partie gauche" de la preuve est encore valide (idem dans l'autre sens); $f(M_\alpha, N_\alpha) = [f(a, c), +\infty)$.

13. Si $M_\alpha = [a, +\infty)$, $N_\alpha = (-\infty, c]$ alors $f(M_\alpha, N_\alpha) = \mathbb{R}$ et donc (19) tient encore

14. (19) et la continuité de f impliquent μ est s.c.s.

15. Puisque $f(N, M)_\alpha$ est toujours un intervalle, ou \mathbb{R} , $f(N, M)$ est convexe

Q.E.D.

REMARQUE : Si on considérait les α -coupes strictes, (19), écrite en ces termes, serait inconditionnellement valide pour toute quantité floue, $\forall f$. Néanmoins, on se ramènerait alors, dans le cas de fonctions d'appartenance s.c.s. à du calcul d'intervalles ouverts

La proposition 1 s'applique encore dans les cas suivants :

. f définie seulement sur un pavé de \mathbb{R}^2 (alors on considère la restriction de μ_M et μ_N à ce pavé.

. $\text{supf} < +\infty$ et/ou $\text{inff} < +\infty$. Remplacer $[f(a,c), +\infty)$ et \mathbb{R} par $[f(a,c), \text{supf}[$ et $]\text{inff}, \text{supf}[$ respectivement aux points 12 et 13.

. f est antitone : $u \geq u'$ et $v \geq v' \rightarrow f(u,v) \leq f(u',v')$

Alors, $f(M_\alpha, N_\alpha) = [f(b,d), f(a,c)]$

. f est hybride : $u \geq u'$ et $v \leq v' \rightarrow f(u,v) \geq f(u',v')$.

Alors $f(M_\alpha, N_\alpha) = [f(a,d), f(c,d)]$

4.2.2 - Calcul pratique dans le cas général

Une conséquence remarquable de la proposition 1 est que le calcul de $f(M,N)$ peut s'effectuer en combinant les parties croissantes de μ_M et μ_N , soit μ_M^+ et μ_N^+ d'une part et les parties décroissantes, soit μ_M^- et μ_N^- , séparément, dans le cas de fonctions f antitones ou isotones. Le tableau ci-dessous résume les 4 cas possibles, pour une fonction f à deux arguments

$f(x,y)$		$\mu_{f(M,N)}^+$ obtenue	
en x	en y	avec	et avec
croissant	croissant	μ_M^+	μ_N^+
croissant	décroissant	μ_M^+	μ_N^-
décroissant	croissant	μ_M^-	μ_N^+
décroissant	décroissant	μ_M^-	μ_N^-

l'équation (19) de commutativité avec les α -coupes permet d'écrire pour f isotone, et $\varepsilon \in \{-, +\}$:

$$\forall \alpha \in]0, 1] \quad \mu_M^\varepsilon(u) = \alpha \quad \mu_N^\varepsilon(v) = \alpha \rightarrow \mu_{f(M,N)}^\varepsilon(f(u,v)) = \alpha \quad (20)$$

Si $(\mu_M^\varepsilon)^{-1}$ est la fonction réciproque de μ_M^ε , alors on déduit de (20) (Dubois et Prade /14/) :

$$\forall \varepsilon \in \{-, +\}, (\mu_{f(M,N)}^\varepsilon)^{-1} = f((\mu_M^\varepsilon)^{-1}, (\mu_N^\varepsilon)^{-1}) \quad (21)$$

NB : Le terme "fonction réciproque" est légèrement abusif si $\exists u', u'', u' \neq u''$ $\alpha = \mu_M^+(u') = \mu_M^+(u'')$, c'est-à-dire μ_M est constant sur $[u', u'']$. Mais dans ce cas les intervalles où μ_M est constant deviennent des discontinuités pour $(\mu_M^+)^{-1}$ et inversement. De même si f n'est pas strictement isotone, $\exists (u'', v'')$,

(u'', v'') : $u' < u''$, $v' < v''$ et $f(u', v') = f(u'' v'') = w$. Alors $\mu_{f(M,N)}^\epsilon$ peut n'être pas continue même si μ_M^ϵ et μ_N^ϵ le sont, car $(\mu_{f(M,N)}^\epsilon)^{-1}$ est constant sur l'intervalle $[\min(\mu_M^\epsilon(u'), \mu_N^\epsilon(v')), \min(\mu_M^\epsilon(u''), \mu_N^\epsilon(v''))]$. Dans ce cas (20) doit se comprendre :

$$\mu_{f(M,N)}^\epsilon(w) = \sup \{ \alpha : w = f(u, v), \mu_M^\epsilon(u) = \mu_N^\epsilon(v) = \alpha \}$$

La relation (21) exprime clairement que l'application du principe d'extension à une classe importante de fonctions numériques mène à des calculs explicites simples. Si on veut obtenir $f(M, N)$ il suffit de résoudre l'équation en α :

$$w = f((\mu_M^\epsilon)^{-1}(\alpha), (\mu_N^\epsilon)^{-1}(\alpha)), \quad \epsilon \in \{+, -\} \quad (22)$$

Néanmoins même si (22) est difficile à résoudre, on peut se contenter

- soit des fonctions réciproques de $\mu_{f(M,N)}^+$ et $\mu_{f(M,N)}^-$ données par (21) sous forme analytique

- soit de discrétiser $[0, 1]$ et d'appliquer (20) pour obtenir un échantillonnage suffisant de la fonction d'appartenance de $f(M, N)$.

4.2.3 - Cas de quantités floues multimodales (Dubois et Prade /14/)

Deux résultats simples permettent de calculer $f(Q, Q')$ où Q et Q' sont des quantités floues multimodales et f satisfait les hypothèses de la proposition 1.

Proposition 4.2 : Si Q et Q' sont deux ensembles flous de \mathbb{R} tels que $\sup \mu_Q \leq \sup \mu_{Q'}$, alors si \underline{Q}' est la troncature de Q' au niveau $\sup \mu_Q$, c'est-à-dire $\mu_{\underline{Q}'} = \min(\sup \mu_Q, \mu_{Q'})$ alors

$$\forall f : f(Q, Q') = f(Q, \underline{Q}')$$

Proposition 4.3 : Si $Q = \bigcup_{i=1, m} M_i$, $Q' = \bigcup_{i=1, m} M_i'$ où M_i et M_i' sont des ensembles flous convexes de \mathbb{R} , à fonction d'appartenance s.c.s alors

$$f(Q, Q') = \bigcup_{i=1, m} \bigcup_{i'=1, m'} f(M_i, M_{i'})$$

Donc on peut toujours se ramener à un calcul sur des intervalles flous.

4.2.4 - Lien avec le calcul d'erreurs

Si deux quantités M et N sont imprécisément connues, sous la forme de sous-ensembles de \mathbb{R} , la quantité $f(M,N)$ est définie par

$$f(M,N) = \{f(u,v) / u \in M, v \in N\} \quad (23)$$

Les opérations engendrées sur les parties de \mathbb{R} par (23) quand f est une opération usuelle (addition, multiplication, ...) ont été étudiées dès 1931 par R.C. Young (/48/). Quand M et N sont des intervalles d'erreur fermés on retrouve le calcul d'erreur bien connu en physique. Le calcul d'erreurs a reçu un regain d'intérêt récemment, avec l'avènement des ordinateurs, pour évaluer les erreurs d'arrondi. Les livres de Moore (/34-35/) fournissent une synthèse sur le sujet.

Il est facile de voir que (23) est équivalent au principe d'extension (18) quand M et N sont des ensembles de réels. On a même le résultat plus fort suivant :

Proposition 4.4 : Le principe d'extension sup-min est la seule façon de définir $f(M,N)$ à partir de M et N , de façon à ce que $f(M,N)$ puisse être reconstitué par du calcul d'intervalles sur M_α, N_α au sens du théorème de représentation (cf. chap.2 section 1), c'est-à-dire

$$\forall w \quad \mu_{f(M,N)}(w) = \sup \{ \alpha / w \in f(M_\alpha, N_\alpha) \} \quad (24)$$

Preuve : Si (24) est vrai alors

$$\mu_{f(M,N)}(w) = \sup \{ \alpha / w = f(u,v), (u,v) \in M_\alpha \times N_\alpha \}$$

On se rappelle (chap.3) que l'ensemble flou d' α -coupes $M_\alpha \times N_\alpha$ a pour fonction d'appartenance $\min(\mu_M, \mu_N)$

donc

$$\mu_{f(M,N)}(w) = \sup \{ \alpha / w = f(u,v), \min(\mu_M(u), \mu_N(v)) \geq \alpha \}$$

d'où (18).

Q.E.D.

Le livre de Moore (/35/) servira de référence, dans la suite, pour montrer que toutes les propriétés structurelles du calcul d'intervalles fermés sont encore valides pour des intervalles flous s.c.s.

NB : On n'a pas évoqué la formule duale du principe d'extension, qui fournit $f(\bar{M}, \bar{N})$. En fait, dans le nombreux cas pratiques, par exemple si f est isotone, et si M et N sont des intervalles flous, $f(\bar{M}, \bar{N}) = |R$ et donc $f(\bar{M}, \bar{N}) = \emptyset$. Néanmoins cette quantité peutparfois être digne d'intérêt, quand elle n'est pas trivialement vide.

L'arithmétique floue définie par (18), en raison de sa coïncidence avec le calcul d'intervalles, peut être qualifiée de "pessimiste" puisque $f(M, N)$ est le plus grand ensemble flou (au sens de l'inclusion de Zadeh) tel que les arguments de f soient contraints par M et N respectivement. On le vérifie aisément en remarquant que chaque α -coupe de $f(M, N)$ est la plus grande possible conformément à (24).

On notera que si f est une fonction réelle de variables réelles, alors son extension à des ensembles flous satisfait la propriété de monotonie

$$M \subseteq M', \quad N \subseteq N' \quad \rightarrow \quad f(M, N) \subseteq f(M', N') \quad (25)$$

Cette propriété a été signalée par Moore (/35/), pour des intervalles, et rappelle que l'imprécision se conserve.

4.2.5 - Application aux opérations usuelles

L'extension d'opérations usuelles à des intervalles flous a été le sujet de travaux déjà publiés (Dubois et Prade /13/, /14/, /16/, Dubois /10-11/) et qui sont résumés ici.

4.2.5.1 - FONCTIONS D'UNE VARIABLE REELLE

Si f n'a qu'un seul argument et M est un intervalle flou, $f(M)$ a pour fonction d'appartenance

$$\begin{aligned} \mu_{f(M)}(w) &= \sup_{w=f(u)} \mu_M(u) \\ &= \mu_M(f^{-1}(u)) \quad \text{si } f \text{ est injective} \end{aligned} \quad (26)$$

Ainsi on peut définir les quantités suivantes, à partir de M :

f	f(M)	$\mu_{f(M)}(u)$
-u	-M (opposé)	$\mu_M(-u)$
1/u	1/M (inverse)	$\mu_M(1/u)$
u^p	M^p (puissance)	$\mu_M(u^{1/p})$
λu	λM (multiplication scalaire)	$\mu_M(u/\lambda) \quad \lambda \neq 0$
u	M valeur absolue	$ M = (M \cup -M) \cap \mathbb{R}^+$
e^u	e^M	$\mu_M(\log(u)); \quad u > 0$

Un intervalle flou M est dit positif (resp : négatif) dès que
 $\inf S(M) \geq 0$ (resp : $\sup S(M) \leq 0$)

Il est facile de voir que 1/M, où M est un intervalle flou, est lui-même un intervalle flou que si M est soit positif soit négatif. Les quantités définies dans les tableaux ci-dessus étendent les définitions de Moore /34/.

4.2.5.2 - EXTENSION DES QUATRE OPERATIONS

Notons d'abord que l'extension de toute fonction f commutative (resp : associative) est commutative (resp : associative).

L'addition M \oplus N de deux quantités floues est définie par

$$\mu_{M \oplus N}(w) = \sup \{ \min(\mu_M(u), \mu_N(w-u)) \mid u \in \mathbb{R} \} \quad (27)$$

C'est une version sup-min de la convolution. L'ensemble $\tilde{\mathcal{J}}(\mathbb{R})$ des intervalles flous s.c.s., muni de \oplus , forme un semi-groupe d'élément neutre 0. Sur les réels, \oplus coïncide avec l'addition usuelle. - M n'est généralement pas l'inverse de M pour \oplus car $(-M) \oplus M$ est un ensemble flou de valeur modale 0. $(\tilde{\mathcal{J}}(\mathbb{R}), \oplus)$ est un sous-semi-groupe de $([0, \tilde{\mathcal{J}}^{\mathbb{R}}, \oplus)$.

La soustraction M \ominus N de deux quantités floues est définie par

$$\mu_{M \ominus N}(w) = \sup \{ \min(\mu_M(w+u), \mu_N(u)) \mid u \in \mathbb{R} \} \quad (28)$$

On vérifie bien sûr $M \otimes N = M \otimes (-N)$

Le produit $M \otimes N$ de quantités floues est défini par

$$\mu_{M \otimes N}(w) = \sup \{ \min(\mu_M(u), \mu_N(w/u)) / u \in \mathbb{R} \} \quad (29)$$

L'ensemble $\tilde{\mathcal{I}}(\mathbb{R}^+)$ des intervalles flous s.c.s. positifs, muni de \otimes forme un semi-groupe d'élément neutre 1. Sur les réels, \otimes coïncide avec le produit usuel. $1/M$ n'est pas un inverse de groupe car $M \otimes 1/M$ est un sous-ensemble flou de valeur modale 1. 0 est élément absorbant pour \otimes . Enfin on vérifie

$$M \otimes N = (-M) \otimes (-N)$$

$$(-M) \otimes N = M \otimes (-N) = -(M \otimes N)$$

La distributivité de \otimes sur \otimes ne tient pas en général. Néanmoins on en a une version affaiblie :

$$M \otimes (P \otimes Q) \subset (M \otimes P) \otimes (M \otimes Q) \quad (30)$$

La distributivité tient dans les cas suivants, au moins.

- . M est un nombre réel
- . M, P, Q sont des intervalles flous s.c.s. P et Q sont ensemble soit négatifs, soit positifs
- . M, P, Q, sont des intervalles flous s.c.s ; P et Q sont des intervalles flous symétriques ($P = -P$, $Q = -Q$).

Ces résultats se démontrent en utilisant (19) et les résultats classiques sur les intervalles (e.g. Moore /35/).

Le quotient $M \oslash N$ de deux quantités floues est défini par

$$\mu_{M \oslash N}(w) = \sup \{ \min(\mu_M(u.w), \mu_N(u)) / u \in \mathbb{R} \} \quad (31)$$

On a bien sûr $M \oslash N = M \otimes 1/N$. Si M et N sont de signe donné (positif ou négatif) alors, si M et N sont des intervalles flous, $M \oslash N$ l'est aussi.

4.2.5.3 - CALCUL PRATIQUE EXACT DES 4 OPERATIONS (Dubois et Prade, /13/; Dubois, /11/)

Si on utilise une représentation paramétrée de la fonction d'appartenance des intervalles flous manipulés, le calcul arithmétique flou

peut devenir très simple.

Une bonne représentation paramétrée de $M \in \tilde{\mathcal{J}}(\mathbb{R})$ est obtenue à partir d'éléments d'une classe de fonction $(\mathbb{R}^+ \rightarrow [0,1])$, soit \mathcal{R} , telle que $\forall L \in \mathcal{R}$; L est décroissante au sens large, s.c.s., $L(0) = 1$; on suppose de plus $\forall u \in [0,1] L(u) \geq 0$ et pour $u > 1$

$$\text{soit } L(u) = 0$$

$$\text{soit } L(u) > 0$$

Alors pour tout intervalle flou s.c.s. M , $\exists L, R \in \mathcal{R}$, $\exists (\underline{m}, \bar{m}) \in \mathbb{R}^2$, $\exists \alpha, \beta \geq 0$ tel que

$$\begin{aligned} \mu_M(u) &= L\left(\frac{m-u}{\alpha}\right) \quad \forall u \leq \underline{m} \\ &= 1 \quad \forall u \in [\underline{m}, \bar{m}] \\ &= R\left(\frac{u-\bar{m}}{\beta}\right) \quad \forall u > \bar{m} \end{aligned} \quad (32)$$

$[\underline{m}, \bar{m}]$ est le noyau de M si $\max(L(u), R(u)) < 1 \quad \forall u > 0$.

$[\underline{m}-\alpha, \bar{m}+\beta]$ est le support de M si M est à support borné et $\min(L(u), R(u)) > 0, \quad \forall u < 1$
 α et β sont appelés étalements. On peut noter sans ambiguïté.

$$M = (\underline{m}, \bar{m}, \alpha, \beta)_{LR}$$

et on dit "M est de type L-R". Par exemple :

$$\bullet \quad M = [a, b] \quad \begin{aligned} L(u) = R(u) &= 1 \quad \forall u \in [0, 1] \\ &= 0 \quad \forall u > 1 \end{aligned}$$

$$\underline{m} = \bar{m} = \frac{a+b}{2} \quad \alpha = \beta = \frac{a+b}{2}$$

$$\bullet \quad M = m \in \mathbb{R} : L \text{ et } R \text{ quelconques, } \alpha = \beta = 0 \quad \underline{m} = \bar{m} = m$$

$$\bullet \quad M \text{ est triangulaire : } L(u) = R(u) = \max(0, 1-u)$$

On peut avoir $L(u) = \max(0, 1-u^p), \max(0, 1-u)^p, e^{-x}, e^{-x^2}, \dots$

Si $M = (m, \alpha, \beta)_{LR}$ et $N = (n, \gamma, \delta)_{LR}$ (pour simplifier on suppose $\underline{m} = \bar{m} = m$)

alors l'équation donnant $f(M, N)$ (c'est-à-dire (22)) devient pour f isotone :
trouver $\lambda \in]0, 1]$ tel que

$$m \geq w = f(m-\alpha L^{-1}(\lambda), n-\gamma L^{-1}(\lambda)) \quad (33)$$

$$m \leq w = f(m+\beta L^{-1}(\lambda), n+\delta L^{-1}(\lambda)) \quad (34)$$

on se ramène à deux équations en $L^{-1}(\lambda)$ dont la complexité dépend seulement de la complexité de l'équation en u :

$$w = f(a+bu, c+du)$$

On a en particulier

. pour $f(u,v) = u+v$ (isotone sur (\mathbb{R}^2)) des équations du 1er degré si M et N sont de type LR. D'où :

$$M \oplus N = (m+n, \alpha+\beta, \gamma+\delta)_{LR}$$

. pour $f(u,v) = u-v$ (antitone sur (\mathbb{R}^2)) des équations du 1er degré si N est de type RL :

$$M \ominus N = (m-n, \alpha+\delta, \gamma+\beta)_{LR}$$

On obtient donc un théorème d'invariance relativement fort pour la somme de nombre flous : si $\tilde{\mathcal{J}}_{LR}(\mathbb{R})$ est l'ensemble des intervalles flous bâtis sur L et R par affinités d'axe vertical Oy suivie de translations, selon (32), alors

$$\begin{aligned} \text{Proposition 4.5 : } M, N \in \tilde{\mathcal{J}}_{LR}(\mathbb{R}) &\rightarrow M \oplus N \in \tilde{\mathcal{J}}_{LR}(\mathbb{R}) \\ M \in \tilde{\mathcal{J}}_{LR}(\mathbb{R}), N \in \tilde{\mathcal{J}}_{RL}(\mathbb{R}) &\rightarrow M \ominus N \in \tilde{\mathcal{J}}_{LR}(\mathbb{R}) \end{aligned}$$

Cette propriété d'invariance est bien plus forte que celle que l'on obtient pour les sommes de variables aléatoires où seules quelques distributions, telles les gaussiennes, sont invariantes.

La proposition 4.5 ne tient bien sûr pas pour le produit et le quotient ; néanmoins (33) et (34) restent simples :

. pour $f(u,v) = u.v$ (isotone sur $(\mathbb{R}^+)^2$) $M > 0$, $N > 0$ des équations du second degré si M et N sont de type L-R.

Par exemple pour

$$w < mn, \mu_{M \otimes N}(w) = L \left(\frac{n\alpha + m\beta + \sqrt{(m\gamma - n\alpha)^2 + 4\alpha\gamma w}}{2\alpha\gamma} \right)$$

. pour $f(u,v) = u/v$ (antitone sur $(\mathbb{R}^+)^2$ $M > 0, N > 0$, M est de type LR et N de type RL alors (33) et (34) sont des équations du 1er degré.

Et par exemple si $w \leq m/n$

$$\mu_{M \ominus N}(w) = L\left(\frac{m-wn}{\alpha+w\delta}\right)$$

On obtient aussi des équations du 2^e degré pour le calcul de sommes du type

$$\sum_{i=1}^n \alpha_i M_i \ominus N_i$$

où M_i et N_i sont des intervalles flous positifs de type LR.

En revanche connaître la fonction d'appartenance de $\prod_{i=1}^p M_i$ où M_i est positif nécessite la résolution d'équations du $p^{\text{ième}}$ degré.

4.2.5.4 - CALCUL APPROCHE DES FONCTIONS D'INTERVALLES FLOUS

Quand on ne sait pas calculer analytiquement $f(M,N)$ on peut soit la calculer point par point (cf § 2.2) soit avoir recours à des formules approchées qui donnent une approximation de type L-R du résultat.

Si f et L sont différentiables alors (33) peut, au second ordre près, et au voisinage de $\lambda = 1$ s'écrire, en posant $\varepsilon = L^{-1}(\lambda) > 0$

$$w \# f(m,n) - f'_x(m,n)\varepsilon\alpha - f'_y(m,n)\varepsilon\gamma \quad (35)$$

ce qui est une équation du 1er degré.

On a donc, au voisinage du noyau ($\underline{m} = \bar{m} = m$), si M et N sont de type L-R et f est isotone, une approximation de type L-R pour $f(M,N)$ sous la forme

$$f_1^\#(M,N) = (f(m,n), f'_x(m,n)\alpha + f'_y(m,n)\gamma, f'_x(m,n)\beta + f'_y(m,n)\delta)_{LR} \quad (36)$$

(36) généralise des formules données par Dubois et Prade (/13/) pour le produit et le quotient, et souligne encore le lien entre le calcul d'intervalles flous et le calcul d'erreur.

Il est facile de voir que

. $\mu_{f_1^\#}(M,N)$ est tangente à $\mu_{f(M,N)}$ en m .

$$\mu_{f_1^\#(M,N)}(w) \geq \mu_{f(M,N)}(w) \quad \forall w \leq m \quad (37)$$

$$\mu_{f_1^\#(M,N)}(w) \leq \mu_{f(M,N)}(w) \quad \forall w \geq m \quad (38)$$

On peut ainsi construire des approximations de $f(M,N)$ en n'importe quel point correspondant à une coupe de niveau λ en écrivant le développement limité (35) au voisinage de $L^{-1}(\lambda)$. On a alors une fonction $f_\lambda^\#(M,N)$. Des expressions de $f_0^\#(M,N)$ au voisinage du support (s'il est borné) de $f(M,N)$ sont fournies dans Dubois (1981).

Une autre approximation consiste à choisir un intervalle flou de même support et de même noyau que $f(M,N)$, mais de type L-R et de poser quand M et N sont à support fini.

$$\tilde{f}(M,N) = (f(m,n), f(m,n) - f(m-\alpha, n-\gamma), f(m+\beta, n+\delta) - f(m,n))_{LR} \approx f(M,N) \quad (39)$$

Quand la fonction $\phi(\epsilon) = f(m-\alpha\epsilon, n-\gamma\epsilon)$ est concave, on peut montrer que $\tilde{f}(M,N)$, $f_1^\#(M,N)$ et $f_0^\#(M,N)$ définissent deux triangles curvilignes qui chacun contiennent la partie gauche ou la partie droite de $\mu_{f(M,N)}$. Voir Dubois (/11/) pour un exemple.

4.2.5.5 - EXTENSION DU MINIMUM ET DU MAXIMUM

(Dubois et Prade /13/)

Il est intéressant de savoir si le principe d'extension parvient encore à donner un sens au maximum ou au minimum de deux intervalles.

Les fonctions $g_1(s,t) = \max(s,t)$ et $g_2(s,t) = \min(s,t)$ sont isotones sur $\mathbb{R} \times \mathbb{R}$; ce qui permet de les étendre aisément puisque l'on a

$$\max([a, a'], [b, b']) = [\max(a, b), \max(a', b')]$$

$$\min([a, a'], [b, b']) = [\min(a, b), \min(a', b')]$$

d'où la construction évidente de $\tilde{\max}(M,N)$ et de $\tilde{\min}(N,M)$, $\tilde{\max}$ et $\tilde{\min}$ désignant les opérations étendues au sens de (18)

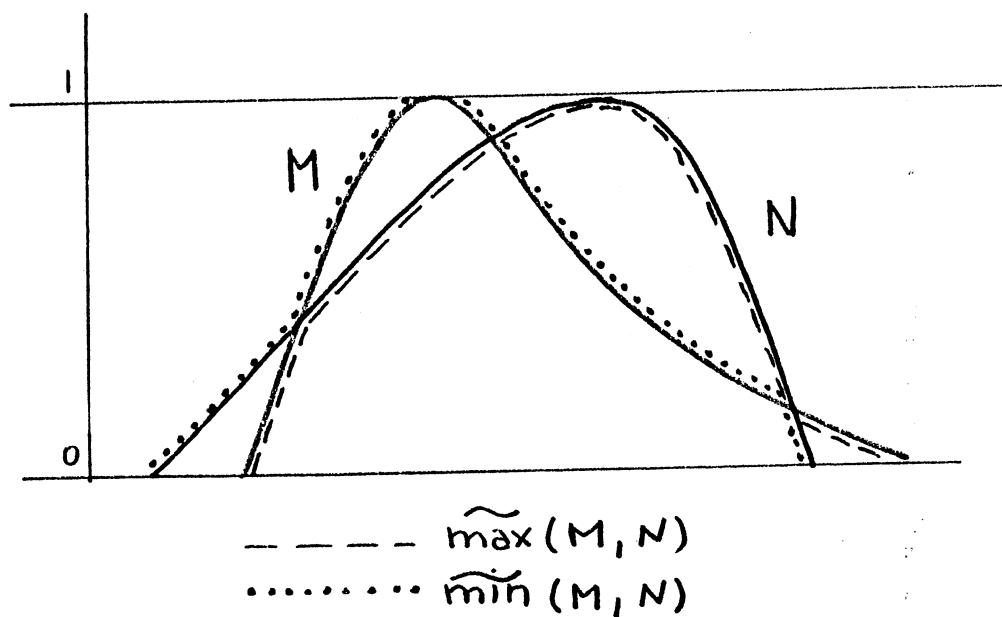


FIGURE 4.2

$\tilde{\max}$ et $\tilde{\min}$ sont commutatifs et associatifs ; on a

$$-\tilde{\max}(M, N) = \tilde{\min}((-M), (-N)) \quad (40)$$

Sur l'ensemble des intervalles flous, $\tilde{\mathcal{F}}(\mathbb{R})$, $\tilde{\max}$ et $\tilde{\min}$ sont mutuellement distributifs et vérifient les propriétés suivantes :

$$\tilde{\min}(M, N) \oplus \tilde{\max}(M, N) = M \oplus N \quad (41)$$

$$M \oplus \tilde{\min}(N, P) = \tilde{\min}(M \oplus N, M \oplus P) \quad (42)$$

$$M \oplus \tilde{\max}(N, P) = \tilde{\max}(M \oplus N, M \oplus P) \quad (43)$$

$$\tilde{\max}(M, N) = M \leftrightarrow \tilde{\min}(N, N) = N \quad (44)$$

$$\tilde{\max}(M, M) = M ; \tilde{\min}(M, M) = M \quad (45)$$

Notons que $\tilde{\max}(M, -\infty) = M$, sur $\mathcal{R}_U\{-\infty\}$. Donc $(\tilde{\mathcal{F}}(\mathcal{R}_U\{-\infty\}), \tilde{\max})$ est un semi-groupe commutatif d'élément neutre $-\infty$. $(\tilde{\mathcal{F}}(\mathcal{R}_U\{-\infty\}), \oplus)$ est aussi un semi-groupe commutatif d'élément neutre $-\infty$. (43) exprime la distributivité à gauche de \oplus sur $\tilde{\max}$; on aurait pu aussi bien l'écrire à droite.

On en conclut que $(\tilde{\mathcal{J}}(\mathbb{R}_U[-\infty]), \tilde{\max}, \emptyset)$ forme un dioïde (Gondran et Minoux, /27/). De plus on a encore l'idem-potence de $\tilde{\max}$. Les seules différences avec le dioïde $(\mathbb{R}_U[-\infty], \max, +)$ étudié par Cuninghame-Greene (/7/) sont le fait que $\tilde{\max}(M, N)$ peut n'être ni M, ni N (cf. figure 2 ; on a même : $\tilde{\max}(M, \tilde{\min}(M, N)) \neq M$ quand M et N ne sont pas convexes), et \emptyset n'est pas inversible.

Si on note

$$M^{(k)} = \tilde{\max}(0, M, \underbrace{M \emptyset M \emptyset M \dots M \emptyset M \emptyset M \dots \emptyset M}_{k \text{ fois}})$$

On rappelle (Gondran, /26/) que M est p-régulier si $M^{(p)} = M^{(p+1)}$. On voit facilement que M est p-régulier si et seulement si $M \leq 0$, et alors $p = 0$.

Dualement, $(\tilde{\mathcal{J}}(\mathbb{R}_U[+\infty]), \tilde{\min}, \emptyset)$ forme un dioïde dont les éléments 0-réguliers sont les $M \geq 0$. Ces remarques seront utiles pour savoir calculer la longueur de plus courts chemins dans un graphe imprécisément valué (cf. Ch. 7).

Les opérations $\tilde{\max}$ et $\tilde{\min}$ sont utiles pour définir l'union et l'intersection de sous-ensembles flous de type 2 (Dubois et Prade (/15/), c'est-à-dire dont les valeurs d'appartenance sont des quantités floues sur $[0, 1]$).

NB : Autres opérations. Toujours dans l'optique de la combinaison d'ensembles flous de type 2, on peut étendre des fonctions isotones telles que, t-normes, t-conormes, moyennes, sommes symétriques à des arguments flous. Par exemple si la fonction ψ monotone croissante génère une co-norme \star , le calcul de $M \star N$ quand M et N sont de type LR se ramènera à des équations du type trouver u :

$$\psi(w) = \psi(a+bu) + \psi(c+du)$$

Ce qui peut être simple pour certains générateurs (fonctions logarithmes, fonctions homographiques, etc...).

Pour des discussions sur l'extension de la fonction puissance x^y , voir Dubois et Prade (/14/, /16/), Dubois⁽¹⁾. Dans certains

(1) dans "Notes sur l'intérêt des sous-ensembles flous en analyse de l'attraction de point de vente" - Rapport N°35, IME, Université de Dijon -

cas on peut démontrer des égalités telles que

$$M^N \otimes M^P = M^{N \oplus P}$$

Mais on a toujours $e^{M \oplus P} = e^M \otimes e^P$

4.2.6 - Le problème des représentations équivalentes d'une fonction

Il existe toujours plus d'une seule représentation mathématique d'une fonction de variable réelle. Soit f et g deux fonctions telles que $f = g$ sur \mathbb{R} . Il est facile de voir qu'en général $f(M, N) \neq g(M, N)$ quand M et N sont des ensembles flous. On a rencontré ce cas à propos de la distributivité du produit sur la somme d'intervalles flous. Moore (/35/) le signale à propos du calcul d'intervalles.

Soit f_1 une représentation mathématique d'une fonction f de deux variables x et y . Soit A_1 l'ensemble des arguments de f_1 , soit I_1 le sous-ensemble (d'arguments) de A_1 qui fait que

$$\left(x_i = x \quad \forall i \in I_1 \quad \text{et} \quad x_i = y \quad \forall i \in A_1 - I_1 \right) \rightarrow f_1 = f$$

Si $|A_1| > 2$, f_1 est dite représentation impropre de f . On peut montrer le résultat suivant

Proposition 4.6 : Si M et N sont des intervalles flous s.c.s.

$$f(M, N) \subseteq f_1(M, N) \quad \text{où} \quad f_1(M, N) \text{ est obtenu}$$

en substituant M aux arguments dans I_1 et N aux arguments de $A_1 - I_1$. De plus si f_1 est isotone (en fonction de $x_i \quad \forall i$) alors

$$f_1(M, N) = f(M, N)$$

(NB = alors f est aussi isotone).

Preuve :

$$\mu_{f_1(M, N)}(w) = \sup \left\{ \min \left(\min_{i \in I_1} \mu_M(x_i), \min_{i \notin I_1} \mu_N(x_i) \right) / f_1(\vec{x}) = w \right\} \quad (46)$$

où $\vec{x} = (x_1, x_2, \dots, x_{|A_1|})$

Il est clair que

$$\begin{aligned} \{\vec{x}/f_1(\vec{x}) = w\} \supseteq \{\vec{x}/f_1(\vec{x}) = f(x,y) = w, x_i = x \text{ pour } i \in I_1 \\ \text{et } x_i = y \text{ pour } i \in A_1 - I_1\} \end{aligned}$$

donc

$$\mu_{f_1(M,N)}(w) \geq \mu_{f(M,N)}(w)$$

Par ailleurs, si f_1 est isotone, $f_1(M_1, M_2 \dots M_{|A_1|})_\alpha$

$$= f_1((M_1)_\alpha, (M_2)_\alpha, \dots, (M_{|A_1|})_\alpha) = [f(a_1 \dots a_{|A_1|}), f(b_1 \dots b_{|A_1|})]$$

$$\text{si } (M_i)_\alpha = [a_i, b_i] \forall i = 1, \dots, |A_1|$$

$$= [f(a,c), f(b,d)] \quad \text{pour } (M_i)_\alpha = M_\alpha = [a,b], i \in I_1$$

$$(M_i)_\alpha = N_\alpha = [c,d], i \in A_1 - I_1$$

$$= f(M_\alpha, N_\alpha) = (f(M,N))_\alpha$$

Q.E.D.

Par exemple $f(x,y) = x-x$ est une représentation impropre de $f(x,y)$ pour l'arithmétique floue car $f(x_1, x_2) = x_3 + x_4$ n'est pas isotone.

$$f(M,N) \subset f(M,N) \oplus M \oplus M \quad (\text{car } M \oplus M \neq 0)$$

Moins artificiel est le cas de polynomes :

si

$$f(x) = \sum_{i=0}^n a_i x^i$$

alors généralement $f(M) \neq \sum_{i=0}^n a_i \overbrace{M \otimes \dots \otimes M}^i$ où l'expression de droite est calculée en effectuant les produits de M par lui-même, puis la somme pondérée des intervalles flous obtenus. Néanmoins on calcule ainsi un encadrement de $f(M)$, d'après la proposition 4.7. Ce résultat est signalé par Moore (/35/). Si l'encadrement est bon, il peut suffire à la connaissance de $f(M)$. $f(M)$ ne peut être connu exactement que si on sait résoudre les équations de $n^{\text{ième}}$ degré:

$$\forall w, w = \sum_{i=0}^n a_i x^i$$

La mise en facteur d'un polynome mène à des représentations non équivalentes pour le calcul d'intervalles flous.

Par exemple

$$\begin{aligned} f_1(M) &= 105M \oplus M \oplus M \oplus M \\ &\neq 10M \circ (5M \oplus M) \end{aligned}$$

De même $102M \oplus M \oplus M \neq (10M)^2$

Mais $(10M)^2$ est réellement ce que vaut $f(x) = 1 - 2x + x^2$ pour $x = M$ car M n'apparaît qu'une fois.

Enfin notons que $M \oplus M = 2M$ est toujours vrai tandis que $M \circ M = M^2$ seulement si M est positif ou négatif.

En effet $S(M^2) \subseteq \mathbb{R}^+$ tandis que si $M = M_+ \cup M_-$ où $S(M_-) \subseteq \mathbb{R}^-$, $S(M_+) \subseteq \mathbb{R}^+$

$$\begin{aligned} M \circ M &= M_+^2 \cup M_-^2 \cup (M_+^+ \circ M_-^-) \\ &= M^2 \cup (M_+^+ \circ M_-^-) \end{aligned}$$

et $S(M_+^+ \circ M_-^-) \subseteq \mathbb{R}^-$.

Ce que dit la proposition 4.6 c'est que si une représentation de f ne fait apparaître qu'une seule fois ses arguments alors cette représentation fournira la bonne fonction d'appartenance de $f(M,N)$. Une telle représentation n'existe pas toujours : par exemple la fonction $f(x,y,z) = xy + yz + zx$ où l'un des arguments apparaît toujours au moins deux fois ! Néanmoins deux représentations isotones s'avèrent toujours équivalentes.

4.2.7 - Analogues possibilistes de la loi des grands nombres

Dans l'étude des variables aléatoires, on sait que si x_1, \dots, x_n sont des variables aléatoires indépendantes de même distribution, et de moyenne \bar{x} alors la distribution de $\frac{x_1 + \dots + x_n}{n}$ tend, lorsque n croît, vers un Dirac de poids 1 sur \bar{x} . C'est-à-dire que l'accumulation d'expériences effectuées dans les mêmes conditions permet peu à peu d'améliorer la précision du résultat (précis mais fluctuant) observé.

Avec des variables possibilistes non interactives le problème revient à se demander vers quoi tend $\frac{M \oplus M \dots \oplus M}{n}$ quand n tend vers l'infini, où M est une quantité floue.

On cite ci-dessous les résultats déjà obtenus sur ce problème :

. (Stein & Talati, /45/). Si M est un ensemble flou convexe alors

$$\forall (p_1 \dots p_n) \text{ tel que } p_1 + p_2 + \dots + p_n = 1$$

$$p_1 M \oplus p_2 M \oplus \dots \oplus p_n M = M \quad (47)$$

. (Stein & Talati, /45/). Si $\forall (p_1 \dots p_n)$ tel que $p_1 + \dots + p_n = 1$, (47) est vrai alors M est un ensemble flou convexe. En particulier (47) est valide pour un intervalle flou,

. En revanche si $(p_1 \dots p_n)$ sont fixés alors (47) peut très bien être vrai sans que M soit convexe ; par exemple $M = \{\text{rationels de } [0,1]\}$ et $p_1 \dots p_n$ est rationel. (Dubois /12/).

Mais on peut trouver des classes de quantités floues sur lesquelles pour un n -uplet $p_1 \dots p_n$ fixé, (47) est équivalent à la convexité de M . C'est le cas quand :

. (Dubois, /12/) M est l'union finie de quantités floues convexes

. (Stein, /44/) μ_M est s.c.s.

Enfin quand M n'est pas convexe, on peut sous certaines hypothèses montrer que si $x_1 \dots x_n$ sont des variables possibilistes non-interactives, alors la distribution de $\frac{x_1 + x_2 + \dots + x_n}{n}$ se convexifie peu à peu.

Si M est une quantité floue son enveloppe convexe est (Lowen /30/)
 \hat{M} telle que

$$\mu_{\hat{M}}(w) = \sup_M \{ \min(\mu_M(u), \mu_M(v)) / u \leq w \leq v \}$$

\hat{M} est un intervalle flou dont les α -coupes sont les enveloppes convexes des M_α . On a alors (Dubois, /12/)

. Si μ_M est continue alors $\lim_{n \rightarrow +\infty} \frac{x_1 + x_2 + \dots + x_n}{n}$ a pour distribution \hat{M} .

L'expression (47) contraste avec la loi des grands nombres probabiliste . Elle ne fait qu'affirmer que si un appareil de mesure est imprécis alors on est limité par son imprécision. On ne peut pas améliorer un résultat de mesure imprécis en répétant l'expérience, si elle doit toujours fournir le même résultat. On retrouve ici la différence entre probabilité (précision et fluctuation) et possibilité (imprécision et cohérence).

On pourrait se reposer le problème de la loi des grands nombres dans le cadre de l'expérience statistique imprécise esquissée au chapitre 1 § 6 c'est-à-dire, supposer que l'instrument de mesure fournit des intervalles aléatoires, dont les tirages sont soumis à la probabilité de base qui définit l'intervalle flou M. On peut conjecturer que

$$\frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \left[\sum_{i=1}^n \underline{x}_i, \sum_{i=1}^n \bar{x}_i \right]$$

(où $X_i = [\underline{x}_i, \bar{x}_i]$ est le résultat de la $i^{\text{ème}}$ mesure) tend vers $[E_*(M), E^*(M)]$, c'est-à-dire la valeur moyenne de la quantité floue M (voir (4.1.3)). En effet les variables aléatoires \underline{x}_i et \bar{x}_i ont respectivement pour fonctions de répartition F_* et F^* associées à la probabilité de base.

On notera au passage la propriété d'additivité de la valeur moyenne d'une somme de deux quantités floues au sens du principe d'extension :

$$E^*(M \oplus N) = E^*(M) \oplus E^*(N)$$

et

$$E_*(M \oplus N) = E_*(M) \oplus E_*(N)$$

Ce résultat est prouvé dans Dubois et Prade (/19/) quand les fonctions de répartition inférieures et supérieures de M et N sont s.c.s. on a alors aussi :

$$E^*(\min(M, N)) + E^*(\max(M, N)) = E^*(M) + E^*(N)$$

et la même chose pour les espérances inférieures.

4.3 - ARITHMETIQUE FLOUE PESSIMISTE AVEC INTERACTION DES VARIABLES

Dans toute la section 4.2 on a supposé que pour calculer le domaine de variation de $f(x,y)$ les variables x et y n'étaient pas interactives, c'est-à-dire que le domaine de variation du couple (x,y) était le produit cartésien $M \times N$ défini à l'aide de l'opérateur minimum.

Quand ce n'est pas le cas, c'est qu'une dépendance relationnelle entre x et y existe sous la forme d'un domaine D du référentiel \mathbb{R}^2 de (x,y) . Le principe d'extension devient :

$$\forall w, \mu_{f(M,N;D)}(w) = \sup \{ \min(\mu_M(u), \mu_N(v)) : w = f(u,v), (u,v) \in D \} \quad (48)$$

D peut éventuellement se réduire à un lien fonctionnel du type

$$\exists g, \quad g(x,y) = 0$$

NB : on remarquera que le calcul de $f(M,N)$ à partir d'une représentation impropre de f au sens de la proposition 4.7, est un cas particulier de (48)

On peut aussi imaginer que D est une relation floue définissant pour les valeurs du couple (x,y) des zones plus ou moins permises. Alors (48) s'écrit

$$\forall w, \mu_{f(M,N;D)}(w) = \sup \{ \min(\mu_M(u), \mu_N(v), \mu_D(u,v)) / f(u,v) = w \} \quad (49)$$

L'étude des propriétés, et a fortiori le calcul de $f(M,N;D)$ peuvent être très difficile à mener en général. On donne ci-dessous deux exemples, considérés comme intéressants, d'interactivité où on sait étudier et calculer $f(M,N;D)$: un cas d'interaction linéaire, et un cas où M et N sont les projections d'une relation floue faiblement séparable (chap.3). Ces deux cas sont étudiés dans Dubois et Prade (/17/,/21/,/23/).

4.3.1 - Dépendance linéaire entre variables

Etant donné n variables $x_1 \dots x_n$ chacune limitée par un intervalle flou M_i ($i=1,n$), on cherche à calculer la distribution de possibilité de la variable $z = a_1 x_1 + \dots + a_n x_n$ où les a_i sont des constantes réelles,

sachant que les x_i sont liés par la contrainte $x_1 + x_2 + \dots + x_n = 1$ qui définit le domaine D. Soit N l'ensemble flou des valeurs possibles de Z. On montre alors (Dubois et Prade (/21/))

. $\sup \mu_N = \mu_{M_1 \oplus M_2 \oplus \dots \oplus M_n}$ (1). Si cette valeur est inférieure à 1 on peut tronquer les M_i pour calculer N.

. Si les M_i sont des intervalles flous s.c.s. à support compact alors N est un intervalle flou s.c.s. et N_α se calcule à partir des $M_{i\alpha} \forall \alpha \in [0, 1]$. On peut calculer explicitement les bornes de N_α en fonction de celles de $M_i (i=1, n)$.

Par exemple si

$$0 \leq a_1 \leq a_2 \leq \dots \leq a_n \quad \text{et} \quad M_{i\alpha} = [b_i, B_i]$$

alors $\forall \alpha \leq \sup \mu_N$:

$$\inf N_\alpha = \max_{k=1, n} \left(\sum_{j=1}^{k-1} B_j a_j + (1 - \sum_{j=1}^{k-1} B_j + \sum_{j=k+1}^n b_j) a_k + \sum_{j=k+1}^n b_j a_j \right) \quad (50)$$

$$\sup N_\alpha = \min_{k=1, n} \left(\sum_{j=1}^{k-1} b_j a_j + (1 - \sum_{j=1}^{k-1} b_j + \sum_{j=k+1}^n B_j) a_k + \sum_{j=k+1}^n B_j a_j \right) \quad (51)$$

. Si les M_i sont tous de type LL alors N est aussi de type LL quand il est normalisé (Dubois et Prade, /17/).

. Si $M_i = (m_i, \alpha_i, \beta_i)_{LL}$ alors $N = (\sum m_i a_i, \gamma, \delta)$ où γ (resp : δ) peut être obtenu en posant $\inf(N_\alpha) = \sum m_i a_i - \gamma$ (resp : $\sup N_\alpha = \sum m_i a_i + \delta$), $B_j = m_j + \beta_j$ et $b_j = m_j - \alpha_j$ dans les formules ci-dessus.

Dans le cas où les variables x_i sont liées par une relation linéaire quelconque on peut par simple changement de variable de type $y_i = k_i x_i$ se ramener à la forme normalisée $\sum y_i = 1$.

Dans Dubois et Prade (/21/), on donne un algorithme pour calculer N quand les a_i 's sont aussi des variables possibilistes de distribution μ_{A_i} . On montre que

. si M_i et A_i sont des intervalles flous positifs et s.c.s., à support compact

alors N_α (somme interactive de produits d'intervalles flous) est aussi un intervalle flou positif s.c.s. et N_α s'obtient à partir des α -coupes $M_{i\alpha}$ et $A_{i\alpha}$

. $\inf N_\alpha$ s'obtient en appliquant (50) avec $a_j = \inf A_{j\alpha}$ et $\sup N_\alpha$ en appliquant (50) avec $a_j = \sup A_{j\alpha}$.

L'intérêt de savoir calculer N est de pouvoir évaluer des espérances mathématiques quand les valeurs de probabilité sont mal connues. On peut aussi calculer des moyennes pondérées (des bary-centres...) quand les poids sont imprécis, définis par des intervalles flous W_i :

Soit la fonction

$$f(x_1 \dots x_n) = \frac{a_1 x_1 + a_2 x_2 + \dots + a_n x_n}{x_1 + x_2 + \dots + x_n}$$

Cette fonction n'est ni isotone, ni antitone, ni hybride pour $n > 2$. Les variables apparaissent deux fois dans l'expression ci-dessus qui est donc impropre. Baas et Kwakernaak (/2/) ont donné une méthode très lourde pour calculer $f(W_1 \dots W_n)$. Dans Dubois et Prade (/21/) on suggère un calcul approché de $f(W_1 \dots W_n)$ plus précis que l'encadrement

$$F(W_1 \dots W_n) = \frac{a_1 W_1 \oplus a_2 W_2 \dots \oplus a_n W_n}{W_1 \oplus W_2 \dots \oplus W_n}$$

Il consiste à

. calculer M_i t.q. $\mu_{M_i} = x_i / \Sigma x_j$, ce qui est facile puisque la fonction $x_i / \Sigma x_j$ est hybride. Si W_k est de type LL, on se ramène à la résolution de deux équations du 1er degré.

. Calculer $\bar{f}(M_1 \dots M_n; D)$ où D est la contrainte linéaire normalisée $\Sigma y_i = 1$ (M_i limite y_i).

On vérifie que :

$$f(W_1 \dots W_n) \subseteq \bar{f}(M_1 \dots M_n; D) \subseteq F(W_1 \dots W_n)$$

En particulier si $n = 2$: $f(W_1 \dots W_n) = \bar{f}(M_1 \dots M_n; D)$.

4.3.2 - Additions de nombres flous au sens d'une t-norme

Considérons une version plus générale du principe d'extension, pour faire l'addition de deux intervalles flous M et N . Etant donné une t-norme \perp , on définit l'addition $M \oplus N$ au sens de \perp par

$$\forall w, \mu_{M \oplus N}(w) = \sup \{ \mu_M(u) \perp \mu_N(w-u), u \in \mathbb{R} \} \quad (52)$$

Si $\perp < \min$, on rappelle que (52) signifie qu'on interdit à x et y de s'éloigner simultanément des noyaux de leurs domaines M et N , l'interdiction étant d'autant moins forte que \perp est proche de \min . Si $\perp = T_w$ alors seule l'une des deux variables peut sortir du noyau. On a le résultat suivant (Dubois et Prade, /23/).

Proposition 4. : Si M et N sont des intervalles flous s.c.s. à support borné, alors $M \oplus N$ est un intervalle flou, quel que soit \perp . La démonstration utilise le fait que, pour tout \perp ,

$$(M \cup P) \oplus N = (M \oplus N) \cup (P \oplus N) \quad (\cup = \max)$$

est encore valide. En décomposant μ_M en parties croissante et décroissante i.e. $M = M^- \cup M^+$, $N = N^- \cup N^+$ on montre alors que $M^\eta \oplus M^\varepsilon$, $(\eta, \varepsilon) \in \{-, +\}^2$ sont tous convexes. Le résultat peut d'ailleurs s'étendre à un certain nombre de fonctions isotones qui définissent une structure de groupe sur un intervalle de \mathbb{R} . La recherche du suprémum dans (52) peut d'ailleurs se restreindre à un intervalle fermé de \mathbb{R} puisqu'on peut facilement montrer (Dubois et Prade, /17/) :

$$\forall w \leq m+n \quad \mu_{M \oplus N}(w) = \sup_{w-n < u < m} \mu_M(u) \perp \mu_N(w-u) \quad (53)$$

où m et n sont des valeurs modales de M et N telles que $n \geq m$. Pour $w \geq m+n$ l'intervalle contenant le maximum est $[m, w-n]$. On a bien sûr

$$\mu_{M \oplus N}(m+n) = 1$$

La résolution de (53) pour les t-normes de base a été menée à bien dans Dubois et Prade (/17/). On en résume ici les conclusions :

. $M \oplus N$ a pour fonction d'appartenance

$$\mu_{M \oplus N}(w) = \max(\mu_M(w-n), \mu_N(w-m)) \quad (54)$$

dans les cas suivants :

. $\perp = T_w$ (la plus petite t-norme)

. $a \perp b = \max(0, a+b-1)$ pourvu que $\mu_M^+, \mu_M^-, \mu_N^+, \mu_N^-$ soient convexes hors des noyaux de M et N respectivement (c'est le cas si M et N ont des distributions linéaires)

. $a \perp b = \text{produit}$, pourvu que $\text{Log } \mu_M^+, \text{Log } \mu_M^-, \text{Log } \mu_N^+, \text{Log } \mu_N^-$ soient convexes hors des noyaux de M et N respectivement. Quand M et N sont de type L-R et $M = (m, \alpha, \beta)_{LR}$, $N = (n, \gamma, \delta)_{LR}$ alors (54) s'écrit $M \oplus N = (m+n, \max(\alpha, \gamma), \max(\beta, \delta))_{LR}$, ce qui contraste avec $M \oplus N$ construit avec $\perp = \min$. On notera que (54) veut encore dire $M \oplus N = (M^- \oplus N^+) \cup (M^+ \oplus N^-)$ où les additions dans le membre gauche sont effectuées au sens de n'importe quelle t-norme, puisque

$$\forall \perp, \quad \mu_{M^- \oplus N^+}(w) = \mu_M(w-n), \quad w \leq m+n \\ = \mu_N(w-m), \quad w \geq m+n$$

Quand $a \perp b = \max(a+b-1, 0)$ (52) a été résolu aussi pour deux nombres flous de type LR où les références sont de la forme $L(x) = \max(0, 1-u^p)$. Par exemple l'étalement gauche de $M \oplus N$ s'écrit

$$\alpha_{M \oplus N} = \left(\alpha^{\frac{p}{p-1}} + \gamma^{\frac{p}{p-1}} \right)^{\frac{p-1}{p}} \quad p > 1 \quad (55)$$

Quand p varie entre 0 et $+\infty$, $\alpha_{M \oplus N}$ passe de $\max(\alpha, \beta)$ à $\alpha + \beta$. Par transformation logarithmique le résultat ci-dessus est aussi valable pour des références de type $L(x) = e^{-x^p}$, et L = produit.

Badard (/3/) a montré que le calcul de (52) était facile quand un nombre flou symétrique, de références $L(x) = e^{-P(x)} = R(x)$, où P est convexe, était additionné avec lui-même au sens de la t-norme produit. Et on a (pour $m = 0$)

$$\mu_{M \oplus M}(w) = e^{-2P\left(\frac{w}{2}\right)}$$

ce résultat est facilement transposable au cas d'une somme de deux nombres flous identiques de fonction d'appartenance concave, au sens de la t-norme $\max(0, a+b-1)$.

Le calcul explicite de $M \oplus N$ peut aussi se faire pour des t-normes strictes ou nil-potentes en explicitant le générateur. On peut alors, moyennant une transformation des fonctions d'appartenance se ramener aux situations résolues pour $\perp = \max(0, a+b-1)$ ou produit. Par exemple si \perp est nil-potente de générateur f , on pourra appliquer (54) dès que $f \circ \mu_M^+$, $f \circ \mu_M^-$, $f \circ \mu_N^+$, $f \circ \mu_N^-$ seront convexes.

NB1 : Alsina et Nguyen (/1/) ont montré que pour des fonctions d'appartenance de type gaussien ou de type fonction gamma, l'additivité des étalements avec conservation de la forme n'est valide que si $M \oplus N$ est calculé au sens de la t-norme min.

NB2 : On notera que généralement, avec une t-norme autre que min, on n'a plus $M \oplus M = 2M$, puisque $2M$ est un nombre flou de fonction d'appartenance $\mu_{2M}(u) = \mu_M(u/2)$. En particulier la somme

$$M_n = \frac{M \oplus N \oplus M \oplus \dots \oplus M}{n}$$

n'est plus invariante.

On peut facilement montrer que dans les cas résolus ci-dessus

$$\lim_{n \rightarrow +\infty} M_n = \overset{\circ}{M} \text{ le noyau de } M \quad (56)$$

résultat déjà montré par Badard (/3/) pour des nombres flous symétriques. Ce type de convergence devrait pouvoir être démontré dans de nombreux cas ; en particulier si (56) est validé pour M au sens de la t-norme \perp alors si $\perp' \leq \perp$ et $N \subseteq M$, (56) est valide pour N au sens de \perp' .

Notons enfin que dans des situations où (55) s'applique on devrait pouvoir trouver des propriétés analogues à la tendance vers la loi normale des probabilités. Un exemple est donné dans (Dishkant, /8/) avec la t-norme nil-potente typique.

4.4 - COMPARAISON D'INTERVALLES FLOUS

A côté du problème des manipulations arithmétiques des quantités floues, se pose celui de leur comparaison en vue d'un classement. Ce qui était fort simple dans le cas de nombres précis devient beaucoup moins trivial. La question se pose à savoir : est ce que la théorie des possibilités nous fournit des informations suffisantes pour comparer des intervalles flous. Un long article (Dubois et Prade, /20/) tente de répondre affirmativement et propose un jeu d'indices de comparaisons dans le cadre de la théorie de Zadeh. Cette section résume les résultats de cet article et présente aussi quelques concepts qui en dépendent tels les domaines à frontière floue.

Le problème de comparaison a été abordé assez souvent dans la littérature des ensembles flous, sans avoir reçu de réponse satisfaisante. (Voir Dubois et Prade (/20/) pour une revue sur le sujet). L'approche ci-dessous tente de combler ce manque.

4.4.1 - Comparaison d'un réel et d'une quantité floue

Avant de savoir comparer deux quantités floues, il convient d'abord de savoir dire comment un nombre réel se place par rapport à un intervalle flou. Le calcul des possibilités nous permet de définir les ensembles des nombres qui peuvent être plus grands ou égaux (resp : sont nécessairement plus grands) que les valeurs d'une variable associée à une quantité floue Q . On les notera respectivement $[Q, +\infty)$ et $]Q, +\infty)$, tels que (Dubois et Prade /18/) :

$$\forall w, \mu_{[Q, +\infty)}(w) = \Pi_Q((-\infty, w]) \quad (57)$$

$$= \sup_{u \leq w} \mu_Q(u)$$

$$\mu_{]Q, +\infty)}(w) = \mathcal{N}_Q((-\infty, w[) \quad (58)$$

$$= \inf_{u > w} 1 - \mu_Q(u)$$

on remarquera que $[Q, +\infty)$ et $]Q, +\infty)$ sont convexes, et leurs fonctions d'appartenance sont les fonctions de répartitions supérieures et inférieures de Q , (cf (10) et (11), section 1.3) (aux points de discontinuité près).

De façon similaire on peut définir $(-\infty, Q]$ (resp : $(-\infty, Q[$), qui sont les ensembles flous de nombres possiblement (resp : nécessairement) plus petits que Q , tels que

$$(-\infty, Q] = \overline{]Q, +\infty)} \quad (59)$$

$$(-\infty, Q[= \overline{[Q, +\infty)} \quad (60)$$

Ce symbolisme est justifié par le fait que si Q se réduit à un réel u , alors $[Q, +\infty)$ (resp : $]Q, +\infty)$) se réduit à la demi-droite $[u, +\infty)$ (resp : $]u, +\infty)$). On vérifie

$$\begin{aligned} (-\infty, Q] \wedge [Q, +\infty) &= \text{enveloppe convexe de } Q \\ &= Q \text{ si } Q \text{ est convexe} \end{aligned} \quad (61)$$

$$(-\infty, Q[\wedge]Q, +\infty) = \emptyset \quad (62)$$

où l'intersection est au sens du minimum dans (61).

On notera que lorsque Q est flou $Q \cup]Q, +\infty) \subset [Q, +\infty)$ généralement (inclusion stricte, sauf si Q est défini par la somme bornée, par exemple).

Les opérateurs $\tilde{\max}$ et $\tilde{\min}$ s'expriment agréablement en fonction des quantités floues ci-dessus introduites (Dubois et Prade /18/).

$$\tilde{\max}(M, N) = ([M, +\infty) \wedge N) \cup (M \wedge [N, +\infty)) \quad (63)$$

$$\tilde{\min}(M, N) = ((-\infty, M] \wedge N) \cup (M \wedge (-\infty, N]) \quad (64)$$

4.4.2 - Comparaison de deux quantités floues

Dans la mesure où elles peuvent se chevaucher, il n'y aura pas forcément de réponse nette à la question de trouver la plus grande de deux quantités floues. Une première idée naturelle consiste à utiliser $\tilde{\max}$ et $\tilde{\min}$ pour discriminer (Dubois et Prade, /13/). Néanmoins on peut facilement voir que deux quantités floues très proches, se chevauchant de façon significative sont malgré tout discriminées par $\tilde{\max}$ et $\tilde{\min}$. Par exemple $\tilde{\max}(M, N) = N$ dès que M est un intervalle flou et $\forall \varepsilon > 0$ $\mu_N(w) = \mu_M(w - \varepsilon)$, $\forall w$. En revanche $\tilde{\max}$ ne discrimine plus dès que les parties croissantes ou décroissantes des fonctions d'appartenance se coupent, quel qu'éloignés que puissent être les noyaux de M et N . Ce dont on a

besoin, c'est de justifier une éventuelle discrimination de façon quantitative, ce qui, comme l'ont fait remarquer Baldwin et Guild (/4/) n'est pas obtenu avec $\overset{\sim}{\max}$ et $\overset{\sim}{\min}$

Dans le cadre de la théorie des possibilités, on peut, pour savoir si M est plus grand que N comparer les ensembles M et $[N, +\infty)$ d'une part, M et $]N, +\infty)$ d'autre part, à l'aide d'un indice de comparaison du type de ceux définis au chapitre 2, §4.2. Pour être complet, il faut envisager un évaluateur existentiel et un évaluateur universel, c'est-à-dire calculer la possibilité et la nécessité au sens de M, des événements flous $[N, +\infty)$ et $]N, +\infty)$ (chap.2 § 4.3.2). Si Π_M et \mathcal{N}_N sont les mesures de possibilité et de nécessité générées par μ_M , on a donc quatre indices de comparaison fondamentaux :

. Possibilité de surclassement au sens large de N par M

$$\begin{aligned} \Pi_M([N, +\infty)) &= \sup_u \min(\mu_M(u), \sup_{v \leq u} \mu_N(v)) \\ &= \sup_{u > v} \min(\mu_M(u), \mu_N(v)) \end{aligned} \quad (65)$$

. Possibilité de surclassement strict de N par M

$$\begin{aligned} \Pi_M(]N, +\infty)) &= \sup_u \min(\mu_M(u), \inf_{v > u} 1 - \mu_N(v)) \\ &= \sup_u \inf_{v > u} \min(\mu_M(u), 1 - \mu_N(v)) \end{aligned} \quad (66)$$

. Nécessité de surclassement au sens large de N par M

$$\begin{aligned} \mathcal{N}_M([N, +\infty)) &= \inf_u \max(1 - \mu_M(u), \sup_{v \leq u} \mu_N(v)) \\ &= \inf_u \sup_{v \leq u} \max(1 - \mu_M(u), \mu_N(v)) \end{aligned} \quad (67)$$

. Nécessité de surclassement strict de N par M

$$\begin{aligned} \mathcal{N}_M(]N, +\infty)) &= \inf_u \max(1 - \mu_M(u), \inf_{v > u} 1 - \mu_N(v)) \\ &= 1 - \sup_{u < v} \min(\mu_M(u), \mu_N(v)) \end{aligned} \quad (68)$$

(65) étudie l'existence d'une intersection entre M et $[N, +\infty)$. Cet indice est déjà connu dans la littérature. Il a été suggéré par Baas et Kwakernaak /2/.
 (68) étudie l'inclusion de M dans $]N, +\infty)$ (pas au sens de Zadeh, mais au sens de \leftarrow , voir chap. 2 § 3.5.1), et de par la relation de dualité possibilité-nécessité on a

$$\Pi_M([N, +\infty)) = 1 - \mathcal{N}_N(]M, +\infty)) \quad (69)$$

parce que (65) implique

$$\Pi_M([N, +\infty)) = \Pi_N((-\infty, M])$$

En revanche $\Pi_M(]N, +\infty))$ et $\mathcal{N}_M([N, +\infty))$ n'avaient jamais été suggérés auparavant ; (mais récemment Tsukamoto et al. ont proposé une famille d'indices de comparaison équivalente, sans référence à la théorie des possibilités /45/).

Le figure ci-contre permet de visualiser les indices proposés, et met en valeur, l'indépendance entre $\Pi_M(]N, +\infty))$ et $\mathcal{N}_M([N, +\infty))$. On peut démontrer les propriétés suivantes :

$$\Pi_M([N, +\infty)) \geq \max(\Pi_M(]N, +\infty)), \mathcal{N}_M([N, +\infty)) \quad (70)$$

$$\min(\Pi_M(]N, +\infty)), \mathcal{N}_M([N, +\infty)) \geq \mathcal{N}_M(]N, +\infty)) \quad (71)$$

Ces quatre indices sont nécessaires et généralement suffisants pour caractériser les positions respectives de deux intervalles fermés usuels. Les configurations pathologiques sont celles où $M \cap N$ est un singleton.

En fait l'approche possibiliste permet de construire 16 valeurs décrivant les positions respectives de M et de N . Il suffit pour les générer de changer M en N ou $[M, +\infty)$ en $(-\infty, M]$ etc... dans (65-68).

L'étude de ces 16 valeurs (voir Dubois et Prade /20/) indique que seules 6 d'entre elles sont indépendantes dans le cas général, comme le suggère le tableau résumé ci-après.

$$\begin{aligned}
\Pi_M([N, +\infty)) &= 1 - \mathcal{N}_M((-\infty, N[) = 1 - \mathcal{N}_N(]M, +\infty)) = \Pi_N((-\infty, M]) \\
\Pi_M(]N, +\infty)) &= 1 - \mathcal{N}_M((-\infty, N]) \leq 1 - \Pi_N(]M, +\infty)) = \mathcal{N}_N((-\infty, M]) \\
\mathcal{N}_M([N, +\infty)) &= 1 - \Pi_M((-\infty, N[) \geq 1 - \mathcal{N}_N([M, +\infty)) = \Pi_N((-\infty, M[) \\
\mathcal{N}_M(]N, +\infty)) &= 1 - \Pi_M((-\infty, N]) = 1 - \Pi_N([M, +\infty)) = \mathcal{N}_N((-\infty, M[)
\end{aligned}$$

Ces 6 valeurs sont nécessaires et suffisantes pour caractériser les positions respectives de deux intervalles réels quelconques. Néanmoins, on peut montrer que, la plupart du temps

$$\Pi_M(]N, +\infty)) = 1 - \Pi_N(]M, +\infty)) \quad (72)$$

$$\mathcal{N}_M([N, +\infty)) = 1 - \mathcal{N}_N([M, +\infty)) \quad (73)$$

(72-73) peuvent être faux quand μ_M et μ_N sont discontinues en un même point qui est justement un extremum de (66) ou (67). Donc on vérifie bien la pertinence et la complétude des quatre indices de comparaison proposés plus haut.

NB : On peut nuancer un indice de comparaison entre deux quantités floues en remplaçant la relation \geq dans (65-68) par une relation floue traduisant à quel point on souhaite qu'une quantité soit plus grande qu'une autre. Cette relation floue de tolérance notée R peut se combiner à l'un des nombres flous, ce qui donnerait les indices.

$$\Pi_M(\text{Ro}[N, +\infty)) = \Pi_{\text{RoM}}([N, +\infty))$$

$$\Pi_M(\text{Ro}]N, +\infty)) = \Pi_{\text{RoM}}(]N, +\infty))$$

$$\mathcal{N}_M(\text{Ro}[N, +\infty)) = \mathcal{N}_{\text{RoM}}([N, +\infty))$$

$$\mathcal{N}_M(\text{Ro}]N, +\infty)) = \mathcal{N}_{\text{RoM}}(]N, +\infty))$$

Cette idée a été suggérée par Baldwin et Guild /5/ et appliquée au premier des indices ci-dessus.

4.4.3 - Classement de n quantités floues

Il y a deux approches envisageables pour classer un ensemble de n intervalles flous $\{M_1 \dots M_n\}$ en tirant parti des indices possibilistes

introduits

- définir des indices globaux de surclassement qui évaluent à quel point M_i domine tous les autres nombres flous
- traiter les relations floues obtenues par comparaison deux à deux des M_i

4.4.3.1 - INDICES GLOBAUX DE SURCLASSEMENT

On peut définir un indice de surclassement global assez naturellement, en interprétant la domination de M_i sur les autres intervalles flous comme le fait que M_i surclasse le plus grand des M_j , $j \neq i$. L'opérateur $\overset{\sim}{\max}$ s'introduit naturellement ici : 4 indices de surclassement peuvent être construits sur (65-68) en comparant M_i et $\overset{\sim}{\max}_{j \neq i} M_j$:

$$PS(M_i) = \Pi_{M_i} (\overset{\sim}{\max}_{j \neq i} M_j, +\infty) \quad (\text{Possibilité de surclassement})$$

$$PSS(M_i) = \Pi_{M_i} (]\overset{\sim}{\max}_{j \neq i} M_j, +\infty) \quad (\text{Possibilité de surclassement strict})$$

$$NS(M_i) = \mathcal{N}_{M_i} (\overset{\sim}{\max}_{j \neq i} M_j, +\infty) \quad (\text{Nécessité de surclassement})$$

$$NSS(M_i) = \mathcal{N}_{M_i} (]\overset{\sim}{\max}_{j \neq i} M_j, +\infty) \quad (\text{Nécessité de surclassement strict})$$

L'ensemble $\{M_1 \dots M_n\}$ peut être classé selon les valeurs de chaque indice. On peut ainsi obtenir 4 ordres linéaires qui s'ils sont cohérents, permettent de définir un classement définitif. Néanmoins, à cause de l'imprécision, ces 4 indices ne fourniront pas forcément le même classement (voir ch.6).

4.4.3.2 - RELATIONS DE SURCLASSEMENT FLOUES

Les quatre indices de comparaison (65 - 68) vont définir trois relations floues seulement, si l'on se rappelle de l'identité (69).

. Relation de possibilité de surclassement

On note PS la relation floue définie par

$$\mu_{PS}(M_i, M_j) = \Pi_{M_i} (]M_j, +\infty) \quad (74)$$

Elle a les propriétés suivantes :

. Réflexivité $\mu_{PS}(M_i, M_i) = 1$

. $\max(\mu_{PS}(M_i, M_j), \mu_{PS}(M_j, M_i)) = 1$ (75)

On remarque qu'en vertu de (69) la relation floue NSS de nécessité de surclassement strict est la complémentaire de PS. i.e. $PS = \overline{NSS}$. NSS est anti-réflexive et (75) indique que NSS est antisymétrique parfaite (Zadeh, /50/).

PS n'est pas transitive ; néanmoins Shimura (/42/) a proposé une méthode pour construire un ordre total à partir d'une relation satisfaisant (75). Un degré de surclassement global est construit, qui est le degré minimum de surclassement de M_i sur les autres M_j $j \neq i$ soit

$$d_{PS}(M_i) = \min_{j \neq i} \mu_{PS}(M_i, M_j)$$

On peut montrer que (Dubois et Prade, /20/)

$$d_{PS}(M_i) = PS(M_i)$$

c'est-à-dire que d_{PS} n'est autre que le degré de possibilité de surclassement global obtenu en comparant M_i et $\max_{i \neq j} M_j$

. Relation de nécessité de surclassement

On note NS la relation floue définie par

$$\mu_{NS}(M_i, M_j) = \int_{M_i}^{(M_j, +\infty)} (76)$$

On vérifie que, sauf situation pathologique

$$\mu_{NS}(M_i, M_j) + \mu_{NS}(M_j, M_i) = 1 \quad (77)$$

Les relations satisfaisant (77) sont connues dans la littérature sous le nom de relation de tournoi. Les liens étroits qu'entretient NS avec l'inclusion ensembliste au sens de l'opérateur \prec (chap. 2 § 3.5.1) suggèrent une approche naturelle pour déduire un classement de NS.

L'ensemble flou des éléments surclassés par M_i est bien sûr $S(i)$ tel que

$$\mu_{S(i)}(M_j) = \mu_{NS}(M_i, M_j), \quad \forall j$$

Pour classer les M_i , on va chercher les relations d'inclusion entre les ensembles flous $S(i)$ au sens de \prec . On rappelle que $F \prec G \leftrightarrow \forall u, \mu_F(u) \leq \frac{1}{2}$ ou $\mu_B(u) > \frac{1}{2}$, et que \prec est transitive. Cette méthode de classification est déjà suggérée dans Dubois et Prade (/16/). Elle échoue s'il y a des incohérences telles que

$$\exists i, j \quad S(i) \prec S(j) \quad \text{et} \quad S(j) \prec S(i)$$

sont faux en même temps. De telles incohérences n'arrivent généralement pas avec la relation NS puisque (Dubois et Prade /20/) :

Proposition 4.8 : si les intervalles flous $M_1 \dots M_n$ ont des fonctions d'appartenance continues alors il existe au moins une permutation σ de $\{1, 2, \dots, n\}$ telle que

$$S(\sigma_n) \prec S(\sigma_{n-1}) \dots \prec S(\sigma_1)$$

Ceci permet de classer les intervalles flous dans l'ordre $M_{\sigma_1} \dots M_{\sigma_n}$. On obtient en fait un ordre partiel.

. Relation de possibilité de surclassement strict

La relation floue bâtie sur l'indice $\Pi_M(\mathbb{N}, +\infty)$ a les mêmes propriétés que NS, et peut donc être traitée comme elle.

4.4.4 - Lien avec la compatibilité de deux quantités floues

On a défini au chapitre 3 (§ 6.3) la compatibilité d'un ensemble flou F par rapport à un ensemble flou G . C'est une quantité floue $\tau(F/G)$ à support sur $[0, 1]$. Elle vérifie les propriétés suivantes (Prade, /40/)

$$\tau(\bar{F}/G) = 1 \ominus \tau(F/G)$$

$$\tau(F_1 \times F_2 / G_1 \times G_2) = \hat{\min}(\tau(F_1/G_1), \tau(F_2/G_2))$$

$$\tau(F_1 + F_2 / G_1 \times G_2) = \hat{\max}(\tau(F_1/G_1), \tau(F_2/G_2))$$

Quand Q est une quantité floue continue, on vérifie que

$$\tau(Q/Q) = V \quad \text{tel que} \quad \mu_V(u) = u, \quad u \in [0,1]$$

$$\tau(\bar{Q}/Q) = \varphi \quad \text{tel que} \quad \mu_\varphi(u) = 1-u, \quad u \in [0,1]$$

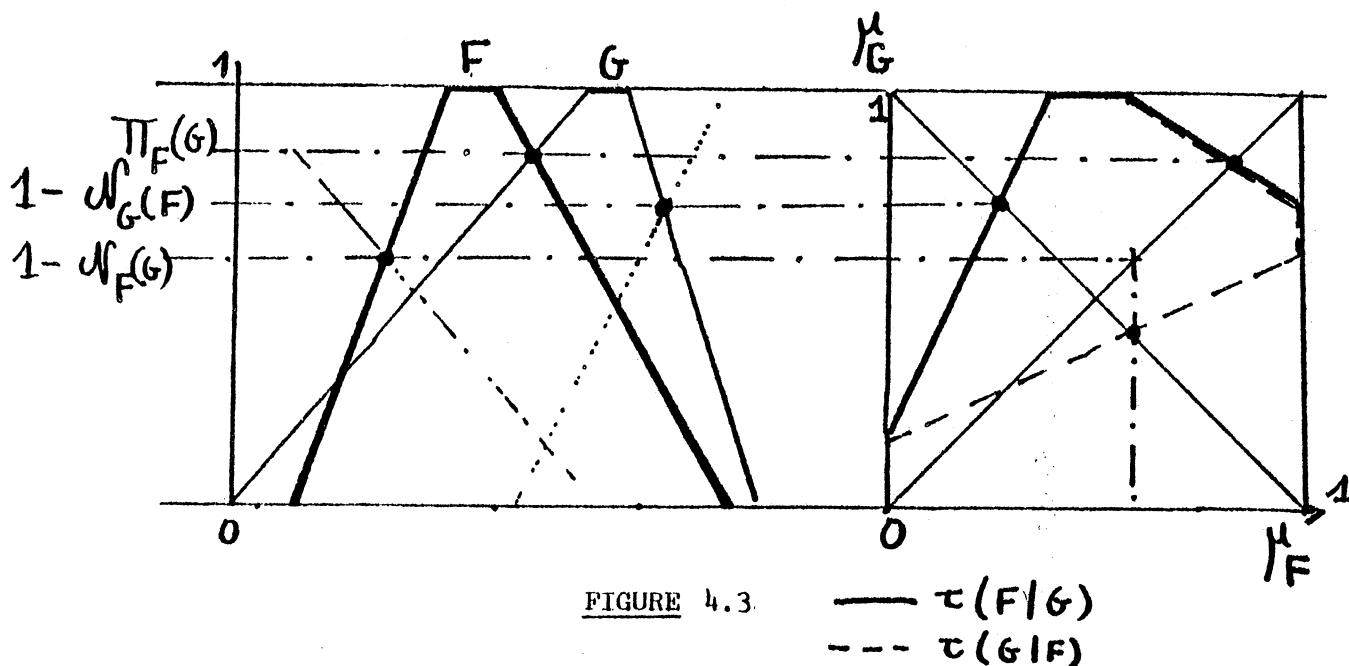
donc les catégories "vrai" et "faux" peuvent être représentées par les quantités floues V et φ respectivement, dans l'optique où $\tau(F/G)$ modélise la vérité de la proposition "X est F" par rapport à "X est G".

Si Π_G et \mathcal{N}_G sont les fonctions de possibilité et de nécessité construites sur μ_G , on a les deux propriétés suivantes

$$\cdot \quad \Pi_G(F) = \Pi_F(G) = \Pi_V(\tau(F/G)) = \Pi_V(\tau(G/F)) \quad (\text{Baldwin et Pilsworth, /6/})$$

$$\cdot \quad \mathcal{N}_G(F) = 1 - \Pi_\varphi(\tau(F/G)) \quad (\text{Prade, /40/})$$

D'où la figure ci-dessous :



On en déduit les relations suivantes entre les quatre indices de comparaison de quantités floues M et N , et les deux valeurs floues de vérité $\tau(M/N)$ et $\tau(N/M)$: (Dubois et Prade /20/)

$$\Pi_N(M) = \Pi_M(N) = \Pi_V(\tau(M/N)) = \min(\Pi_N([M, +\infty)), 1 - \mathcal{N}_N([M, +\infty))) \quad (78)$$

$$1 - \mathcal{N}_N(M) = \Pi_\varphi(\tau(M/N)) = \max(1 - \mathcal{N}_N([M, +\infty)), \Pi_N([M, +\infty))) \quad (79)$$

$$\mathcal{N}_M(N) = 1 - \Pi_\varphi(\tau(N/M)) = \min(\Pi_N([M, +\infty)), 1 - \mathcal{N}_N([M, +\infty))) \quad (80)$$

En remarquant que dans (78), l'une des deux quantités apparaissant en argument du minimum vaut 1, on voit que toute l'information fournie dans les 4 indices est contenue dans les deux valeurs de compatibilité $\tau(M/N)$ et $\tau(N/M)$.

NB1 : La difficulté du calcul des indices introduits dépend de la représentation adoptée pour les quantités floues. Basiquement il s'agit de trouver les coordonnées de l'intersection de deux courbes. Si on se restreint à des fonctions d'appartenance linéaires ou paraboliques, on peut même donner des formules analytiques.

NB2 : On peut bien sûr vouloir sortir du cadre de la théorie "max-min" des possibilités pour comparer deux intervalles flous. On peut d'abord introduire des normes triangulaires pour définir la possibilité d'un événement flou. On peut aussi vouloir mesurer le chevauchement de M et N en mesurant la surface limitée par $M \wedge N$. Ce type d'évaluation est de nature probabiliste ; voir Willmott /46/ pour une étude systématique de ces indices. Le cadre proposé au chapitre 2 (§ 4.2) contient les deux approches, possibiliste et probabiliste, et peut en suggérer d'autres.

4.4.5 - Domaines délimités par des bornes floues

On a vu qu'une quantité floue M mène à généraliser les demi-droites ouvertes et fermées. Dans le même esprit, la donnée de deux quantités floues M et N, peut donner l'occasion d'étendre la notion d'intervalle ouvert et d'intervalle fermé délimités par M et N, si on pose

$$[M, N] = [M, +\infty) \wedge (-\infty, N] \quad (81)$$

$$]M, N[=]M, +\infty) \wedge (-\infty, N[\quad (82)$$

Ces quantités sont telles que

$$\mu_{[M, N]}(w) = \sup \{ \min(\mu_M(u), \mu_N(v)) / u \leq w \leq v \}$$

$$\mu_{]M, N[}(w) = \inf \{ \max(1 - \mu_M(u), 1 - \mu_N(v)) / u \geq w \geq v \}$$

Leur étude systématique a été menée dans Dubois et Prade (/18/, /24/). Les propriétés suivantes ont été vérifiées quand M et N sont des intervalles flous s.c.s.

- . le support de $]M,N[$ est contenu dans le noyau de $[M,N]$ (quand ce dernier n'est pas vide.
- . le noyau de $]M,N[$ n'est pas vide si et seulement si $M \wedge N = \emptyset$
- . si on définit la frontière du domaine flou délimité par M et N par

$$\Phi(M,N) = [M,N] \wedge \overline{]M,N[}$$

alors

$$\Phi(M,N) = M \cup N \leftrightarrow \overset{\sim}{\max}(M,N) = N \leftrightarrow \overset{\sim}{\min}(M,N) = M$$

On voit qu'un tel domaine à frontière floue revient à la donnée de deux ensembles flous I et F tels que $I \overset{\circ}{\subseteq} F$, c'est-à-dire fortement emboîtés. I est l'intérieur, et F la fermeture du domaine. Dans le cas de quantités floues on a $I =]M,N[$ et $F = [M,N]$ comme exemple d'un tel concept. Dubois et Prade /18/ donnent des conditions pour construire M et N à partir de I et F, ils étudient aussi les propriétés des paires d'ensembles flous (qui généralisent les ensembles flous de Gentillomme /25/ - voir chap.2 § 1), vues comme des "ensembles à frontière floue".

L'intérêt des quantités $[M,N]$ et $]M,N[$ est de fournir une approche canonique au problème de localisation d'un intervalle flou P entre M et N. On peut ainsi évaluer à quel point il est possible que P se trouve strictement entre M et N (e.g. $\Pi_P(]M,N[)$), il est nécessaire que P soit entre M et N (e.g. $\mathcal{N}_P([M,N])$) etc... (voir Dubois et Prade /24/).

4.5 - ARITHMETIQUE FLOUE OPTIMISTE

Une des principales caractéristiques du calcul des intervalles flous c'est l'absence d'inverse, c'est-à-dire que si une opération sur les réels est un groupe, alors son extension aux intervalles flous n'est plus qu'un semi-groupe. Cet état de choses a un impact direct sur la possibilité de résoudre des équations dans le cadre de l'arithmétique floue, qui peut donc être contestée à cause de cela (Yager, /47/). Dans cette section on montre que le principe d'extension de Zadeh, pessimiste, au vu de son interprétation, possède un dual optimiste, directement lié au problème des équations d'intervalles flous. Ce sujet a été exploré récemment et indépendamment par Sanchez /41/ et Dubois et Prade /23/. Ce qu'on présente ci-après emprunte à ces deux références.

4.5.1 - Addition et soustraction de Minkowski

L'addition de deux ensembles A et B, eux-mêmes sous-ensembles d'un référentiel Ω a été définie au début du siècle par Minkowski (cf. Matheron, /31/).

$$A \oplus B = \{w / \exists u \in A, \exists v \in B, w = u + v\} \quad (83)$$

On peut réécrire (83) en faisant apparaître l'ensemble $a \ominus B = \{a-b/b \in B\}$ comme

$$A \oplus B = \{w / (w \ominus B) \cap A \neq \emptyset\} \quad (84)$$

(84) peut faire apparaître, comme dual naturel de (83), l'ensemble

$$A \succ (B = \{w / w \ominus B \subseteq A\} \quad (85)$$

\succ est appelé "soustraction de Minkowski".

Au sens de la théorie des possibilités, $A \oplus B$ est l'ensemble des valeurs possibles de $A \oplus v$, quand $v \in B$, et $A \succ (B$ est l'ensemble des valeurs que recouvre nécessairement $A \oplus v$ quand $v \in B$; (i.e. $\bigcap \{A \oplus v / v \in B\}$)

On a $\overline{A \succ (B} = \overline{A \oplus B}$, c'est-à-dire par contraste avec (83)

$$A \succ (B = \{w / \forall v \in B, \exists u \in A, u + v = w\} \quad (86)$$

En particulier $A \succ (B \neq B \succ (A$

Si

$$A = [a, a'] , \quad B = [b, b']$$

$$A \oplus B = [a+b, a'+b']$$

$$A \succ (B = [a+b', a'+b] \quad \text{si} \quad a+b' \leq a'+b$$

$$= \emptyset \quad \text{sinon}$$

On a toujours $A \succ (B \subseteq A \oplus B$. Si A et B sont les intervalles définis ci-dessus et $l(A)$ est la longueur de l'intervalle A

$$l(A \oplus B) = l(A) + l(B)$$

$$l(A \succ (B) = l(A) - l(B)$$

Donc $A \succ (B$ est plus précis que $A \oplus B$. Par exemple $A \succ (A = \{a+a'\}$. Si A et B sont des nombres réels, les deux opérations \oplus et $\succ ($ coïncident avec l'addition

usuelle. On peut vérifier les inclusions

$$(A) \dot{+} (B) \ominus B \subseteq A \subseteq (A \oplus B) \dot{+} (-B) \quad (87)$$

Les inégalités sont des égalités quand A et B sont des intervalles fermés.

En particulier $(A) \dot{+} (-A) = \{0\}$

Donc il est clair que l'opérateur $\dot{+}$ peut permettre de résoudre l'équation

$$A \oplus X = B$$

sous la forme : $X = B \dot{+} (-A)$ quand A et B sont des intervalles fermés.

De même que la formule (83) peut être étendue à n'importe quelle opération f à deux arguments pour construire $f(A,B)$, (86) peut aussi être utilisé pour construire l'ensemble

$$\underline{f}(A,B) = \{w / \forall v \in B, \exists u \in A, f(u,v) = w\} \quad (88)$$

Proposition 4.9 : si $f(u,v') = f(u,v) \rightarrow v = v'$

alors

$$\underline{f}(A,B) = \overline{f(\bar{A},B)} \quad (89)$$

si de plus f est isotone et continue, et si $A = [a, a']$, $B = [b, b']$

alors

$$\begin{aligned} \underline{f}(A,B) &= [f(a, b'), f(a', b)] \quad \text{si } f(a, b') \leq f(a', b) \\ &= \emptyset \quad \text{sinon} \end{aligned}$$

Preuve : soit $w \in \underline{f}(A,B)$. Supposons $\exists (u,v) \in \bar{A} \times B$, $f(u,v) = w$.

Alors on voit aussi que $\exists u' \in A, f(u', v) = f(u,v)$ et donc $u = u'$ ce qui

est une contradiction. La réciproque est similaire. Le reste de la proposition

est évident en utilisant (89).

Q.E.D.

Notons f^* la fonction qui à (u,w) fasse correspondre l'ensemble des v tels que $f(u,v) = w$, et considérons

$$\begin{aligned} \underline{f}^*(A,B) &= \{v / \forall u \in A, \exists w \in B, f(u,v) = w\} \\ &= \{v / f(A,v) \subseteq B\} \end{aligned} \quad (90)$$

$\underline{f}^*(A,B)$ est, quand elle existe, solution de l'équation $f(A,X) = B$, où A et B sont des intervalles fermés et f satisfait les hypothèses de la proposition 4.9

4.5.2 - Equations d'intervalles flous

Soit M et N deux ensembles flous de \mathbb{R} , et considérons l'équation

$$f(M, X) = N \quad (91)$$

où X est un ensemble flou à trouver. $f(M, X)$ est obtenu par application du principe d'extension, au sens d'une t-norme continue \perp (cf. § 3.2).

Si X_0 est solution de (91) alors son enveloppe convexe (cf. § 2.7) \hat{X}_0 satisfait $f(M, \hat{X}_0) \supseteq N$ à cause de la monotonie de f au sens de l'inclusion. Supposons que f soit isotone continue et que l'équation $f(a, x) = b$ ait une solution unique notée $f^*(a, b)$. Dans ce cas on montre

$$f(M, \hat{X}_0) \subseteq \hat{N}$$

La démonstration, donnée pour $f(a, b) = a+b$ par Dubois et Prade /23/ s'étend sans problème aux hypothèses ci-dessus.

Dans ce cas on voit que si X_0 est une solution de (91) et N est convexe, \hat{X}_0 est aussi solution de (91).

Puisque pour n'importe quelle t-norme, et $U = \max$

$$f(M, X_1 \cup X_2) = f(M, X_1) \cup f(M, X_2) \quad (92)$$

on en déduit que si X_1 et X_2 sont solutions de (91) alors $X_1 \cup X_2$ est aussi solution de (91). Donc si (91) a une solution, elle a une plus grande solution au sens de l'inclusion de Zadeh.

Donc l'ensemble des solutions de (91) forme un demi-treillis supérieur avec un élément dominant.

4.5.3 - Solution des équations d'intervalles flous

L'équation (91) se développe en : trouver X tel que

$$\begin{aligned} \forall w, \mu_N(w) &= \sup_{u, x} \{ \mu_M(u) \perp \mu_X(x) / f(u, x) = w \} \\ &= \sup_x \{ \mu_{f(M, x)}(w) \perp \mu_X(x) \} \end{aligned} \quad (93)$$

puisque

$$\mu_{f(M,x)}(w) = \sup \{ \mu_M(u) / f(u,x) = w \}$$

Soit R la relation floue telle que

$$\mu_R(x,w) = \mu_{f(M,x)}(w)$$

On voit que (91) se ramène à une équation de relation floue du type

$$N = X \Phi R \tag{94}$$

Φ désigne la composition sup- \perp de relations floues (chap.3 § 8.2). Ces équations ont été résolues par Sanchez /41/ pour $\perp = \min$ et Pedrycz /39/ pour les t-normes continues. Si une solution de (94) existe alors la plus grande solution est

$$\mu_{\hat{X}}(x) = \inf_w \mu_R(x,w) \mathcal{Q} \mu_N(w) \tag{95}$$

où l'opérateur \mathcal{Q} est une implication multivalente déjà rencontrée au chap.2 § 3.5.2 ; \mathcal{Q} est déduite de la t-norme \perp par

$$u \mathcal{Q} v = \sup \{ w / u \perp w \leq v \}$$

NB : Dans le cas où $\perp = \max(0, a+b-1)$ l'équation (94) est très proche d'une équation résolue par Cuninghame-Green /7/ : pour cet auteur μ_X, μ_R, μ_N sont des fonctions à valeurs dans \mathbb{R} , $\perp = \text{somme}$, et $\mathcal{Q} = \text{différence}$.

L'application des équations de relations floues à la résolution des équations en arithmétique floue a été entreprise par Sanchez /41/ pour $\perp = \min$, Pedrycz /39/ et Dubois-Prade /23/ pour l'addition et une t-norme continue.

(95) peut-être explicitée comme

$$\begin{aligned} \mu_{\hat{X}}(x) &= \inf_{w,u} (\sup \{ \mu_M(u) / f(u,x) = w \} \mathcal{Q} \mu_N(w)) \\ &= \inf_u \{ \mu_M(u) \mathcal{Q} \mu_N(f(u,x)) \} \end{aligned} \tag{96}$$

en utilisant $(\sup_i a_i) \underline{d} b = \inf_i (a_i \underline{d} b)$, si \underline{d} est continue.

Si M et N sont des ensembles au sens usuel, on sait que $\mu_{\hat{X}}(x) = 1$ dans (95) veut dire $f(M, x) \subseteq N$. Donc (96) généralise (90) à des ensembles flous, i.e. si l'équation a une solution

$$\hat{X} = \underline{f}^*(M, N) \rightarrow f(M, \hat{X}) = N$$

On notera que $\hat{X} = \underline{f}^*(M, N)$ est la plus grande solution de l'équation $f(M, X) \subseteq N$. Si aucune solution de $f(M, X) = N$ n'existe, alors $f(M, \hat{X}) \subset N$. Eventuellement $\hat{X} = \emptyset$.

4.5.4 - Equation $M \oplus X = N$

On suppose ici que M et N sont des intervalles flous s.c.s. .

Dans le cas de l'addition, la plus grande solution de l'équation $M \oplus X = N$ s'écrit, quand elle existe, $\hat{X} = N \oplus (-M)$, tel que

$$\mu_{\hat{X}}(x) = \inf_u \mu_M(u) \underline{d} \mu_N(u+x) \quad (97)$$

où M et N sont des intervalles flous s.c.s.. On rappelle que \underline{d} est tel que

$$a \underline{d} b = 1 \leftrightarrow a \leq b$$

quel que soit la t -norme dont il est issu.

Une condition nécessaire d'existence de X solution de $M \oplus X = N$ est

$$\exists x, \forall u \mu_M(u) \leq \mu_N(u+x) \quad (98)$$

sinon \hat{X} ne serait pas normalisé ; or puisque N et M le sont, toute solution de $M \oplus X = N$ doit l'être aussi.

(98) signifie que M est plus précis que N , puisque par translation on peut ramener M à l'intérieur de N . Quand \underline{d} dérive du minimum, la condition (98) est nécessaire et suffisante pour que

$$\forall \alpha \in]0, 1], \exists \hat{X}_\alpha, M_\alpha \oplus \hat{X}_\alpha = N_\alpha$$

En effet on vérifie d'après 4.5.1, $\hat{X}_\alpha = N_\alpha \oplus (-M_\alpha) = \left[\frac{n_\alpha - \bar{m}_\alpha}{\alpha}, \frac{\bar{n}_\alpha - \bar{m}_\alpha}{\alpha} \right]$,
 Si $N_\alpha = \left[\frac{n_\alpha}{\alpha}, \frac{\bar{n}_\alpha}{\alpha} \right]$, $M_\alpha = \left[\frac{m_\alpha}{\alpha}, \frac{\bar{m}_\alpha}{\alpha} \right]$.

\hat{X}_α est non vide si et seulement si $\frac{n_\alpha - \bar{m}_\alpha}{\alpha} > \frac{m_\alpha - \bar{m}_\alpha}{\alpha}$, i.e. $1(M_\alpha) \leq 1(N_\alpha)$. Or il est clair que (98) $\leftrightarrow 1(M_\alpha) \leq 1(N_\alpha)$, $\forall \alpha$

Pour pouvoir construire une solution \hat{X} à partir des \hat{X}_α il faut qu'ils soient α -coupes. En particulier

$$\alpha < \beta \rightarrow N_\alpha \oplus (-M_\alpha) \supseteq N_\beta \oplus (-M_\beta)$$

En conséquence l'existence de la quantité $N \uparrow (-M$ ne garantit pas celle de X tel que $M \oplus X = N$, puisque $N \uparrow (-M$ peut exister sans que (98) soit vérifiée. En revanche $N \uparrow (-M = \emptyset$ assure l'absence de solution.

La plus grande solution de $M \oplus X = N$ peut être trouvée quand \oplus est une addition étendue au sens de $\underline{\downarrow} = \min$ en résolvant les deux équations déduites de (22)

$$(\mu_M^\varepsilon)^{-1}(\alpha) + x_\varepsilon(\alpha) = (\mu_N^\varepsilon)^{-1}(\alpha) \quad \varepsilon \in \{-, +\}$$

pourvu que $\forall \alpha < \beta \quad x_-(\alpha) \leq x_-(\beta) \leq x_+(\beta) \leq x_+(\alpha)$

Cette solution est unique quand elle existe.

Si $M = (m, \alpha, \beta)_{LR}$ et $N = (n, \gamma, \delta)_{LR}$ on voit facilement que

$$\hat{X} = N \uparrow (-M = (n-m, \gamma-\alpha, \beta-\delta)_{LR}$$

et que \hat{X} existe seulement si $\gamma - \alpha > 0, \beta - \delta > 0$.

$N \uparrow (-M$ peut encore s'écrire $N \downarrow (-M$ car c'est une soustraction généralisée.

Bien sûr $\hat{X} = N \downarrow (-M$ sera solution de $X \oplus M = N$ quand elle existe. Cette addition d'intervalles flous est bien "optimiste" puisque \hat{X} est toujours plus précis que N , car si N est de type RL et M de type LR alors $N \downarrow (-M = (m+n, \gamma-\alpha, \beta-\delta)_{RL}$, ce qui contraste avec les addition et soustraction \oplus et \ominus pour lesquelles les étalements s'ajoutent. $\downarrow (-$ correspond à la compensation maximale des erreurs.

On vérifie (Dubois et Prade/23/), que pour toute t-norme continue générant \square on a les propriétés suivantes (M, N, P sont des intervalles flous s.c.s.)

$$M \downarrow (N \cup P) = (M \downarrow N) \wedge (M \downarrow P)$$

$$(N \wedge P) \downarrow M = (N \downarrow M) \wedge (P \downarrow M)$$

$$M \downarrow (N \wedge P) \supseteq (M \downarrow N) \cup (M \downarrow P)$$

$$(N \cup P) \downarrow M \supseteq (N \downarrow M) \cup (P \downarrow M)$$

Si $N \subseteq P$ alors

$$M \downarrow N \supseteq M \downarrow P$$

$$N \downarrow M \subseteq P \downarrow M$$

Ces propriétés sont valables pour n'importe quelle opération à la place de +; l'union et l'intersection sont au sens du minimum. Le calcul de (97) quand \mathcal{L} est généré par une t-norme autre que le minimum est plus difficile. Par exemple si \mathcal{L} = produit alors

$$\mu_X(x) = \inf_u \min(1, \frac{\mu_N(u+x)}{\mu_M(u)})$$

Néanmoins, si M et N sont de type L-R on peut résoudre l'équation $M \otimes X = N$ dans les cas où on sait calculer $M \otimes X$ (cf. 4.3.2). Par exemple si \otimes est basée sur T_w on a

$$\begin{aligned} M \otimes X &= (m + \bar{x}, \max(\alpha, \alpha_x), \max(\beta, \beta_x))_{LR} \\ &= (n, \gamma, \delta)_{LR} = N \end{aligned}$$

où X est cherché sous la forme $(\bar{x}, \alpha_x, \beta_x)_{LR}$
on en déduit $\bar{x} = n - m$

$$\alpha_x = \begin{cases} \gamma & \text{si } \alpha < \gamma \\ [0, \gamma] & \text{si } \alpha = \gamma \end{cases} \quad \beta_x = \begin{cases} \delta & \text{si } \beta < \delta \\ [0, \delta] & \text{si } \beta = \delta \end{cases}$$

sinon X n'existe pas. On retrouve la condition que M doit être plus précis que N.

4.5.5 - Equation $M \otimes X = N$

On suppose dans cette section que \otimes est le produit au sens du principe d'extension sup-min. M et N sont des intervalles flous s.c.s. positifs. La plus grande solution de $M \otimes X = N$ quand l'équation a une solution définit une division optimiste notée \mathcal{L} , telle que

$$\mu_N \mathcal{L} (v) = \inf_{u > 0} \mu_M(u) \mathcal{L} \mu_N(u.v)$$

où

$$\begin{aligned} a \mathcal{L} b &= 1 \quad \text{si } a \leq b \\ &= b \quad \text{si } a > b \end{aligned}$$

On a encore la propriété suivante :

$$\exists X : M \circ X = N \rightarrow \exists v, \forall u, \mu_M(u) \leq \mu_N(u, v) \quad (99)$$

Pour le montrer, on utilise la même démonstration que pour l'addition.

On a d'ailleurs plus généralement :

Proposition 4.10 : Si M et N sont des quantités flous normalisées,

$$\exists X, f(M, X) = N \rightarrow \exists v, \forall u, \mu_M(u) \leq \mu_N(f(u, v))$$

Preuve : si X existe alors puisque M et N sont normalisés, X doit l'être aussi et on a, de (96)

$$\exists v : \mu_M(u) \leq \mu_N(f(u, v)) = 1, \forall u ; \quad \text{d'où la condition.}$$

(on note que la condition est nécessaire $\forall v$).

Si f est isotone, continue et telle que f(., v) est une bijection de son domaine sur lui-même alors si M et N sont des intervalles flous s.c.s., on sait que $f(M_\alpha, X_\alpha) = N_\alpha \forall \alpha \in]0, 1[$ de par la proposition 4.1. Donc on peut construire une solution X à partir de ses α -coupes. On est donc ramené, avec les notations de la section précédente aux équations

$$\begin{aligned} f(\underline{m}_\alpha, \underline{x}_\alpha) &= \underline{n}_\alpha \\ f(\bar{m}_\alpha, \bar{x}_\alpha) &= \bar{n}_\alpha \end{aligned} \quad (100)$$

il est clair qu'une solution X_α de $f(M_\alpha, X_\alpha) = N_\alpha$ existe si et seulement si la solution de (100) (unique si f(u, .) est bijective) satisfait $\underline{x}_\alpha \leq \bar{x}_\alpha$.

Soit v tel que $\mu_M(u) \leq \mu_N(f(u, v)) \forall u$; on notera que $\mu_N \circ f$ est la fonction d'appartenance de $f^*(\cdot, N)$. On a donc

$$\forall \alpha, M_\alpha \subseteq \{u / f(u, v) \in N_\alpha\} \triangleq [f^*(\underline{n}_\alpha, v), f^*(\bar{n}_\alpha, v)]$$

avec $N_\alpha = [\underline{n}_\alpha, \bar{n}_\alpha]$ et $f^*(\cdot, v)$ est l'inverse de f(., v)

Donc $\forall \alpha, [f(\underline{m}_\alpha, v), f(\bar{m}_\alpha, v)] \subseteq [\underline{n}_\alpha, \bar{n}_\alpha]$

donc \bar{x}_α et \underline{x}_α , solutions de (100) sont, $\forall \alpha$, telles que

$$\underline{x}_\alpha \leq v \leq \bar{x}_\alpha$$

Pour construire \hat{X} telle que $f(M, \hat{X}) = N$ à partir de \underline{x}_α et \bar{x}_α il faut de plus que les intervalles $\{[\underline{x}_\alpha, \bar{x}_\alpha], \alpha \in]0, 1[\}$ forment un ensemble d' α -coupes.

(99) découle donc de la proposition 4.10, car le produit xy est bien une bijection de $[0, +\infty)$ sur lui-même. On notera que (100) n'est autre que l'équation (22), dont la solution directe permet de résoudre l'équation $M \circ X = N$.

Par exemple si $M = (M, \alpha, \beta)_{LR}$ et $N = (n, \gamma, \delta)_{LR}$ alors on doit résoudre

$$(m - \alpha L^{-1}(\lambda)) \underline{x}_\lambda = (n - \gamma L^{-1}(\lambda)), \quad \underline{x}_\lambda \leq \frac{n}{m}$$

$$(m + \beta L^{-1}(\lambda)) \bar{x}_\lambda = (n + \delta L^{-1}(\lambda)), \quad \bar{x}_\lambda \geq \frac{n}{m}$$

soit

$$\begin{aligned} \mu_N : (M(x) &= L\left(\frac{mx-n}{\alpha x-\gamma}\right) \quad 0 \leq x \leq \frac{n}{m} \\ &= R\left(\frac{n-mx}{\beta x-\delta}\right) \quad x \geq \frac{n}{m} \end{aligned}$$

On obtient ainsi la plus grande solution de $M \circ X = N$ si (99) est vérifié ; on montre facilement que

$$(99) \leftrightarrow \frac{\beta}{\delta} > \frac{m}{n} > \frac{\alpha}{\gamma}$$

4.5.6 - Autres équations d'intervalles flous

A partir des résultats précédents on peut résoudre des équations plus compliquées, par exemple une équation du second degré de la forme $f(M, X) = N$ où $f(u, v) = v^2 + uv$. On n'aura aucun problème pour appliquer la proposition 4.1 quand par exemple M et N sont des intervalles flous positifs. Si ça n'est pas le cas on peut être amené à partitionner les intervalles flous pour travailler sur des domaines où f est monotone, puis regrouper les résultats à l'aide des formules de distributivité données au § 5.4.

Pedrycz (/39/) a fait remarquer que l'on peut étendre la solution générale d'une équation d'intervalles flous (95) à des systèmes d'équations floues. Par exemple considérons le système, d'inconnues X et Y

$$\begin{aligned} M_1 \circ X \oplus N_1 \circ Y &= P_1 \\ M_2 \circ X \oplus N_2 \circ Y &= P_2 \end{aligned} \tag{101}$$

La i ème équation revient à l'équation fonctionnelle en μ_X, μ_Y

$$\sup \{ \mu_{M_i}(u) \wedge \mu_{N_i}(v) \wedge \mu_X(x) \wedge \mu_Y(y) / ux+vy = w \} = \mu_{P_i}(w)$$

Si R_i est une relation floue sur \mathbb{R}^4 telle que

$$\mu_{R_i}(v, w, x, y) = \mu_{M_i}\left(\frac{w-vy}{x}\right) \wedge \mu_{N_i}(v)$$

on est ramené à l'équation de relations floues

$$P_i = R_i \circ X Y_i \quad (102)$$

où XY_i est une relation floue sur \mathbb{R}^2 . Pedrycz suggère comme méthode de résolution:

. $\forall i$ résoudre (102) : XY_i est obtenu par composition inf- \wedge de R_i par P_i (cf 95).

. $XY = XY_1 \wedge XY_2$ on réalise l'intersection des solutions obtenues pour chaque contrainte.

Des méthodes pratiques de résolution, basées sur des résultats du type de la proposition 4.11, peuvent sans doute être développées pour ramener (101) à un système d'équations usuelles notamment quand M_i , N_i et P_i sont des intervalles flous de type L-R.

On peut aussi considérer des équations où l'inconnue apparaît dans les deux membres par exemple trouver X tel que

$$MX \oplus M' = NX \oplus N'$$

Il est clair que si $N = 1$, $N' = 0$, l'équation ci-dessus n'a aucune solution. On voit facilement que l'équation ci-dessus se ramène à une équation de relation floue du type

$$R \circ X = R' \circ X$$

qui n'a, jusqu'ici pas été étudié dans la littérature.

4.5.7 - Autres types d'équations floues

Les équations d'intervalles flous définies ci-dessus ont un intérêt théorique dans la mesure où elles sont l'occasion de définir des opérations "inverses" pour les opérations entre intervalles flous : par exemple $M \ominus M \neq 0$ mais $M) - (M = 0$, etc... . La solution d'une telle équation floue fournit

la plage de variation de l'inconnue, de façon à ce que l'incertitude sur l'un des membres d'une égalité ne dépasse pas celle de l'autre membre.

On notera qu'une autre généralisation possible de l'équation $f(u,x) = v$ est $f(M,X) \leq N$ où l'inclusion est au sens de Zadeh. De telles équations floues sont traitées de la même façon que les équations de type $f(M,X) = N$. Cette fois, l'ensemble des solutions forme un treillis, et il n'est jamais vide (\emptyset fait partie des solutions !). La plus grande solution est donnée par (95) (voir Sanchez (/41/)).

Néanmoins, il est des situations où de telles formulations s'avèrent trop rigides. Compte tenu de la nature d'une fonction d'appartenance, réclamer l'égalité de deux telles fonctions peut mener à des impossibilités (l'équation $A \circ X = B$ n'a pas toujours de solution) alors qu'une égalité approchée fournirait une solution.

Un concept d'égalité approchée peut être obtenu à l'aide d'indices de comparaison entre les ensembles flous qui définissent les deux membres d'une égalité.

Conformément à la section 4 (eq. 78-80) l'égalité de M et N peut s'étudier à l'aide de trois quantités : $\Pi_M(N)$, $\mathcal{N}_M(N)$, $\mathcal{N}_N(M)$.

. La première évalue la possibilité d'égalité entre M et N (hauteur de $M \cap N$)
 . Les deux autres évaluent respectivement l'inclusion (au sens affaibli de \leftarrow , cf. chap.2 §?) de M dans N et de N dans M. La nécessité d'égalité peut donc être évaluée par $\min(\mathcal{N}_M(N), \mathcal{N}_N(M))$.

Donc une version très affaiblie d'une équation floue $f(M,X) = N$ peut être obtenue en se fixant un seuil π de possibilité et en imposant

$$\Pi_N(f(M,X)) = \sup_u \min(\mu_{f(M,X)}(u), \mu_N(u)) \geq \pi \quad (103)$$

Si même avec $\pi = 1$, la solution X s'avère trop imprécise, on peut alors imposer un seuil ν tel que

$$\mathcal{N}_{f(M,X)}(N) = \inf_u \max(1 - \mu_{f(M,X)}(u), \mu_N(u)) \geq \nu \quad (104)$$

Pour simplifier le problème, en pratique, on peut se contenter de chercher X sous la forme d'un intervalle au sens usuel, c'est-à-dire chercher x tel que $f(M,x)$ et N soient approximativement égaux au sens de (103) ou (104). Si on veut résoudre un ensemble d'équations à coefficients imprécis, on sera alors ramené à un ensemble d'inégalités de type (103) ou (104). On pourra alors, en introduisant une variable supplémentaire (qui sera le degré minimum de satisfaction des équations) transformer un système d'égalités approchées en un problème d'optimisation. C'est le propos d'une des sections du chapitre 7.

NB : Ce type d'égalité approchée, tel qu'il est décrit en (103) et (104) peut être utile pour définir des notions affaiblies de convergence pour des suites de nombres flous. Si $\{M_n\}_n$ est une suite de nombres flous, on peut penser à de nombreuses façons de la voir converger, plus satisfaisantes intuitivement que la convergence point par point des fonctions d'appartenance. On peut définir une convergence α -possible vers le nombre flou M si

$$\exists \alpha, \exists n_0 \forall n \geq n_0 \quad \Pi_{M_n}(M) \geq \alpha$$

Une convergence α -nécessaire vers M serait

$$\exists \alpha, \exists n_0 \forall n \geq n_0 \quad \min(\hat{\mathcal{A}}_{M_n}(M_n), \hat{\mathcal{A}}_{M_n}(M)) \geq \alpha$$

Un exposé préliminaire sur cette question, et l'illustration de l'intérêt de ces notions sur le problème de la stabilité d'un système flou sont donnés dans Dubois et Prade (/19/). On montre qu'on peut généraliser des théorèmes tels que la limite d'une somme de deux suites est la somme des limites. On a par exemple : si $\{M_n\}_n$ est α -possiblement convergent vers M et $\{N_n\}_n$ est β -possiblement convergent vers N alors $\{M_n \oplus N_n\}_n$ est $\min(\alpha, \beta)$ -possiblement convergent vers $M \oplus N$.

REFERENCES

- /1/ ALSINA C. NGUYEN H.
A note on the addition of fuzzy variables.
Fuzzy Sets and Syst.
- /2/ BAAS S., KWAKERNAAK H. (1977)
Rating and ranking of multiple aspect alternatives using fuzzy sets. *Automatica*, B, 47-58
- /3/ BADARD R. (1981)
The law of large numbers for fuzzy processes and the estimation problem, Dept. D'Informatique INSA, Lyon
- /4/ BALDWIN J.F., GUILD N. (1979)
Comments on the fuzzy max operator of Dubois and Prade.
Int. J. Systems Sciences 10, 1063-1064
- /5/ BALDWIN J.F., GUILD N.C.F. (1979)
Comparison of fuzzy sets on the same decision space.
Fuzzy Sets and Systems 2, n°3, 213-233
- /6/ BALDWIN J.F., PILSWORTH B.W. (1979)
Fuzzy truth definition of possibility measure for decision classification. *Int. J. Man-Machine Studies* 11, 447-463
- /7/ CUNINGHAME-GREEN R.A. (1979)
Minimax Algebra, Lecture notes in Economics and Math. Syst. N°166, Springer Verlag.
- /8/ DISHKANT H. (1981)
About membership function estimation.
Fuzzy Sets and Systems 5, 141-147
- /9/ DUBOIS D. (1980)
Ensembles flous et Conception Assistée par Ordinateur :
Rappels théoriques, perspectives et exemple. Rapport de Recherche IMAG N°199, Grenoble

- /10/ DUBOIS D. (1980)
La théorie des nombres flous : panorama des résultats actuels.
Table ronde CNRS "Quelques applications utilisant les derniers perfectionnements de la théorie du flou, Université de Lyon.
- /11/ DUBOIS D. (1981)
Résultats d'arithmétique floue in "Détermination des conditions de coupe sur une chaîne automatisée d'usinage".
 Rapport CERT-DERA n°027 - Toulouse - pp.87-101.
- /12/ DUBOIS D. (1982)
A law of large numbers for non convex fuzzy sets of the real line BUSEFAL n°9, 31-38
- /13/ DUBOIS D., PRADE H. (1978)
Operations on fuzzy numbers, Int. J. Syst. Sci. 9, 613-626
- /14/ DUBOIS D., PRADE H. (1979)
Fuzzy real algebra : some results. Fuzzy Sets & Systems 2, 327-348.
- /15/ DUBOIS D., PRADE H. (1979)
Operations in a fuzzy-valued logic. Information and Control 43 224-240
- /16/ DUBOIS D., PRADE H. (1980)
Fuzzy Sets and Systems : Theory and Applications, Academic Press, New York.
- /17/ DUBOIS D., PRADE H. (1981)
Addition of Interactive fuzzy numbers, IEEE Trans. Automatic Control, 26, 926-936
- /18/ DUBOIS D., PRADE H. (1981)
Fuzzily-bounded domains : a topological point of view.
Proc. 3d Int. Seminar on Fuzzy set theory (E.R. Klement, Ed)
 J. Kepler Univ. Linz. Autriche, 191-209.
- /19/ DUBOIS D., PRADE H. (1982)
What does "convergence" mean for fuzzy numbers.
Présenté à IFAC Symposium on Theory and Applications of Digital Control New Delhi.

- /20/ DUBOIS D., PRADE H. (1982)
Ranking fuzzy numbers in the setting of possibility theory, Information Sciences (à paraître)
- /21/ DUBOIS D., PRADE H. (1982)
The use of fuzzy numbers in decision Analysis in "Fuzzy Information and Decision Processes" (M.M. Gupta, E. Sanchez Eds) North Holland 309-321.
- /22/ DUBOIS D., PRADE H. (1982)
Upper and lower possibilistic expectations and applications. 4th Int. Seminar on Fuzzy Set theory, Linz, Autriche.
- /23/ DUBOIS D., PRADE H. (1983)
Inverse operations for fuzzy numbers. IFAC Symposium on Fuzzy Information, Knowledge representation and Decision Analysis, Marseille.
- /24/ DUBOIS D., PRADE H. (1983)
Twofold fuzzy sets : an approach to the representation of sets with fuzzy boundaries, based on possibility and necessity measures. A paraître dans Fuzzy Mathematics (Chine)
- /25/ GENTILLOMME Y. (1968)
Les ensembles flous en linguistique. Cahiers de Linguistique Théorique et Appliquée, Bucarest, 38, 154-169.
- /26/ GONDRAN M. (1979)
Les éléments p-réguliers dans les dioïdes. Discrete. Math. 25, 33-39.
- /27/ GONDRAN M., MINOUX M. (1979)
Graphes et Algorithmes, Eyrolles.
- /28/ HÖHLE H. (1981)
Representation theorems for L-fuzzy quantities, Fuzzy Sets and Systems, 5, 83-107.
- /29/ KUZ'MIN V.B. (1981)
A parametric approach to description of linguistic values of variables and hedges, Fuzzy Sets and Syst. 6, 27-41, and 205
- /30/ LOWEN R. (1980)
Convex fuzzy sets, Fuzzy Sets & Syst. 3, 291-310

- /31/ MATHERON G. (1975)
Random sets and Integral Geometry, Wiley
- /32/ MIZUMOTO M., TANAKA K. (1976)
The 4 operations of arithmetic on fuzzy numbers.
Systems, Computers, Controls 7 n°5, 73-81
- /33/ MIZUMOTO P., TANAKA K. (1979)
Some properties of fuzzy numbers, in Advances in Fuzzy set theory and Applications (M.M. Gupta, R.K. Ragade, R.R. Yager eds) North Holland pp.153-164.
- /34/ MOORE R. (1966)
Interval analysis, Prentice Hall
- /35/ MOORE R. (1979)
Methods and Applications of Interval Analysis, SIAM Studies on applied mathematics, vol.2.
- /36/ NAHMIAS S. (1978)
Fuzzy variables, Fuzzy Sets & Systems 1, 97-111
- /37/ NAHMIAS S. (1979)
Fuzzy variables in a random environment, in Advances in Fuzzy Set Theory and Applications (M.M. Gupta, R.K. Ragade, R.R. Yager Eds) North Holland 165-180
- /38/ NGUYEN H.T. (1978)
A note on the extension principle for fuzzy sets. J. Math. Anal. Appl. 64, N°2, 369-380
- /39/ PEDRYCZ W. (1982)
Fuzzy relational equations with triangular norms and their resolutions, BUSEFAL n°11, 24-32.
- /40/ PRADE H. (1980)
Compatibilité - Qualification - Modification - Niveau de précision. BUSEFAL n°4, 71-78
- /41/ SANCHEZ E. (1982)
Solution of fuzzy equations with extended operations.
Memo UCB/ERL M 82-63, University of California, Berkeley
- /42/ SHIMURA M. (1973)
Fuzzy set concepts in the rank-ordering of objects : J. Math. Anal. and Appl. 43, 717-733.

- /43/ STEIN W. (1983)
A note on convexity and fuzzy random variables.
 pp.43-46 - BUSEFAL N°14
- /44/ STEIN W., TALATTI K. (1981)
Convex fuzzy random variables. Fuzzy Sets and Systems vol.6,
 pp.271-283.
- /45/ TSUKAMOTO Y. NIKIFORUK P.N., GUPTA M.M. (1981)
On the comparison of fuzzy sets using fuzzy chopping. Proc.
8th Triennial World Congress IFAC, Kyoto 5, 46-52.
- /46/ WILLMOTT R. (1981)
Mean measures of containment and equality between fuzzy sets.
Proc. 11th IEEE Int. Symp. on Multiple Valued Logic,
Oklahoma City, 183-190
- /47/ YAGER R.R. (1980)
On the lack of inverses in fuzzy arithmetic, Fuzzy Sets & Syst.
 4, 73-82.
- /48/ YOUNG R.C. (1931)
The algebra of many valued quantities, Math. Ann., 104, 260-290
- /49/ ZADEH L.A. (1965)
Fuzzy Sets, Information and Control, 8, 338-353.
- /50/ ZADEH L.A. (1971)
Similarity relations and fuzzy orderings, Information Sciences,
 3, 77-200
- /51/ ZADEH L.A. (1975)
*The concept of a linguistic variable and its application to
 approximate reasoning.* Information Sciences Part 1 : 8, 199-249 ;
 Part 2 : 8, 301-357 ; Part 3 : 9, 43-80.

CHAPITRE 5

FONCTIONS FLOUES

Après avoir particularisé la notion d'ensemble flou au cas d'une quantité floue, et montré qu'on peut définir un calcul arithmétique sur ces quantités floues, qui généralise le calcul d'erreurs, l'étape naturelle suivante est celle de l'analyse. Pour cela il convient de se donner une définition précise d'une "fonction floue". Ce chapitre passe en revue un certain nombre de définitions possibles, et examine dans quelles circonstances on peut être amené à construire des fonctions floues. Le concept d'intégrale d'une fonction floue au sens de Riemann est ensuite étudié, généralisé, et illustré sur la notion de variable aléatoire floue. On étudie aussi ce que peut vouloir dire l'intégrale d'une fonction entre deux bornes imprécises.

5.1 - QUELQUES DEFINITIONS POSSIBLES D'UNE FONCTION FLOUE

Soient Ω et U deux ensembles. On connaît dans la littérature quatre définitions possibles d'une "fonction floue" de Ω vers U .

A. Une fonction f dont le domaine de définition est un ensemble flou F sur Ω et dont les images forment un ensemble flou G sur U . Pour exprimer rigoureusement cette propriété on impose sur les α -coupes :

$$\forall \alpha > 0, \forall \omega \in F_\alpha, f(\omega) \in G_\alpha \quad (1)$$

ce qui équivaut à

$$\mu_F \leq \mu_G \circ f \quad (2)$$

Une telle définition a été posée par Negoita et Ralescu /17/. Au vu de (1), G est une "cible" floue qu'on atteint d'autant plus précisément que l'on choisit ω proche du noyau de f . On remarquera que (2) est une caractérisation très faible de f qu'on peut choisir de nombreuses façons différentes. Une fonction f satisfaisant (2) peut être appelée fonction d'un ensemble flou F dans un ensemble flou G . Pour une étude systématique, voir Ponasse /20/.

B. Une relation floue sur $\Omega \times U$. Si R est une telle relation floue, l'image floue de $\omega \in \Omega$ est un ensemble flou, noté $\tilde{f}(\omega)$ dont la fonction d'appartenance est

$$\mu_{\tilde{f}(\omega)}^v = \mu_R(\omega, \cdot) \quad (3)$$

De telles fonctions floues peuvent être directement définies à partir d'une famille $\{f_\lambda\}_{\lambda \in \mathbb{R}}$ paramétrée de fonctions, et une valeur floue $\tilde{\lambda}$ du paramètre. On a alors l'image $f_\lambda^\vee(\omega)$ de ω , calculée à l'aide du principe d'extension : (Chang and Zadeh, /4/).

$$\forall u \in U, \mu_{f_\lambda^\vee(\omega)}(u) = \sup \{ \mu_\lambda^\vee(v) / v: f_v(\omega) = u \} \quad (4)$$

$$\triangleq \mu_R(\omega, u)$$

La composition de deux fonctions floues de ce type est une composition de relations floues (cf. chap.3).

C. Un ensemble flou Φ de fonctions de Ω dans U (Negota et Ralescu /17/) Φ est donc défini par une fonction d'appartenance $\mu_\Phi : U^\Omega \rightarrow [0,1]$. $\forall f \in U^\Omega$, $\mu_\Phi(f)$ est le degré d'appartenance de la fonction f au faisceau flou de fonctions Φ . Une fonction à paramètre flou f_λ^\vee est un exemple d'ensemble flou de fonctions dont le support est

$$S(f_\lambda^\vee) = \{f_\lambda / \lambda \in S(\tilde{\lambda})\}$$

et

$$\mu_{f_\lambda^\vee}(f) = \mu_\lambda^\vee(\lambda) \quad \text{si } \exists \lambda, f = f_\lambda \quad (5)$$

$$= 0 \text{ sinon}$$

On notera que la donnée de Φ ne suffit pas à définir l'image d'un élément $\omega \in \Omega$ par Φ . Par exemple, si $\Omega = \mathbb{R} = U$

$$S(\Phi) = \{f_1, f_2\} \quad \text{tel que } \mu_\Phi(f_1) \neq \mu_\Phi(f_2)$$

avec

$$f_1(\omega) = \omega, \quad f_2(\omega) = \omega^2$$

$$S\Phi(\omega) \triangleq \{f(\omega) / f \in S(\Phi)\}, \quad \text{on voit que } S\Phi(1) = \{1\}$$

mais quelle valeur d'appartenance attribuer au couple $(1,1)$? $\mu_\Phi(f_1)$ ou $\mu_\Phi(f_2)$? Pour le savoir on a besoin d'une règle, telle que (4), qui définisse la valeur d'appartenance de u à l'image de ω par Φ en fonction des valeurs d'appartenance $\mu_\Phi(f)$, $\forall f \in \{f \in S(\Phi), u = f(\omega)\}$.

La composition de deux faisceaux flous Φ et ψ est définie à l'aide du principe d'extension, si Φ est un faisceau flou de Ω sur U , et ψ de U sur V :

$$\forall h : \Omega \rightarrow V, \mu_{\psi \circ \Phi}(h) = \sup \{ \min(\mu_{\psi}(g), \mu_{\Phi}(f)) \mid h = f \circ g \}$$

D. Une fonction φ de $[0,1]^{\Omega}$ dans $[0,1]^U$, c'est-à-dire une fonction qui à chaque ensemble flou de Ω , fait correspondre un ensemble flou de U . Par exemple si $f : \Omega \rightarrow U$ est une fonction usuelle, le principe d'extension permet de définir $f(F)$, $\forall F \in [0,1]^{\Omega}$ (cf. chap.3). Si R est une relation floue sur $\Omega \times U$, alors la composition sup- \perp ou \perp est une t-norme permet de définir l'image de F par R de façon cohérente avec la notion habituelle d'image d'un sous-ensemble de Ω : $\varphi(F) = R \circ F$ (cf chap.3); on notera que le principe d'extension appliqué à une fonction f est un cas particulier de composition sup- \perp d'un ensemble flou par une relation floue.

Si φ est telle qu'il existe R , relation floue de $\Omega \times U$, telle que

$$\forall F \in [0,1]^{\Omega}, \varphi(F) = R \circ F \quad (6)$$

alors : φ est croissante au sens de l'inclusion de Zadeh : $\forall F, F'$

$$\mu_F \leq \mu_{F'} \rightarrow \mu_{R \circ F} \leq \mu_{R \circ F'}$$

$$\varphi(F \cup F') = \varphi(F') \cup \varphi(F) \quad ; \quad \varphi(F \cap F') \leq \varphi(F) \cap \varphi(F')$$

où \cup et \cap sont au sens de maximum et minimum respectivement

Une relation R vérifiant (6) n'existe pas toujours. Par exemple, si $U = \Omega$ et $\varphi(F) = \bar{F}$, le complémentaire de F , alors φ n'est certainement pas croissante au sens de l'inclusion ! On voit donc que ce point de vue est très général.

Le point de vue A est trop particulier ; D est trop général; avec l'intuition d'une fonction qui à chaque élément $\omega \in \Omega$ associe une image imprécise, on définira une fonction floue \tilde{f} à l'aide d'une relation floue R , par l'intermédiaire de (3). Une application floue sera une fonction floue vérifiant de plus la condition :

$$\forall \omega \in \Omega, \tilde{f}(\omega) \neq \emptyset \quad (7)$$

Etant donné un faisceau flou de fonctions, Φ , on peut, au sens de la théorie des possibilités de Zadeh, calculer une fonction floue \tilde{f}_Φ qui, pour chaque ω , définisse une image floue $\tilde{f}_\Phi(\omega)$, de façon cohérente avec (4) :

$$\forall (\omega, u), \mu_{\tilde{f}_\Phi(\omega)}(u) = \sup \{ \mu_\Phi(f) / u = f(\omega) \} \quad (8)$$

On obtient ainsi le degré de possibilité pour que u soit image de ω par Φ . Le problème inverse : étant donnée une fonction floue \tilde{f} , trouver un faisceau flou de fonction Φ tel que $\tilde{f}_\Phi = \tilde{f}$ au sens de (8) a été traité par Dubois et Prade /7/.

Si f est une fonction de Ω dans U , on peut, conformément au chapitre 2, section 1, définir à quel point f est un bon représentant de \tilde{f} par l'indicateur :

$$\mu_\gamma(\tilde{f})(f) = \inf \{ \mu_{\tilde{f}(\omega)}(f(\omega)) / \omega \in \Omega \} \quad (9)$$

(9) définit un faisceau flou de fonctions $\gamma(\tilde{f})$. On notera que si \tilde{f} n'est pas application floue, la violation de (7) implique $\gamma(\tilde{f}) = \emptyset$.

De plus $\gamma(\tilde{f})$ est normalisé si et seulement si $\forall \omega \in \Omega, \tilde{f}(\omega)$ est normalisé. Enfin on vérifie que, en général $\gamma(\tilde{f} \circ \tilde{g}) \supseteq \gamma(\tilde{f}) \circ \gamma(\tilde{g})$.

Il est facile de vérifier, qu'en général, il n'est pas possible de reconstruire \tilde{f} à partir de $\gamma(\tilde{f})$; en particulier par une règle telle que (8).

Néanmoins il existe une classe importante de fonctions floues, nommées applications floues fortes dans /7/, telles que \tilde{f} peut se déduire de $\gamma(\tilde{f})$ par (4). Selon la terminologie de Ponasse /20/ on définit la ligne de flou d'un ensemble flou F sur Ω , l'ensemble $L_\alpha F$ tel que

$$L_\alpha F = \{ \omega / \mu_F(\omega) = \alpha \}$$

Et pour une fonction floue on a

$$L_\alpha \tilde{f} = \{ (\omega, u) / \mu_{\tilde{f}(\omega)}(u) = \alpha \}$$

Une application floue forte est une fonction floue \tilde{f} dont chaque ligne de flou est soit vide, soit contient une application. Cette dernière condition s'écrit

$$\forall \alpha \in]0, 1[\exists f_\alpha : \Omega \rightarrow U, \forall \omega, (\omega, f_\alpha(\omega)) \in L_\alpha \tilde{f}$$

Il est facile de voir qu'une application floue forte est une application floue. On a alors le résultat suivant (Dubois et Prade /7/) :
Si $\overset{\gamma}{f}$ est une application floue forte alors

$$\forall (\omega, u), \mu_{\overset{\gamma}{f}(\omega)}^{\gamma}(u) = \sup_{\omega'} \{ \inf_{\omega'} \mu_{\overset{\gamma}{f}(\omega')}^{\gamma}(\omega') / f : u = f(\omega) \} \quad (10)$$

Une application floue forte peut donc se voir comme une famille d'applications f telles que

$$\forall \omega \neq \omega' \quad \mu_{\overset{\gamma}{f}(\omega)}^{\gamma}(f(\omega)) = \mu_{\overset{\gamma}{f}(\omega')}^{\gamma}(f(\omega')) = \mu_{\gamma(\overset{\gamma}{f})}^{\gamma}(f)$$

5.2 - CONSTRUCTION DE FONCTIONS FLOUES

A côté du point de vue mathématique, on peut chercher les situations pratiques où l'on est amené à spécifier une fonction de façon imprécise, et des méthodes pour construire, à partir de données brutes, une application floue. On considère ici essentiellement des fonctions réelles à valeurs dans \mathbb{R} .

5.2.1 - Spécifications incomplètes de fonctions (Dubois et Prade /10/)

Expérimentalement, une fonction peut nous être donnée sous l'une des formes suivantes :

i) une spécification exacte mais incomplète : C'est-à-dire qu'on ne connaît f que pour un sous-ensemble (généralement fini) de \mathbb{R} , sous la forme $D = \{(x_i, f(x_i)) / i = 1, n\}$. La restitution de f à partir d'un tel échantillonnage est un problème de lissage dont les techniques sont bien connues. On signalera malgré tout le travail de Pavlidis et Chang /18/ qui utilisent une définition floue de l'adéquation d'une fonction spline à une ligne polygonale (construite sur les données) aux points anguleux de celle-ci. Ce critère est utilisé, en plus du critère traditionnel d'erreur quadratique moyenne, pour affiner la représentation de la fonction.

ii) Une spécification imprécise : Par exemple un instrument de mesure fournit un ensemble de paires d'intervalles $\{(X_i, Y_i) / i = 1, n\} = D$ tels que

$$\forall r \in X_i, f(r) \in Y_i \quad (11)$$

La spécification est complète dès que $\bigcup_{i=1,n} X_i = \mathfrak{X}$, domaine de définition de F.

Pour qu'elle ne soit pas contradictoire, il faut de plus que

$$\forall X_i, X_j, \text{ si } X_i \cap X_j \neq \emptyset, \text{ alors } Y_i \cap Y_j \neq \emptyset \quad (12)$$

(12) est une condition nécessaire et suffisante d'existence de f. Si (12) est faux on doit chercher à représenter D par quelque chose de plus faible qu'une fonction : une relation, une fonction aléatoire, par exemple.

iii) Une spécification linguistique : Une spécification imprécise peut être fournie par un individu sous forme linguistique (Zadeh, /33/). On peut la représenter sous la forme d'une collection de propositions de la forme "Si X est X_i alors Y est Y_i ", $i = 1, n$ où X_i et Y_i sont des ensembles flous sur les référentiels de X et Y respectivement. Ce type de spécification linguistique est courant dans les systèmes experts en intelligence artificielle, et a été utilisé pour décrire la façon dont on doit piloter un système complexe (cf. Mamdani, /16/) ou pour décrire les relations entre l'entrée et la sortie d'un tel système (Tong, /29/), perçues par un observateur. On a donc ici un ensemble $\tilde{D} = \{(X_i, Y_i) / i = 1, n\}$ qui généralise la spécification imprécise ii) en une spécification floue.

Si Ω est le domaine de définition de la fonction spécifiée linguistiquement par \tilde{D} , le degré de complétude de la spécification sera

$$C(\tilde{D}) = \inf_{r \in \mathfrak{X}} \mu \bigcup_{i=1,n} X_i (r) \quad (13)$$

où \bigcup s'exprime par un opérateur d'union d'ensembles flous. En particulier si on utilise la somme bornée, $C(\tilde{D}) = 1$ s'exprime par

$$\forall r \in \mathfrak{X}, \sum_{i=1}^n \mu_{X_i}(r) \geq 1$$

La condition de non-contradiction (12) peut s'étendre au cas d'une spécification floue en

$$\forall i, j, \sup_{r' \in \mathfrak{Y}} \mu_{Y_i}(r') \perp \mu_{Y_j}(r') \geq \sup_{r \in \mathfrak{X}} \mu_{X_i}(r) \perp \mu_{X_j}(r) \quad (14)$$

où \perp est une t-norme (par exemple $\perp = \min$).

Enfin l'existence de redondance dans la spécification peut s'exprimer par

$$\exists (X_i, Y_i), (X_j, Y_j), X_i \subseteq X_j \text{ et } Y_i \subseteq Y_j$$

Dans ce cas l'une des deux paires peut être supprimée, pourvu que le degré de complétude de \tilde{D} ne diminue pas trop.

iv) une spécification précise mais incohérente : On obtient cette fois D sous la forme d'un ensemble $\{(x_i, y_i) / i=1, n\}$ de points précis tels qu'à une valeur x_i puisse correspondre plusieurs valeurs y_i . On a ce qu'on appelle habituellement, un "nuage de points" dans \mathbb{R}^2 . Le caractère non-déterministe de la spécification peut provenir de ce que l'on a négligé des variables d'état secondaires dans le modèle. C'est-à-dire qu'au lieu d'avoir les valeurs $y \in Y$ fonctions de $x \in X$ uniquement on a

$$y = f(x, z)$$

où z est une variable d'effet limité, et mal connu, sur y . On est alors amené à représenter la relation entre x et y soit par une fonction f qui minimise un critère d'erreur entre $f(x_i)$ et $y_i \forall i = 1, N$, soit par une fonction aléatoire de type $y = f(x) + \omega$ où ω est une variable aléatoire de moyenne nulle ; une autre possibilité est d'utiliser une application floue.

5.2.2 - Approximation d'une spécification linguistique par une fonction

On ne considère ici que le cas iii) puisqu'il subsume ii). Par ailleurs les méthodes d'obtention de fonction dans les cas i) et vi) sont déjà connues dans la littérature : il s'agit respectivement des techniques de lissage, et des méthodes d'analyse de données de type régression, etc...

Considérons la situation où l'on possède une spécification linguistique sous la forme d'un ensemble de "granules" (Zadeh /34/) $\{(X_i, Y_i), i = 1, n\} = \tilde{D}$. On veut trouver une fonction $y = f(x)$ qui approxime au mieux \tilde{D} . Ce problème a été fort peu étudié dans la littérature des ensembles flous.

Une première méthode, très empirique, a été suggérée par

Rao et Rutherford /24/. Elle comporte les étapes suivantes

- . Calculer le "graphe" de D sous la forme d'une relation floue

$$R = \bigcup_{i=1, n} X_i \times Y_i$$

- . Prendre l' α -coupe R_α de ce graphe pour un seuil α -donné.

- . Calculer

$$\forall u, f_\alpha^+(u) = \sup \{v / (u,v) \in R_\alpha\}$$

$$f_\alpha^-(u) = \inf \{v / (u,v) \in R_\alpha\}$$

quand ces quantités existent.

- . Poser $f_\alpha(u) = \frac{1}{2} (f_\alpha^+(u) + f_\alpha^-(u))$ (ce qui est arbitraire)

- . Lisser afin de raccorder les morceaux de fonction obtenus.

Une étude plus rigoureuse du problème a été esquissée par Dubois et Prade /10/. Une fonction f sera une bonne représentation de \tilde{D} , si, au sens du principe d'extension $Y_i = f(X_i) \forall i = 1, n$, c'est-à-dire

$$\forall i, \forall y \in Y_i, \mu_{Y_i}(y) = \sup \{ \mu_{X_i}(x) / y = f(x) \} \quad (15)$$

on notera que (15) implique

$$\forall i, \mu_{X_i} \leq \mu_{Y_i} \circ f \quad (16)$$

c'est-à-dire que la restriction de f au support de X_i est une fonction de l'ensemble flou X_i dans l'ensemble flou Y_i , au sens de la section 1, cas A.

L'équation fonctionnelle ($\forall i, Y_i = f(X_i)$) est un cas particulier d'équation de relations floues (cf. chap.3) c'est-à-dire que parmi les relations floues R qui satisfont l'égalité $Y_i = X_i \circ R$ on cherche une solution qui soit une fonction.

On suppose que le support de X_i et Y_i est fini, ou que μ_{X_i} et μ_{Y_i} sont continues. La solution f_i de $Y_i = f_i(X_i)$ est telle que

$$\forall u \in S(X_i), f_i(u) \in \{v / \mu_{X_i}(u) = \mu_{Y_i}(v)\} \triangleq \Phi_i(u) \quad (17)$$

Si $S(X_i)$ et $S(Y_i)$ sont finis on peut facilement avoir $\Phi_i(u) = \emptyset$. Mais si μ_{X_i} et μ_{Y_i} sont continus alors $\Phi_i(u) \neq \emptyset \quad \forall u \in S(X_i)$.

Pour obtenir une fonction f continue, en raccordant des fonctions f_i , $i = 1, n$, il faut que les conditions suivantes soient respectées sur les $\Phi_i(u)$:

$$\begin{cases} \forall u \in S(X_i) \quad \Phi_i(u) \neq \emptyset \\ \forall i, \forall j, \forall u \in S(X_i \cap X_j) \neq \emptyset, \quad \Phi_i(u) \cap \Phi_j(u) \neq \emptyset \end{cases} \quad (18)$$

Dans ce cas toute solution f de (15) est telle que

$$\forall u, u \in S(X_i) \rightarrow f(u) \in \Phi_i(u) \quad (19)$$

On vérifie que (18) implique la non-contradiction de la spécification \tilde{D} , c'est-à-dire (14) avec $\perp = \min$. Pour qu'une fonction f spécifiée par (19) existe il faut que $\forall i, j \quad i \neq j$

$$\forall u \in S(X_i \cap X_j), \mu_{Y_i}(f(u)) = \mu_{X_i}(u) ; \mu_{Y_j}(f(u)) = \mu_{X_j}(u) \quad (20)$$

Cette condition, très stricte, sera très rarement vérifiée en pratique. Cela indique que (15) n'a généralement pas de solution sous la forme d'une fonction. Si on se contente de la condition (16), la solution f_i de (16) est telle que

$$\forall u \in S(X_i), f_i(u) \in \{v / \mu_{X_i}(u) \leq \mu_{Y_i}(v)\} = \hat{\Phi}_i(u) \quad (21)$$

on notera que $\Phi_i(u)$ est la ligne de flou $L\mu_{X_i}(u) Y_i$, qui si Y_i est un intervalle flou, contient au plus 2 points, tandis que $\hat{\Phi}_i(u)$ est l' α -coupe de Y_i pour $\alpha = \mu_{X_i}(u)$; c'est-à-dire que si Y_i est un intervalle flou, $\hat{\Phi}_i(u)$ est un intervalle, la fermeture convexe de $\Phi_i(u)$.

Réécrite avec $\hat{\Phi}_i$, la condition (18) devient beaucoup plus raisonnable, et la condition d'existence de f , devient, au lieu de (20), $\forall i, \forall j \quad i \neq j$:

$$\forall u \in S(X_i \cap X_j), f(u) \in (Y_i)_\alpha \cap (Y_j)_\beta$$

avec

$$\alpha = \mu_{X_i}(u), \quad \beta = \mu_{X_j}(u)$$

Quand la représentation obtenue est jugée non satisfaisante on peut chercher à représenter \tilde{D} sous la forme d'une relation floue qui soit à la fois fidèle à la spécification initiale, et facile à utiliser (par exemple à stocker dans la mémoire d'un ordinateur). En effet, avec une représentation purement fonctionnelle de \tilde{D} on perd de la souplesse ; notamment si les règles de type "si X_i alors Y_i " décrivent un processus décisionnel. L'avantage d'une fonction "floue" est de pouvoir en moduler les images de façon à s'adapter à une situation donnée, tout en restant cohérent avec le modèle du processus décisionnel.

5.2.3 - Approximation d'une spécification linguistique par une relation floue

On cherche maintenant une relation floue R telle que

$$\forall i = 1, n \quad Y_i = R \circ X_i \quad (22)$$

où "o" est une composition sup- \perp au sens d'une t-norme \perp continue. On décompose le problème en n sous-problèmes de la forme $Y_i = R_i \circ X_i$, on suppose que Y_i et X_i ont des fonctions d'appartenance s.c.s. et normalisées. On vérifie les résultats suivants :

. L'équation $Y_i = R_i \circ X_i$ a toujours une solution. Une solution particulière est $R_i = X_i \times Y_i$ où le produit cartésien est défini au sens de la t-norme présente dans la composition o.

. La plus grande solution est $\hat{R}_i = X_i \rightarrow Y_i$ de fonction d'appartenance

$$\mu_{\hat{R}_i}(u, v) = \mu_{X_i}(u) \mathcal{Q} \mu_{Y_i}(v) \quad (23)$$

(cf. chap.2)

En particulier si $\circ =$ composition sup-min, alors

$$\begin{aligned} a \mathcal{Q} b &= 1 \quad \text{si } a \leq b \\ &= b \quad \text{si } a \geq b \end{aligned}$$

et le résultat est dû à Sanchez /25/.

Dans le cas d'une t-norme continue le résultat est dû à Pedrycz /19/.

. Définissons l'opérateur σ par

$$\begin{aligned} a \sigma_{\perp} b &= 0 & \text{si } a < b \\ &= a \wedge b & \text{sinon} \end{aligned} \quad (24)$$

alors la relation floue $\underline{R}_i = X_i \sigma Y_i$ est solution de $Y_i = R_i \circ X_i$; pour $\perp = \min$, Sanchez /26/ a montré que sous certaines conditions \underline{R}_i était la plus petite solution de $Y_i = R_i \circ X_i$ (au sens de l'inclusion de Zadeh). Néanmoins cette plus petite solution n'existe pas toujours. On remarque que

$$\mu_{\underline{R}_i}(u,v) = \mu_{X_i \times Y_i}(u,v) \quad \text{ou} \quad 0.$$

On a donc un ensemble de solutions de $Y_i = R_i \circ X_i$ sous la forme

$$\mathcal{R}_i = \{R_i / \underline{R}_i \leq R_i \leq \hat{R}_i\} \quad (25)$$

Donc on trouve une solution du problème (22) sous la forme

$$R \in \bigcap_{i=1,n} \mathcal{R}_i = \{R / \bigcup_i \underline{R}_i \leq R \leq \bigcap_i \hat{R}_i\} \quad (26)$$

(les symboles \bigcup et \bigcap entre accolades sont respectivement max et min).

Si $\bigcap_{i=1,n} \mathcal{R}_i$ est vide cela ne prouve pas que (22) n'ait pas de solution

Si $\bigcap_{i=1,n} \mathcal{R}_i$ n'est pas vide, il contient la plus grande solution de (22)

qu'on peut considérer comme la plus valide (car la plus imprécise !)

NB : On notera que si $u \notin S(X_i)$ alors $\mu_{\hat{R}_i}(u,v) = 1 - \mu_{\underline{R}_i}(u,v) = 1$ donc (26) définit un ensemble vide seulement si

$$\exists u, \max \{ \mu_{\underline{R}_i}(u) / i \in K(u) \} > \min \{ \mu_{\hat{R}_i}(u) / i \in K(u) \}$$

où

$$K(u) = \{i / u \in S(X_i)\}$$

Donc les seuls conflits entre ces règles apparaissent dans les zones de chevauchement des X_i .

On remarquera la différence d'interprétation entre des solutions du type $\hat{R}_i = X_i \rightarrow Y_i$ qui présupposent une causalité de X_i vers Y_i et des solutions du type produit cartésien qui n'expriment qu'une corrélation entre les X_i et les Y_i . En particulier :

$$Y_i \circ (X_i \rightarrow Y_i) \neq X_i$$

tandis que

$$Y_i \circ (X_i \times Y_i) = X_i$$

Cette différence se traduit aussi au niveau du principe d'agrégation des solutions R_i :

. On fait l'intersection des \hat{R}_i dont les fonctions d'appartenance valent 1 dès que $u \notin S(X_i)$

. On fait l'union des \underline{R}_i qui représentent des "îlots" de support inclus dans $S(X_i) \times S(Y_i)$.

Connaissant R , solution de (22), on peut calculer la réponse du système décrit par R à une entrée X , définie par un intervalle flou, sous la forme $Y = R \circ X$ et réaliser ainsi une interpolation entre les règles "si X_i alors Y_i " qui forment \tilde{D} .

On aurait pu aussi, comme le font Baldwin et Guild /2/ calculer l'ensemble des $\{R_i \circ X / i=1, n\}$ et agréger les "réponses" de chaque règle après coup. Savoir dans quels cas les deux processus sont équivalents reste une question ouverte.

5.2.4 - Régression linéaire floue

Tanaka, Uejima, et Asai /27-28/ ont proposé une méthode pour approximer un nuage de points par une fonction linéaire floue.

Les données se présentent sous la forme $D = \{(\vec{x}_i, y_i) / i=1, n\}$ où les \vec{x}_i sont des vecteurs à m composantes. On cherche à représenter D sous la forme

$$\hat{f}(\vec{x}) = \sum_{j=1}^m M_j x^j$$

où les M_j sont des intervalles flous à définir.

Le degré de cohérence entre \hat{f} et D au point \vec{x}_i est défini par $\mu_{\hat{f}(\vec{x}_i)}^{\hat{f}}(y_i)$.

La précision de la solution est d'autant meilleure que les supports des M_j sont étroits. Une mesure de précision est donc

$$\sum_{j=1}^m l(S(M_j)) \quad \text{ou} \quad l([a, b]) = b - a.$$

Le problème de régression linéaire floue se pose donc sous la forme d'une optimisation de fonctionnelles

Trouver $\mu_{M_1} \dots \mu_{M_m}$ telles que $\sum_{j=1}^m l(S(M_j))$ soit minimum

sous les contraintes

(27)

$$\mu_{\hat{f}(\vec{x}_i)}^{\hat{f}}(y_i) \geq h \quad \forall i = 1, n$$

où h est un seuil de cohérence minimale entre le modèle et les données. Si $M_1 \dots M_m$ sont des nombres flous de type L-R et que $x_i^j \geq 0 \forall i, j$ on a

$$M_j = (m_j, \alpha_j, \beta_j)_{LR}$$

$$\hat{f}(\vec{x}) = \left(\sum_{j=1, m} m_j x^j \quad \sum_{j=1, m} \alpha_j x^j, \quad \sum_{j=1, m} \beta_j x^j \right)_{LR}$$

Et le problème (27) se ramène à de la programmation linéaire, si L et R sont données :

$$\min \sum_{j=1}^m \alpha_j + \beta_j$$

$$-\left(\sum_{j=1, m} \beta_j x_i^j \right) R^{-1}(h) \leq \sum_{j=1, m} m_j x_i^j - y_i \leq \left(\sum_{j=1, m} \alpha_j x_i^j \right) L^{-1}(h)$$

$$\alpha_i, \beta_i, m_i \geq 0$$

On notera que Tanaka et al. supposent L et R linéaires et M_j symétriques, mais comme on le voit ci-dessus, ces hypothèses sont parfaitement inutiles. On peut donc résoudre le problème avec plusieurs formes L-R, afin d'améliorer la solution du problème ci-dessus.

NB : il est possible de maximiser la cohérence modèle-données avec une contrainte sur la précision; néanmoins le paramètre h devient alors une inconnue et on ne peut plus espérer se ramener à de la P.L. sauf si les μ_{M_i} ont des formes linéaires.

L'idée de représenter un nuage de points par une fonction floue d'équation analytique donnée a été suggérée d'abord par Chang /3/. Si la fonction floue paramétrée est $y = f_{\beta}(u)$ où β est un vecteur de paramètres flous, la coupe de niveau 1 de f_{β} est la meilleure estimée f_{β_0} de la fonction au sens des moindres carrés, tandis que la valeur d'appartenance $\mu_{f_{\beta}(u)}(v)$ décroît avec la distance entre les points (u,v) et $(u, f_{\beta_0}(u))$. Cette approche diffère de celle de Tanaka et al.

Une autre méthode, plus ancienne, est celle de Gitman et Levine /13/, qui ne présuppose pas de formulation analytique. A partir d'un ensemble de points $D = \{\omega_i, i = 1, n\}$ d'un espace métrique Ω on définit la fonction d'appartenance d'un ensemble flou F^{θ} sur Ω comme une mesure de concentration; si d est une distance sur Ω

$$\forall \omega \in \Omega, \mu_{F^{\theta}}(\omega) = \frac{1}{n} \cdot |\{\omega_i / d(\omega, \omega_i) \leq \theta\}|$$

On remarquera néanmoins que l'ensemble flou obtenu est généralement multimodal.

5.2.5 - Lissage d'une application floue

Une représentation analytique particulièrement simple à manipuler, bien que relativement générale, pour des applications floues fortes de \mathbb{R} dans \mathbb{R} a été suggérée par Dubois et Prade /5/ ; elle tire parti de la représentation L.R. Elle suppose que l'application floue \tilde{f} définie sur \mathbb{R} est telle que

$\rightarrow \forall u, \tilde{f}(u)$ est un intervalle flou de type LR

$\rightarrow \forall \alpha > 0$ l' α -coupe \tilde{f}_{α} est une bande limitée par deux fonctions f_{α}^{+} et f_{α}^{-} dites lignes de niveaux supérieures et inférieures : $\tilde{f}_{\alpha} = \{f / f_{\alpha}^{-} < f < f_{\alpha}^{+}\}$

\tilde{f} est donc définie à l'aide des deux lignes de niveau 1, f_1^{+} et f_1^{-} , et de deux fonctions A et B : $\mathbb{R} \rightarrow \mathbb{R}$, positives, et on a

$$\forall u, \tilde{f}(u) = (\tilde{f}_1^-(u), \tilde{f}_1^+(u), A(u), B(u))_{LR} \quad (28)$$

Etant donné une relation floue R sur \mathbb{R}^2 , qui soit telle que $\forall u, \mu_R(u_i, \cdot)$ définisse un intervalle flou, on peut mettre R sous la forme (28) par la procédure suivante :

. Choisir un échantillonnage $\{u_1 \dots u_n\}$ de \mathcal{X} ; faire un lissage de $\mu_R(u_i, \cdot)$ par un intervalle flou de type LR, au sens des moindres carrés. On obtient des nombres flous $\{M_i = (m_i^-, m_i^+, \alpha_i, \beta_i)_{LR} / i = 1, n\}$

. Déterminer des représentations analytiques des fonctions $u_i \rightarrow m_i^+, u_i \rightarrow m_i^-, u_i \rightarrow \alpha_i, u_i \rightarrow \beta_i$, par lissage au sens des moindres carrés, pour obtenir $\tilde{f}_1^-, \tilde{f}_1^+, A, B$.

Une telle représentation va s'avérer utile pour calculer efficacement des intégrales de fonctions floues.

5.3 - ELEMENTS D'ANALYSE FLOUE

5.3.1 - Intégration de Riemann d'une application floue

L'exposé qui suit résume l'article de Dubois et Prade /7/, dont une première version se trouve dans /5/ et /6/.

Soit \tilde{f} une application floue de \mathbb{R} dans \mathbb{R} , définie sur un intervalle $I = [a, b]$. On suppose que $\forall u, \tilde{f}(u)$ est à support borné. Soit \mathcal{J} l'ensemble des fonctions intégrales de $[a, b]$ dans \mathbb{R} . On note $\int_I f$ l'intégrale au sens de Riemann de f sur $[a, b]$; c'est une notation allégée pour $\int_I f(s) ds$.

L'intégrale de l'application floue à support borné \tilde{f} , sur l'intervalle I est un ensemble flou de \mathbb{R} , noté $\int_I \tilde{f}$, défini par :

$$\forall v \in \mathbb{R}, \mu_{\int_I \tilde{f}}(v) = \sup \{ \mu_{\gamma(\tilde{f})}(f) / \int_I f = v, f \in \mathcal{J} \} \quad (29)$$

où $\gamma(\tilde{f})$ est le faisceau flou de fonctions qui représentent f , défini par l'équation (9). $\int_I \tilde{f}$ est la surface floue de l'aire limitée par les axes $y = 0$, $x = a$, $x = b$, et la frontière floue \tilde{f}

Si \tilde{f} est une fonction intégrable usuelle, $\int_I \tilde{f}$ est l'intégrale au sens usuel.
Si \tilde{f} est une multiapplication \bar{f} (29) se ramène à l'intégrale d'Aumann /1/ :

$$\int_I \bar{f} = \{ \int_I f / f \in \bar{f} \cap \mathcal{F} \} \quad (30)$$

En particulier si $\int_I \tilde{f}_\alpha$ est l'intégrale d'Aumann de l' α -coupe \tilde{f}_α , on a

$$\forall v, \mu_{\int_I \tilde{f}_\alpha}(v) = \sup \{ \alpha / v \in \int_I \tilde{f}_\alpha \} \quad (31)$$

Néanmoins on n'a pas toujours $(\int_I \tilde{f})_\alpha = \int_I \tilde{f}_\alpha$. Quand cette condition est vérifiée, on l'appelle condition de commutativité. On rappelle (cf. chap.3) qu'une condition nécessaire et suffisante de cette commutativité est que le suprémum soit atteint en (29). Mais, avec les α -coupes strictes, $(\int_I \tilde{f})_\alpha = \int_I \tilde{f}_\alpha$ est toujours vrai.

On montre en /7/ les propriétés suivantes ⁽¹⁾ :

. linéarité : $\int_I \tilde{f} \oplus \tilde{g} \supseteq \int_I \tilde{f} \oplus \int_I \tilde{g}$ où \oplus est l'addition de quantités floues (cf. chap.4) parce que $\gamma(\tilde{f} \oplus \tilde{g}) \supseteq \gamma(\tilde{f}) \oplus \gamma(\tilde{g})$. L'égalité est prouvée en utilisant la condition de commutativité.

. Si I et I' sont des intervalles adjacents alors

$$\int_{I \cup I'} \tilde{f} = \int_I \tilde{f} \oplus \int_{I'} \tilde{f}$$

. Convexité : Si la condition de commutativité tient alors $\int_I \tilde{f}$ est un ensemble flou convexe (i.e. $\mu_{\int_I \tilde{f}}$ est quasi-concave).

Le dernier résultat généralise un théorème d'Aumann /1/.

Considérons des applications floues:

continues : $\mu_{\tilde{f}(u)}$ et $\mu_{\tilde{f}(\cdot)}(v)$ sont continues

normalisées : $\tilde{f}(u)$ est normalisé $\forall u$

convexes : $\forall u$ $\tilde{f}(u)$ est convexe

Alors ce sont des applications floues fortes, au sens de la section 1.

Si \tilde{f} est continue alors la condition de commutativité est valide et $\int_I \tilde{f}$ peut être calculée comme la limite d'une somme (floue) de Riemann (convergence des α -coupes).

(1) En utilisant la représentation par α -coupes strictes, on peut vérifier qu'en fait linéarité et convexité sont toujours valides.

Soit f_λ une courbe de niveau, contenue dans la ligne de flou $L_\lambda f$, et continue ; on montre que

$$\mu_{f_I}^{\tilde{v}}(f_I f_\lambda) = \lambda.$$

Réciproquement, $\forall v \in S(f_I f)$, $\exists f_\lambda \in L_\lambda \tilde{f}$, f_λ continue et $v = f_I f_\lambda$.
On peut donc calculer $f_I \tilde{f}$ en intégrant toutes les courbes de niveau et on a

$$S(f_I \tilde{f}) = \{f_I f_\lambda / f_\lambda \in L_\lambda \tilde{f}, \lambda \in]0, 1]\}$$

qui est le support de la quantité floue $f_I \tilde{f}$

Enfin si \tilde{f} n'est pas normalisée, on définit la quantité $b(\tilde{f})$.

$$b(\tilde{f}) = \inf_{u \in I} \sup_{v \in \mathbb{R}} \mu_{\tilde{f}(u)}^{\tilde{v}}(v)$$

et si \tilde{f}_b est définie par $\mu_{\tilde{f}_b}^{\tilde{v}} = \min(\mu_{\tilde{f}}^{\tilde{v}}, b(\tilde{f}))$
alors

$$f_I \tilde{f} = f_I \tilde{f}_b$$

si $\tilde{f}(u)$ n'est pas convexe, soit \tilde{g} tel que $\tilde{g}(u)$ soit l'enveloppe convexe de $\tilde{f}(u)$, alors $f_I \tilde{f} = f_I \tilde{g}$.

Donc si \tilde{f} est une application floue continue quelconque alors il existe une application floue forte équivalente \tilde{g} , telle que les $\tilde{g}(u)$ sont convexes et de même hauteur, et l'intégrale de \tilde{g} est égale à celle de \tilde{f} .
Ce résultat généralise aussi un théorème d'Aumann /1/.

Les propriétés des applications floues continues sont encore valides pour des applications floues fortes (à support borné) telles que $\mu_{\tilde{f}(u)}^{\tilde{v}}$ est s.c.s.

On peut définir une primitive d'une application floue à support borné \tilde{f} . C'est une application floue $\tilde{F} : I \rightarrow \mathbb{R}$ telle que

$$\forall u \in [a, b], \exists c \in [a, b], \tilde{F}(u) = \int_c^u \tilde{f}$$

Néanmoins la propriété classique des primitives: $F(v) - F(u) = \int_u^v f$ ne tient plus. Si \tilde{f} est une application floue de type L-R (cf. § 2.5), alors toute primitive de \tilde{f} est encore de type LR et

$$\tilde{F}(u) = \left(\int_J \tilde{f}_1^-, \int_J \tilde{f}_1^+, \int_J A, \int_J B \right)_{LR}$$

avec

$$J = [c, u] \quad \text{ou} \quad [u, c]$$

Ce résultat indique clairement que le calcul pratique d'une intégrale de fonction floue peut être facile à mener.

Récemment, les résultats présentés ci-dessus ont été étendus à des fonctions de \mathbb{R} dans \mathbb{R}^n par Puri et Ralescu /23/, qui fournissent en plus un théorème de convergence de suites d'intégrales du type $\left\{ \int_I \tilde{f}_k, k \in \mathbb{N} \right\}$ au sens d'une généralisation de la distance de Hausdorff.

5.3.2 - Dérivée d'une application floue (Dubois et Prade /9/)

Soit une application \tilde{f} continue, à valeur dans les intervalles flous de \mathbb{R} , à support borné. Supposons de plus que toute ligne de flou de \tilde{f} ne contienne que deux lignes de niveaux f_λ^+ et f_λ^- différentiables si $\lambda < 1$ et une seule si $\lambda = 1$. On note $\omega(\tilde{f})$ le faisceau flou de fonctions tel que

$$\forall f, \mu_{\omega(\tilde{f})}(f) = \lambda \quad \text{si} \quad \exists \lambda > 0, \varepsilon \in \{-, +\} f = f_\lambda^\varepsilon$$

La dérivée d'une telle application floue au point u est une quantité floue notée $\tilde{f}'(u)$ telle que

$$\forall v, \mu_{\tilde{f}'(u)}(v) = \sup \{ \mu_{\omega(\tilde{f})}(f) / v = f'(u) \} \quad (32)$$

La dérivation conserve une propriété de linéarité

$$\tilde{f}'(u) \oplus \tilde{g}'(u) \underset{\geq}{=} (\tilde{f} \oplus \tilde{g})'(u)$$

où \oplus est l'addition au sens du principe d'extension sup-min.

La dérivation d'applications floues de type LR est relativement facile à effectuer, mais en général $\tilde{f}'(u)$ ne sera pas un nombre flou de

type L-R (voir /9/), selon le signe des dérivées $A'(u)$ et $B'(u)$:

Si $A'(u), B'(u)$ sont positifs alors $f'(u) = (f'_1(u), A'(u), B'(u))_{LR}$

Si $A'(u), B'(u)$ sont négatifs alors $f'(u) = (f'_1(u), -B'(u), -A'(u))_{RL}$

Si $A'(u) > 0, B'(u) < 0$ alors $f'(u) = M_1 \cup M_2$

$$\text{avec } M_1 = (f'_1(u), A'(u), 0)_{LR}$$

$$M_2 = (f'_1(u), -B'(u), 0)_{RL}$$

$\tilde{f}(u)$ décrit la façon dont sont inclinées les courbes de niveaux f_λ , $\lambda \neq 1$ par rapport à la ligne de crête de \tilde{f} , au voisinage de u .

Un autre concept, plus général, de différentielle d'une fonction floue est défini par Puri et Ralescu /29/. Notons d'abord que la définition usuelle de la dérivée à droite d'une fonction f en u :

$$f'(u_+) = \lim_{h \rightarrow 0_+} \frac{f(u+h) - f(u)}{h}$$

ne peut s'étendre en

$$\tilde{f}'(u_+) = \lim_{h \rightarrow 0_+} \frac{\tilde{f}(u+h) \ominus \tilde{f}(u)}{h}$$

où \ominus est la soustraction "pessimiste" de nombres flous, au sens du principe d'extension sup-min. Puri et Ralescu /21/ ont proposé d'utiliser la soustraction "optimiste" (cf. chap. 4) $M) - (N$, définie par la plus grande solution de l'équation $M = X \ominus N$, quand elle existe :

Si $\exists \varepsilon > 0, \forall h < \varepsilon, h > 0$, la quantité $\tilde{f}(u+h) - (\tilde{f}(u))$ existe, alors la dérivée à droite de l'application floue $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$ est

$$\tilde{f}'(u_+) = \lim_{h \rightarrow 0} \frac{\tilde{f}(u+h) - (\tilde{f}(u))}{h} \quad (33)$$

quand la limite existe.

Pour comparer (32) et (33), supposons que f est une application floue de type LR. En utilisant les résultats du chapitre 4 on voit que (33) n'est défini que si

$$\exists \varepsilon > 0, \forall h < \varepsilon, A(u+h) \geq A(u) \quad \text{et} \quad B(u+h) \geq B(u),$$

c'est-à-dire si $A'(u)$ et $B'(u)$ sont positifs. Dans ce cas, (32) et (33) coïncident puisque

$$\frac{\tilde{f}(u+h) - \tilde{f}(u)}{h} = \left(\frac{f_1(u+h) - f_1(u)}{h}, \frac{A(u+h) - A(u)}{h}, \frac{B(u+h) - B(u)}{h} \right)_{\mathbb{LR}}$$

(en remplaçant $f'(u)$ par $f'(u_+)$ dans (32)).

On voit donc que si (32) et (33) coïncident quand elles sont toutes deux définies il y a des cas où (32) a un sens alors que (33) n'en a pas ($A'(u) > 0$ et $B'(u) < 0$) et inversement ((33) ne suppose rien sur le contenu des lignes de flou de \tilde{f}).

5.3.3 - Intégration sur un intervalle flou

La première définition d'une intégrale sur un domaine flou a été formulée par Zadeh /32/ : c'est la probabilité scalaire d'un événement flou (cf. chapitre 2). Cette définition a été reprise par Puri et Ralescu /21/ hors du contexte probabiliste. Si F est un ensemble flou d'un espace $(\Omega, \mathcal{A}, \mu)$ mesurable alors l'intégrale d'une fonction mesurable $f = \Omega \rightarrow [0, +\infty)$ sur F , tel que μ_F est mesurable, vaut

$$\int_F f d\mu = \int_{\Omega} f \mu_F d\mu \quad (34)$$

on montre que si μ est σ -finie,

$$\exists \alpha \in]0, 1[, \int_F f d\mu = \int_{F_{\alpha}} f d\mu \leftrightarrow \forall \alpha \in]0, 1[\mu(L_{\alpha}(u)) = 0$$

C'est-à-dire que l'intégrale de f peut se calculer sur une α -coupe de F . Par contraste Yager /30/ a défini une probabilité floue d'un événement flou, rappelée au chapitre 3 § 6.2, basée sur la représentation d'un ensemble flou par ses α -coupes, puis modifiée /8/ pour inclure d'autres ensembles dans la représentation (cf. chap. 2 § 1).

Dubois et Prade /8/ ont défini l'intégrale d'une fonction f entre deux bornes floues A et B , qui sont des intervalles flous, définissant un domaine imprécis \mathcal{D} , par $\int_A^B f$:

$$\forall w, \mu_{\int_A^B f}(w) = \sup \left\{ \min(\mu_A(u), \mu_B(v)) / \int_u^v f(s) ds = w \right\} \quad (35)$$

f est supposée définie, intégrable sur $[\inf S(A), \sup S(B)]$

Si φ est une primitive de f , on montre que

$$\int_A^B f = \varphi(B) \ominus \varphi(A) \quad (36)$$

où \ominus est la soustraction pessimiste \sup, \min de nombres flous.

Les propriétés de cette intégrale sont

. linéarité $\int_A^B (f+g) \subseteq \int_A^B f \oplus \int_A^B g$. L'égalité tient si f et g sont toutes deux positives ou négatives.

. Si $\mathcal{D} = (A, B)$, $\mathcal{D}' = (C, D)$ sont deux domaines imprécis, et si

$$\overset{\vee}{\min}(A, C) = A \quad \text{et} \quad \overset{\wedge}{\max}(B, D) = B$$

alors

$$\overset{\wedge}{\max}\left(\int_A^B f, \int_C^D f\right) = \int_A^B f$$

où $\overset{\wedge}{\max}$ et $\overset{\vee}{\min}$ sont les opérateurs maximum et minimum étendus au sens du principe d'extension (cf. chap.4)

$$\int_A^C f \subseteq \int_A^B f \oplus \int_B^C f$$

Si M est un intervalle flou, la définition de Yager /30/, modifiée, de la probabilité d'un événement flou, revient à proposer comme intégrale de f sur l'intervalle flou M , la quantité $\int_M f$ telle que

$$\forall v, \mu_{\int_M f}(v) = \sup_I \left\{ \inf_{u \in S} \mu_M(u) \mid \int_I f(s) ds = v \right\} \quad (37)$$

où I est un intervalle.

On montre que si $M \subseteq M'$ alors

$$\overset{\wedge}{\max}\left(\int_M f, \int_{M'} f\right) = \int_M f$$

et que si

$$M_- = M \wedge \left] \sup M, +\infty \right), \quad M_+ = M \wedge (-\infty, \inf M \left[$$

alors

$$\int_M f = \int_{M_-}^{M_+} f \quad \text{au sens de (35)}$$

M_+ et M_- sont les bornes floues de l'intervalle flou M .

Ce résultat permet de relier de façon précise la définition de Yager /30/ et celle de /8/.

Parallèlement, on peut définir la dérivée en un point flou M d'une fonction, qui n'est autre que $f'(M)$ si f' est la dérivée de f sur le support de M /8/. Enfin on peut intégrer une fonction floue sur un intervalle flou : si $I(x,y) = \int_x^y \tilde{f}$, alors on peut poser

$$\int_A^B \tilde{f} = I \circ (AxB) \quad (\text{composition sup-min}). \text{ cf /8/}.$$

5.4 - VARIABLES ALEATOIRES FLOUES

Une variable aléatoire floue \tilde{X} sera une application floue \tilde{f} d'un ensemble d'événements Ω dans \mathbb{R} . On suppose dans un premier temps qu'il existe une partition finie $A_1 \dots A_n$ de Ω telle que $\forall \omega \in A_i, \tilde{f}(\omega) = M_i$ où M_i est une quantité floue. Si (Ω, \mathcal{A}, P) est un espace de probabilité où \mathcal{A} est induit par $\{A_1 \dots A_n\}$, on peut définir l'espérance mathématique de \tilde{X} par (Kwakernaak, /15/) :

$$E(\tilde{X}) = P(A_1)M_1 \oplus P(A_2)M_2 \dots \oplus P(A_n)M_n$$

Si les M_i sont des ensembles classiques, il est clair qu'on obtient sur \mathbb{R} un ensemble aléatoire, c'est-à-dire que la valeur de probabilité des événements tels que $\tilde{X} \in I$, où I est un intervalle réel n'est connue que de façon imprécise. Une variable aléatoire floue généralise donc la théorie de Shafer au cas d'éléments focaux flous.

Récemment Yager /3/ a suggéré une définition de la probabilité d'un événement flou quand les observations (c'est-à-dire ici les M_i ou les éléments focaux) sont floues.

Soit Q une quantité floue. On rappelle que la compatibilité de Q avec un autre ensemble flou N sur le même référentiel est $\tau(Q/N) = \mu_Q(N)$

obtenue par le principe d'extension, comme la valeur d'appartenance du point flou N à Q (cf. chap. 3 et 4).

La probabilité de l'événement flou " $\tilde{X} \in Q$ " est alors définie comme :

$$\tilde{P}(Q) = P(A_1) \tau(Q/M_1) \oplus \dots \oplus P(A_n) \tau(Q/M_n) \quad (38)$$

Elle est telle que $\tilde{P}(R) = 1$ et $\tilde{P}(\emptyset) = 0$

On vérifie que si M_i est un ensemble usuel $\forall i$, de même que Q , alors

$$\inf \tilde{P}(Q) = P_{\star}(Q) \quad \sup \tilde{P}(Q) = P^{\star}(Q)$$

c'est-à-dire que (38) généralise bien les probabilités supérieures et inférieures de Dempster (1967). La super-additivité (resp : sub-additivité) de P_{\star} (resp : P^{\star}) se généralise en : $\forall Q_1 \dots Q_n$ quantités floues

$$\begin{aligned} \tilde{P}\left(\bigcup_{i=1}^n Q_i\right) \leq & \tilde{P}(Q_1) \oplus \tilde{P}(Q_2) \oplus \dots \oplus \tilde{P}(Q_n) \ominus \sum_{i \neq j} P(Q_i \cap Q_j) \\ & \dots \oplus (-1)^{n+1} P\left(\bigcap_{i=1}^n Q_i\right) \end{aligned}$$

L'intérêt de cette notion de probabilité floue, trop récente pour qu'on ait pu l'étudier en détail, est évident pour modéliser certains schémas de raisonnement approché.

L'étude des variables aléatoires floues a été suggérée pour la première fois dans un cadre plus général par Féron /11,12/ en 1976. Les travaux de Kwakernaak /14,15/ et Puri-Ralescu /23/ considèrent des variables aléatoires floues de Ω dans R^n pour une mesure de probabilité non-atomique. L'espérance mathématique de la variable aléatoire floue est alors définie à l'aide de l'intégrale d'une fonction floue.

Une variable aléatoire floue se prête naturellement à la modélisation de résultats de recueils d'information de type sondage quand on admet que les réponses fournies peuvent être imprécises (par exemples des qualificatifs linguistiques).

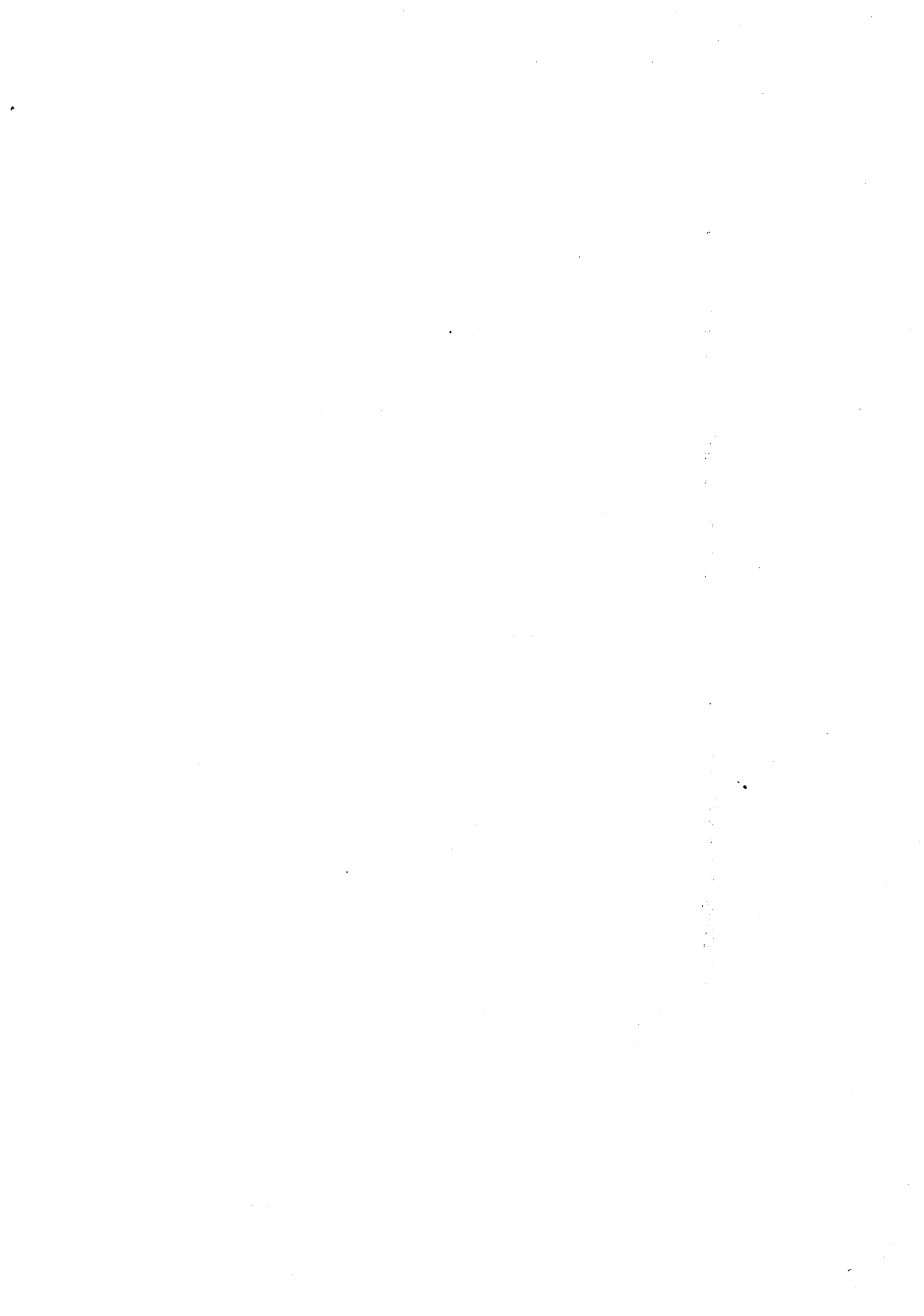
REFERENCES

- /1/ AUMANN R.J. (1965)
Integrals of set-valued functions.
J. Math. Anal. Applications 12, 1-12
- /2/ BALDWIN J.F., GUILD NCF (1980)
Modelling controllers using fuzzy relations.
Kybernetes 9, 223-229
- /3/ CHANG S.S.L. (1977)
Application of fuzzy set theory to economics.
Kybernetes 6, 203-207
- /4/ CHANG S.S.L., ZADEH L.A. (1972)
On fuzzy mapping and control.
IEEE Trans. Systems Man and Cybernetics, 2, 30-35.
- /5/ DUBOIS D., PRADE H. (1978)
*Towards fuzzy analysis : integration and differentiation of
 fuzzy mappings; in "Fuzzy Algebra, Analysis, Logics".*
 Tech. Rep. TREE 78-13, School of Electrical Engineering Purdue
 University, West Lafayette, IN
- /6/ DUBOIS D., PRADE H. (1980)
Fuzzy Sets and Systems : Theory and Applications,
 Academic Press, N.Y.
- /7/ DUBOIS D., PRADE H. (1982)
*Towards fuzzy differential calculus Part. 1 : intégration of
 fuzzy mappings, Fuzzy Sets & Systems 8, 1-17*
- /8/ Idem
*Part 2 : Integration on fuzzy intervals, Fuzzy Sets and
 Systems 8, 105-116*
- /9/ Idem
Part 3 : Differentiation, Fuzzy Sets and Systems 8, 225-233
- /10/ DUBOIS D., PRADE H. (1982)
*Toward the analysis and the synthesis of fuzzy mappings in :
Fuzzy Set and Possibility theory : Recent developments
 (R. Yager Ed.) Pergamon Press, 301-315.*

- /11/ FERON R. (1976)
Ensembles aléatoires flous, CRAS, vol. 282, serie A, 903-907
- /12/ FERON R. (1976)
Ensembles flous, ensembles aléatoires flous et économie aléatoire floue. Publications économétriques IX fasc. 2
- /13/ GITMAN I., LEVINE M.D. (1970)
An algorithm for detecting unimodal fuzzy sets and its application as a clustering technique. IEEE Trans. on Computers 19, 583-593
- /14/ KWAKERNAAK H. (1978)
Random fuzzy variables part I Definitions and Theorems, Information Sciences 15, 1-29
- /15/ Idem
Part II : Algorithms and Examples for the discrete case Information Sciences 17, 253-278, 1979.
- /16/ MAMDANI E. (1977)
Application of fuzzy logic to approximate reasoning using linguistic systems. IEEE Trans. Computers 26, 1182-1191
- /17/ NEGOITA C.V., RALESCU D. (1975).
Applications of fuzzy sets to systems analysis, Birkäuser Verlag, Basel,
- /18/ PAVLIDIS T., CHANG R.L.P. (1979)
Application of fuzzy sets in curve-fitting, Fuzzy Sets and Systems 2, 67-74
- /19/ PEDRYCZ, W. (1982)
Fuzzy Control and Fuzzy Systems, Tech. Rep. Dept. of Mathematics, Delft University of Technology, Pays-Bas
- /20/ PONASSE D. (1980)
Ensembles et Applications flous : généralités, étude catégorique, étude ponctuelle in "Seminaire Mathématique Floue 1979-80". (D. Ponasse Ed.) 149-180, Université de Lyon
- /21/ PURI M. , RALESCU D. (1979)
Integration on fuzzy sets, Indiana University, Bloomington , I.N.

- /22/ PURI M., RALESCU D. (1981)
Differentielle d'une fonction floue, C.R.A.S. 293, Série I,
 237-239
- /23/ PURI M., RALESCU D. (1982)
Fuzzy Random Variables, Indiana University, Bloomington.
- /24/ RAO G.P., RUTHERFORD D.A. (1981)
Approximate reconstruction of mapping functions from linguistic description, in problems of fuzzy logic applied to System Control, Cybernetics and Systems 12, 225-236.
- /25/ SANCHEZ E., (1976)
Resolution of composite fuzzy relation equations.
Inf. & Control, 30, 38-48
- /26/ SANCHEZ E. (1977)
Solution of composite fuzzy relation equations : application to medical diagnosis in Browerian logic in "Fuzzy Automata and Decision Processes" (M.M. Gupta, G.N. Saridis, B.R. Gaines eds) North Holland 221-234.
- /27/ TANAKA H., UEJIMA S., ASAI K. (1981)
Fuzzy linear regression model, in Applied Systems and Cybernetics (G.E. Lasker Ed.) vol.VI, Pergamon Press, 2933-2939.
- /28/ TANAKA H., UEJIMA S., ASAI K (1982)
Linear regression analysis with fuzzy model.
IEEE Trans. Systems Man and Cybernetics, 12, 903-907
- /29/ TONG R.M. (1978)
Synthesis of fuzzy models for industrial processes : some recent results.
Int. J. General Systems, 4, 143-162
- /30/ YAGER R.R. (1979)
A note on probabilities of fuzzy events.
Information Sciences 18, 113-129
- /31/ YAGER R.R. (1982)
Probabilities from fuzzy observations.
Tech. Rep. MII -207, Iona College, New Rochelle, N.Y.

- /32/ ZADEH L.A. (1968)
Probability measures of fuzzy events.
J. Math. Anal. and Appl. 10, 421-427
- /33/ ZADEH L.A. (1973)
*Outline of a new approach to the analysis of complex systems
and decision processes.*
IEEE Trans. on Systems, Man, and Cybernetics 3, 28-44.
- /34/ ZADEH L.A. (1979)
*Fuzzy Sets and information Granularity in Advances in Fuzzy
Set Theory (M.M. Gupta, R.K. Ragade, R.R. Yager, Eds) -
North-Holland, 3-18.*



PARTIE II

APPLICATIONS

CHAPITRE 6.

MODÈLES DE PROCESSUS DÉCISIONNELS

"So the second change we propose in utility theory is to admit that we shall be dealing with fuzzy subjective probabilities, not sharp objective ones".

R.D. Luce H. Raiffa Games and Decisions

(Wiley 1957) p. 373.

Le but de n'importe quelle théorie de la décision est de proposer une méthode et des outils mathématiques pour l'analyse rationnelle des choix que peut faire un individu (ou un groupe d'individus) confronté à une situation donnée, en vue de lui suggérer quelles sont les "meilleures" options. Une situation de choix est généralement décrite à l'aide des éléments suivants :

- Un environnement E que l'on peut caractériser par ses états $e \in E$
- Un ensemble d'actions possibles A dans l'environnement E .
- Un ensemble de conséquences possibles X des actions. Ces conséquences dépendent de l'état dans lequel se trouve l'environnement quand l'action est effectuée.
- Des critères d'évaluation des actions. Les actions seront en fait jugées sur leurs conséquences.
- Des contraintes qui limitent le domaine des actions, et qui sont elles aussi généralement formulées sur l'ensemble des conséquences.

Le problème d'un décideur est donc de choisir une action, compte tenu de l'état de l'environnement, en vue de produire les conséquences qu'il souhaite au sens des critères et des contraintes qu'il s'est fixé. Il est donc supposé détenir des connaissances sur

- . ses préférences, au sens de chaque critère, quant aux conséquences de son action
- . la façon d'agréger les critères
- . l'état réel de l'environnement
- . les conséquences de son action.

Dans ce chapitre, on tente de montrer l'intérêt de certains des outils mathématiques introduits précédemment pour développer de nouvelles méthodes d'agrégation de critères, et d'évaluation en environnement incertain ou face à des conséquences imprécisément perçues. On supposera que les conséquences, évaluées selon chaque critère forment un ensemble muni d'une relation de pré-ordre complet, comme c'est le cas dans les modèles de type utilité /27/.

Néanmoins, l'incertitude et l'imprécision se répercuteront sur l'évaluation globale des actions ; c'est-à-dire que cette évaluation globale donnera naissance à une ou plusieurs relations floues de préférence sur A , et la décision finale se fera au vu de ces relations floues.

Trois sections composent ce chapitre : la première est consacrée à l'utilisation des opérateurs ensemblistes flous étudiés au chapitre 2 pour l'agrégation de critères. La seconde cherche à exploiter certaines classes de mesures d'incertitude introduites au chapitre 1 (notamment les mesures de possibilité), et l'arithmétique floue développée au chapitre 4 pour la modélisation de la connaissance partielle de l'état de l'environnement par le décideur. La troisième section traite le problème du classement d'un ensemble d'actions évaluées de façon imprécise.

Le but de ce chapitre n'est pas de présenter en détail une étude de cas (même si certaines applications pratiques seront évoquées). Il ne prétend pas être exhaustif en développant d'un bout à l'autre une théorie de la décision. Il cherche surtout à démontrer l'intérêt des outils mathématiques décrits précédemment, quant à leur capacité à décrire et résoudre des situations nouvelles non prises en compte par les méthodes cardinales classiques fondées sur la notion d'utilité et de probabilité.

6.1 - MODÈLES D'AGRÉGATION DE CRITÈRES DÉCRITS PAR DES ENSEMBLES FLOUS

Dans cette section, on suppose connu l'état de l'environnement, de sorte que chaque action a possède une conséquence connue décrite par une suite de points $(m_1(a), m_2(a), \dots, m_p(a))$. $m_i(a)$ est une mesure (objective) de l'action a au sens du critère $n^{\circ}i$, par exemple un coût, une taille, une consommation, etc... m_i définit une application de A dans une échelle objective X_i , et l'ensemble X des conséquences est identifié avec le produit cartésien $X_1 \times X_2 \times \dots \times X_p$.

L'échelle de mesure X_i est soit une échelle continue, soit une échelle discrète, éventuellement finie. C'est dans ce cadre que Bellman et Zadeh /2/ ont formulé, en termes ensemblistes flous, leur modèle d'agrégation de critères, qui sert de point de départ à l'approche présentée ci-dessous. Notre apport consiste à proposer un panorama de fonctions qui sont autant de façons possibles d'agréger des critères, ainsi qu'une méthodologie pour déterminer, dans une situation donnée, quelle est la fonction qui agrège les critères en respectant au mieux la façon dont le décideur évalue globalement ses actions potentielles. Cette démarche s'étend facilement au cas où l'on

ne dispose pas de renseignements précis sur les mesures $m_i(a)$ qui sont alors des sous-ensembles de X_i , éventuellement flous. On tentera de montrer les relations qui existent entre cette approche et la théorie multi-dimensionnelle de l'utilité (Keeney et Raiffa, /27/)

6.1.1 - REPRESENTATION D'UN CRITERE PAR UN ENSEMBLE FLOU (Bellman et Zadeh, 1970).

Supposons dans un premier temps $X = X_1$. Afin de posséder plus d'information quant aux souhaits que peut formuler un décideur sur les conséquences de son action, on lui demande, pour chaque valeur $x \in X$ de fournir une évaluation quant à l'opportunité de décider d'effectuer l'action $a \in A$ telle que $m(a) = x$. On suppose ici que les préférences du décideur se traduisent par un préordre complet c'est-à-dire

$$(I) \left\{ \begin{array}{l} \forall x, x \text{ est indifférent à } x, \text{ que l'on note } x \sim x \\ \forall x, x' \neq x \text{ est soit préféré, soit indifférent, soit préférable à } x' \\ \text{que l'on note } x \succsim x' \text{ ou } x' \succsim x \\ \text{le décideur est cohérent : } \succsim \text{ est une relation transitive} \end{array} \right.$$

Questionner le décideur sur l'opportunité de chaque conséquence n'est certainement pas toujours possible. En fait, et ce surtout dans le cas d'échelles X continues, il lui sera plus naturel de s'exprimer linguistiquement. Par exemple, si $A =$ ensemble de voitures, $X =$ échelle de prix, le décideur souhaitera choisir une voiture "de coût modéré". Il est donc naturel de modéliser le souhait du décideur quant aux conséquences de son choix par un ensemble flou G sur X .

On appellera objectif du décideur selon le critère m , un ensemble flou G sur X tel que $\forall x \in X$, $\mu_G(x)$ est le degré de compatibilité entre la valeur x décrivant la conséquence, et l'objectif du décideur. G doit être vu comme un point de mire /34/ flou. Dans la mesure où les axiomes (I) sont vérifiés on peut choisir pour $\mu_G(x)$ un domaine qui soit une échelle numérique, $[0, 1]$ par convention, avec l'interprétation suivante :

$$\mu_G(x) = 1 \leftrightarrow x \text{ est une conséquence parfaitement compatible avec l'objectif}$$

$$\mu_G(x) > \mu_G(x') \leftrightarrow x \text{ est préféré à } x'$$

$$\mu_G(x) = 0 \leftrightarrow x \text{ est une conséquence à rejeter absolument.}$$

donc toutes les conséquences situées dans le noyau $\overset{\circ}{G}$ sont indifférentes les unes aux autres (et sont les préférées) ainsi que toutes celles qui figurent hors du support $S(G)$ (mais sont également rejetées).

On notera que le rôle de la fonction d'appartenance μ_G est de traduire numériquement l'intensité des préférences du décideur entre les diverses mesures des conséquences x , et par là, joue un rôle semblable aux seuils d'indifférence et de présomption de préférence suggérés par Roy /34/ pour classer les types de critères.

A titre d'exemple, l'objectif "coût modéré" peut être représenté comme dans la figure ci-dessous.

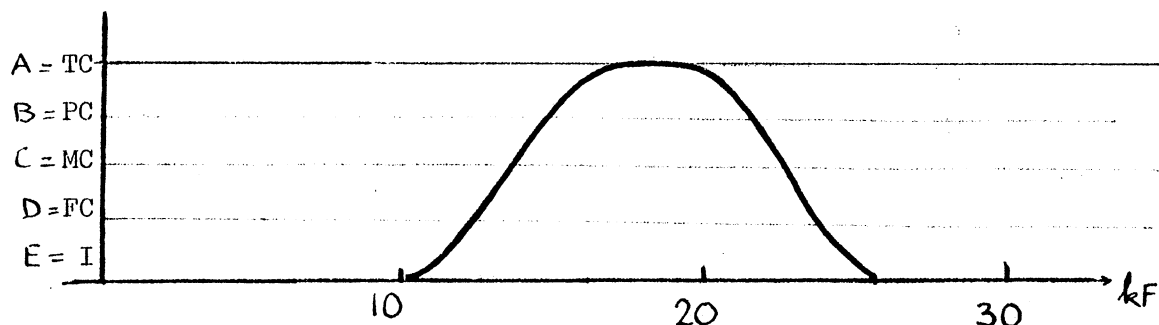


FIGURE 1

L'estimation de μ_G ne peut se faire de façon exacte. Néanmoins la forme de la courbe permet d'exprimer certains comportements du décideur. Pour saisir ce comportement il n'est pas question de lui demander de s'exprimer à l'aide de l'échelle $[0,1]$ (qui est arbitraire !) mais sur une échelle discrète de notation comportant de 5 à 7 niveaux selon le seuil de perception du décideur. Des études expérimentales ont montré que au delà de 7 niveaux, on comprend mal la nuance de sens entre deux niveaux consécutifs. Une idée simple est d'exprimer linguistiquement les niveaux de compatibilité entre objectif et conséquence, et de projeter ces niveaux sur $[0,1]$, comme sur le tableau ci-dessous (voir aussi fig.1):

	Niveau de compatibilité conséquence-objectif	Convention numérique
A	Totalement compatible	1
B	Plutôt compatible	0.75
C	Moyennement compatible	0.5
D	Faiblement compatible	0.25
E	Incompatible	0

Pour l'identification de μ_G on peut penser à deux approches (voir par exemple Dubois, /10/).

- discrétiser X en un ensemble fini X' et demander au décideur de fournir une appréciation ($\in \{A,B,C,D,E\}$) de chaque conséquence $x' \in X'$, puis lisser le résultat obtenu
- représenter G sous forme d'un nombre flou de type L-R (cf. chapitre 4) dont le décideur fournit directement les paramètres (en fixant les limites de $\overset{\circ}{G}$ et $S(G)$) et la forme, parmi un ensemble de fonctions-types L et R (par questionnaire sur des conséquences appartenant à $S(G)-\overset{\circ}{G}$, on peut trouver la forme la plus fidèle).

La figure ci-dessous donne des exemples de fonctions types (voir aussi Dubois /10/).

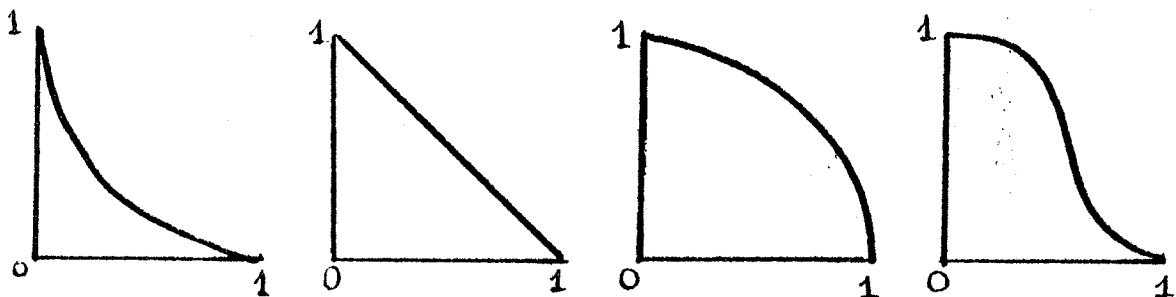


FIGURE 2

On peut définir, connaissant l'objectif G et le critère m , un jugement de chaque action $a \in A$, quant à son opportunité relativement à l'objectif G par la fonction d'appartenance μ_{G_j} telle que

$$\mu_{G_j}(a) = \mu_G(m(a)) \quad (1)$$

Il est clair que le caractère réaliste de ce type de modèle est lié à la nature du critère envisagé. (1) est raisonnable pour des critères élémentaires, c'est-à-dire définis sur une échelle objective évidente (coût, dimension ...) ; pour des critères plus complexes, l'échelle n'est plus unique (sécurité, robustesse, ...) voir objectivement inexistante (beauté). Dans ce cas il faut utiliser

- soit une approche hiérarchique par décomposition en critères élémentaires ; c'est l'approche envisagée ci-dessous
- soit une approche ordinale, par construction d'une relation de préférence, et analyse de celle-ci (Roy /33/).

Comme l'ont fait remarquer Bellman et Zadeh /2/, une contrainte, définissant un sous-ensemble de conséquences admissibles, ne se distingue mathématiquement plus d'un objectif, tel qu'il est défini par l'équation (1). On peut convenir d'appeler contrainte, un partage strict de X en deux sous-ensembles usuels G et \bar{G} où G est le domaine admissible. L'ensemble des actions possibles se définit bien par l'équation (1). En fait le concept d'objectif proposé ici, d'après Bellman et Zadeh /2/ correspond à l'idée de contrainte souple, et $\mu_G(x)$ peut être vu comme une évaluation du respect de la contrainte par la conséquence x . Cette démarche a d'ailleurs été utilisée par certains chercheurs apparemment isolés (Kendall /28/ par exemple).

6.1.2 - ENSEMBLE FLOU ET FONCTION D'UTILITE

Il est clair que le concept d'utilité, en tant que représentation cardinale des préférences, est très proche de celui d'ensemble flou, et que la fonction d'appartenance d'un ensemble flou joue le rôle d'une fonction d'utilité (normalisée).

Néanmoins la fonction d'utilité, telle qu'elle a été introduite par Von Neumann et Morgenstern /43/, et reprise par Savage /38/, Keeney et Raiffa /27/ etc... présente des particularités qui font qu'un ensemble flou n'est pas toujours stricto sensu une fonction d'utilité déguisée.

D'abord, une fonction d'utilité est toujours définie sur un continuum. En revanche la fonction d'appartenance d'un objectif, telle qu'elle est définie à la section précédente peut très bien être construite sur un

référentiel X fini. L'approche proposée à la section suivante pour l'agrégation des critères ne présupposera rien sur la nature de X .

Ensuite la fonction d'utilité a toujours une connotation monétaire. X est très souvent évalué par une échelle de coûts et on est toujours supposé exprimer l'intérêt que présente une action par une somme d'argent équivalente. De par sa nature abstraite, la fonction d'appartenance est neutre.

Enfin, et c'est peut être le point important, la fonction d'utilité fait largement appel à la théorie des probabilités pour la spécification des axiomes qui la caractérisent (voir le livre de Von Neumann et Morgenstern¹). La notion d'"équivalent certain" est fondée sur l'hypothèse que si deux conséquences $x < x'$ sont données avec leur probabilités d'occurrence p et $p' = 1-p$ respectivement, alors il existe $x'' \in]x, x'[$ tel que le décideur considère que cette situation incertaine est indifférente à la certitude que x'' se produira. Les ensembles flous ne présupposent rien sur le caractère certain ou incertain des objets du référentiel où ils sont définis.

Au lieu de faire appel au concept de probabilité, on peut, pour établir l'unicité d'une fonction d'appartenance, à une classe de transformation près, munir le produit cartésien $X \times X$ d'une relation de préférence qui soit, elle aussi un pré-ordre complet, en plus de la relation sur X qui compare les conséquences entre elles.

Cette approche, originellement proposée indépendamment dans Krantz et al. /29/, adaptée par Norwich et Turksen /30/, et rappelée au début du chapitre 2, est beaucoup plus satisfaisante qu'un recours quelque peu artificiel à la notion de probabilité. En substance, on montre que si on peut comparer les préférences relatives sur un continuum X , et si la relation de préférence relative (qui compare des paires (x, x') entre elles) satisfait certains axiomes de structure, alors il existe une fonction d'évaluation, unique à une transformation affine près, à valeur sur un intervalle réel, qui permet de classer les objets de X . Cette fonction d'évaluation là correspond mieux à l'idée d'ensemble flou que la fonction d'utilité de Von Neumann.

1) Cf aussi le § 2.1 de ce chapitre.

6.1.3 - METHODOLOGIE POUR L'AGREGATION DE CRITERES

(Dubois et Prade /12/).

Supposons l'espace X des conséquences identifié à un produit cartésien $X_1 \times X_2 \dots \times X_p$. On a donc p critères élémentaires associés à des indices de mesure m_1, \dots, m_p , où m_i est une fonction $A \rightarrow X_i$ qui mesure les actions selon le critère i . L'ensemble \mathcal{D} des décisions acceptables est défini comme un ensemble flou de A obtenu par agrégation des ensembles de meilleures actions \mathcal{U}_i déduits des objectifs partiels G_i , $i = 1, p$, construits par (1). La fonction d'appartenance $\mu_{\mathcal{D}}$ est donc telle que

$$\forall a, \mu_{\mathcal{D}}(a) = h(\mu_{G_1}(m_1(a)), \dots, \mu_{G_p}(m_p(a))) \quad (2)$$

On est donc ramené à trouver un opérateur ensembliste flou H qui combine les objectifs $G_1 \dots G_p$. Le chapitre 2 constitue le fondement théorique de la méthodologie d'agrégation de critères, puisque, dans le cas $p = 2$, il a présenté un éventail assez large de possibles pour le choix de H , avec de surcroît une interprétation sémantique. Dans les paragraphes ci-dessous, on développe l'approche par les ensembles flous, en évoquant successivement le cas de 2 critères, de $p > 2$ critères, le problème (pratique) d'identification des opérations, et celui des objectifs d'importance inégale. Ce type d'approche a été indépendamment adopté par Dubois et Prade /12/ et Zimmermann et Zysno /5/.

6.1.3.1 - Conditions minimales sur l'opération h.

On trouvera naturel d'exiger que l'opération h satisfasse les conditions suivantes, pour être crédible :

A1 Conditions aux limites $h(0,0\dots 0) = 0$; $h(1,1\dots 1) = 1$

A2 $\forall (u_i, v_i) \in [0,1]^2$, si $u_i \geq v_i, \forall i$, alors

$$h(u_1 \dots u_p) \geq h(v_1 \dots v_p)$$

A1 exprime qu'une action totalement compatible (resp : incompatible) avec chacun des objectifs est une décision totalement acceptable (resp : à rejeter absolument).

A2 indique que h ne doit pas contrarier l'ordre partiel défini par les vecteurs $(\mu_{\mathcal{U}_1}(a) \dots \mu_{\mathcal{U}_p}(a))$ sur l'ensemble A . Si \mathcal{D} est le noyau de \mathcal{D} ,

c'est-à-dire l'ensemble des meilleures actions induit par l'agrégation h , et si \mathcal{M} est l'ensemble des éléments maximaux de l'ordre partiel sur A , on doit avoir $\mathcal{D} \subseteq \mathcal{M}$

On s'intéressera dans la suite aux opérateurs d'agrégation symétriques c'est-à-dire :

A3 : h est une fonction symétrique de ses arguments.

On notera que le caractère symétrique de h ne signifie pas que l'agrégation obtenue soit symétrique : en effet ce qu'on agrège ce sont les ensembles flous G c'est-à-dire qu'on peut très bien introduire une certaine forme de dissymétrie entre les critères en jouant sur les fonctions d'appartenance μ_G . L'axiome A3 est justifié quand les objectifs sont d'importance égale et peuvent donc s'interchanger dans le processus d'agrégation. Dire que les objectifs sont d'égale importance ne signifie pas que les critères (= les m_i) aient tous la même criticité. On module celle-ci à l'aide des G_i . Enfin on suppose que h est continu.

6.1.3.2 - Agrégation de deux critères

On suppose $p = 2$. Quand les objectifs sont définis de façon tout ou rien, c'est-à-dire que G_i est un sous-ensemble de X_i (qui revient à une contrainte) les deux seules agrégations possibles sont la conjonction des objectifs ($D = G_1 \cap G_2$) ou la disjonction (s'ils sont redondants) ($D = G_1 \cup G_2$). Dans ce cas on exclut tout compromis entre les deux critères. Quand les objectifs deviennent nuancés, le compromis devient une des attitudes naturelles d'un décideur. Néanmoins les deux autres attitudes, à savoir la satisfaction simultanée des deux objectifs ou la redondance de l'un d'eux, sont aussi très naturelles. Il est à noter que la théorie de l'utilité s'affirme essentiellement comme une analyse de la notion de compromis (cf. le titre de Keeney et Raiffa /27/). Avec l'approche par les ensembles flous on obtient donc (et ceci est lié à l'utilisation d'un intervalle de calibration $[0,1]$ unique) un enrichissement sémantique par rapport à la théorie de l'utilité, puisqu'on peut représenter trois attitudes fondamentales du décideur devant l'agrégation de critères. A côté de ces trois attitudes "pures" on peut imaginer des attitudes hybrides. Prenons l'exemple suivant : soit une action potentielle "a".

- . Si "a" est mal notée au sens des deux critères, elle bénéficie d'une attitude relativement indulgente (redondance admise)
- . Si a est bien notée au sens du premier critère et mal notée au sens du second, le jugement global est un compromis.
- . Si a est bien notée au sens des deux critères alors le jugement global est sévère (conjonction).

Ce type d'attitude se rencontre chez certains enseignants qui doivent noter leurs élèves. C'est l'idée d'agrégation dépendant des niveaux d'évaluations partielles qu'on met ici en valeur.

Un cas d'attitude hybride que l'on va bien savoir modéliser est celui où on demande que le classement final des actions potentielles ne dépende pas du fait qu'on ait agrégé les ensembles de bonnes actions au sens des deux critères (\mathcal{Q}_1 et \mathcal{Q}_2) ou les ensembles de mauvaises actions ($\overline{\mathcal{Q}}_1$ et $\overline{\mathcal{Q}}_2$). Dans ce dernier cas l'ensemble des actions globalement préférées est obtenu par élimination des actions globalement mauvaises ainsi construites.

On peut utiliser directement les résultats du chapitre 2 (section 3) pour définir les opérations d'agrégation évoquées ci-dessus.

Pour une opération exprimant la satisfaction simultanée des objectifs un axiome naturel est :

$$A4 : \forall u, v \quad h(u,v) \leq \min(u,v)$$

c'est-à-dire que l'évaluation globale d'une action ne peut être meilleure que la plus mauvaise des évaluations partielles. On appellera ces opérations "conjonctions". Une sous-classe importante est l'ensemble des conjonctions associatives qui ont servi de modèle aux intersections d'ensembles flous (chap.2 section 3.2). La propriété d'associativité est importante pour savoir étendre ces opérations pour la combinaison de $n > 2$ critères. On rappelle les principales conjonctions associatives (cf. chap.2).

$$\min(u,v) ; u.v ; \max(0,u+v-1) ; T_w(u,v).$$

$f^*(f(u)+f(v))$ où f est une fonction décroissante de $[0,1]$ sur \mathbb{R}^+ telle que $f(1) = 0$. f^* est la pseudo-inverse de f .

Une opération exprimant la redondance de deux objectifs satisfera un axiome dual de A4 :

$$A5 : \forall u, v \quad h(u, v) \geq \max(u, v)$$

C'est-à-dire que c'est la meilleure des évaluations partielles qui conditionne l'évaluation globale. On appellera ces opérations "disjonctions". Une sous-classe importante est celle des disjonctions associatives servant de modèle aux unions d'ensembles flous (chap.2 section 3.2), et dont les principales sont

$$\max(u, v) ; u+v-uv ; \min(1, u+v) ; T_W^*(u, v)$$

$$\varphi^*(\varphi(u) + \varphi(v)) \quad \text{où } \varphi \text{ est une fonction croissante de } [0, 1] \text{ dans } \mathbb{R}^+ \\ \text{telle que } \varphi(0) = 0.$$

Une opération exprimant le compromis satisfera l'axiome suivant, complémentaire de A4 et A5

$$A6 : \min < h < \max$$

C'est-à-dire que l'évaluation globale se situe à un niveau intermédiaire entre les évaluations partielles. On est ramené aux opérations de moyenne (cf. chap.2). On sait que les seules moyennes associatives sont de la forme :

$$h_\alpha(u, v) = \text{mediane}(u, v, \alpha) \quad , \quad \alpha \in]0, 1[$$

Pour être capable d'étendre h à plus de 2 arguments on est amené à utiliser la propriété de bisymétrie, plus faible que l'associativité (cf. chap.2, § 3.4.1) ; les principales opérations sont alors les moyennes harmoniques géométriques, arithmétiques, etc... qui toutes sont de la forme

$$h(u, v) = f^{-1} \left(\frac{f(u) + f(v)}{2} \right) \quad \text{où } f \text{ est une fonction monotone de } [0, 1] \text{ dans } \mathbb{R}.$$

Enfin l'agrégation qui demeure invariante selon que l'on opère sur les objectifs ou sur leurs complémentaires sera modélisée par les sommes symétriques étudiées au chapitre 2 § 3.4.2, et leur axiome de base est

$$A7 : h(u, v) = 1 - h(1-u, 1-v)$$

On a vu que certains de ces opérateurs sont des moyennes

$(\frac{\min(u,v)}{1-|u-v|}, \frac{\max(u,v)}{1+|u-v|}, \frac{u+v}{2}, \text{médiane}(u,v,1/2))$. En revanche toutes les sommes symétriques associatives (sauf la médiane) expriment une agrégation hybride, à savoir :

- . si $\max(u,v) < \frac{1}{2}$ $h(u,v) \leq \min(u,v)$
- . si $\min(u,v) > \frac{1}{2}$ $h(u,v) \geq \max(u,v)$
- . si $u \leq \frac{1}{2} \leq v$ $u \leq h(u,v) \leq v$

Des exemples de sommes symétriques associatives sont

$$\frac{uv}{uv+(1-u)(1-v)} ; f^{-1}(f(u)+f(v))$$

où f est de la forme $f(u) = h(u) - h(1-u)$ et h est n'importe quelle fonction continue décroissante $[0,1] \rightarrow \mathbb{R}^+$ telle que $h(0) = +\infty$. Voir chapitre 2, § 3.4.2 pour d'autres résultats. On peut y noter des sommes symétriques dont le comportement d'agrégation hybride est différent (par exemple σ_+).

L'intérêt de posséder des représentations fonctionnelles pour ces opérations est de pouvoir disposer de familles paramétrées ce qui va faciliter l'identification des opérations. Voir annexe pour des exemples de familles paramétrées.

Signalons enfin que les trois façons d'agréger ont été axiomatiquement mises en évidence par Fung et Fu /23/. Néanmoins, leur système d'axiomes contient les hypothèses d'idempotence ($h(u,u) = u$) et d'associativité simultanément, ce qui les mène à ne trouver comme solutions que les médianes $\text{med}(u,v,\alpha)$, $\alpha \in [0,1]$. Conjonction et disjonction sont obtenues pour $\alpha = 0$ et 1 respectivement.

6.1.3.3 - Agrégation de n critères par combinaison symétrique des objectifs.

Les opérations évoquées ci-dessus peuvent s'étendre au cas de $n > 2$ objectifs d'égale importance, pourvu qu'elles possèdent une structure qui se prête "naturellement" à une augmentation du nombre de leurs arguments, qui préserve l'axiome de commutativité A3.

Les conjonctions et disjonctions associatives en sont de tels exemples.

Versions à n arguments des opérateurs d'agrégation rencontrés

Nature	Opération binaire	Extension
Conjonction	$\max(0, u+v-1)$	$\max(0, \sum_{i=1}^n u_i - n + 1)$
Conjonction	$u \cdot v$	$\prod_{i=1}^n u_i$
Conjonction	$\min(u, v)$	$\min_{i=1, n} u_i$
Compromis	$\text{med}(u, v, \alpha)$	$\text{med}(\max_{i=1, n} u_i, \min_{i=1, n} u_i, \alpha)$
Compromis	$\frac{2xy}{x+y}$	$(n \prod_{i=1}^n u_i) / (\sum_{i=1}^n u_i)$
Compromis	$\sqrt{u, v}$	$\sqrt[n]{u_1 \dots u_n}$
Compromis	$\frac{\min(u, v)}{1+ u-v }$	$\frac{\min_i u_i}{1 - \max_i u_i + \min_i u_i}$
Compromis	$\frac{u+v}{2}$	$\frac{1}{n} \sum_{i=1}^n u_i$
Compromis	$1 - \sqrt{(1-u)(1-v)}$	$1 - \sqrt[n]{\prod_{i=1}^n (1-u_i)}$
Compromis	$\frac{\max(u, v)}{1+ u-v }$	$\frac{\max_i u_i}{1 - \min_i u_i + \max_i u_i}$
Disjonction	$\max(u, v)$	$\max_i u_i$
Disjonction	$u+v-uv$	$1 - \prod_{i=1}^n (1-u_i)$
Disjonction	$\min(1, u+v)$	$\min(1, \sum_{i=1}^n u_i)$
Somme symétrique	$\sigma_-(u, v)$	$\left\{ \frac{1}{1 + \frac{\prod_{i=1}^n (1-u_i)}{n}} \right.$
Somme symétrique	$\sigma_+(u, v)$	$\left. \frac{1}{1 + \frac{\prod_{i=1}^n u_i}{n}} \right\}$
Conjonction ou disjonction ou somme symétrique	$f^{-1}(f(u)+f(v))$	$f^{-1}(\sum_{i=1}^n f(u_i))$

Définissons récursivement l'opération à n arguments $h^{(n)}$ par

$$\begin{aligned} h^{(2)} &= h \\ h^{(n)}(u_1, u_2, \dots, u_n) &= h(h^{(n-1)}(u_1, u_2, \dots, u_{n-1}), u_n) \end{aligned} \quad (3)$$

Si h est associatif et commutatif alors $h^{(n)}$ vérifie A3 et est l'extension naturelle de h pour l'agrégation de n critères. On peut utiliser (3) sur les conjonctions, disjonctions et sommes symétriques associatives, et les médianes paramétrées.

Pour les opérateurs de compromis autre que la médiane, c'est la propriété de bisymétrie qui sera utile pour l'extension à n arguments. En général on aura donc :

$$h(u_1 \dots u_n) = f^{-1}\left(\frac{1}{n} \sum_{i=1}^n f(u_i)\right) \quad (4)$$

Certaines sommes symétriques (non associatives éventuellement) ont aussi une extension naturelle, que l'on peut définir grâce à la forme canonique décrite au chapitre 2 :

$$\exists g, \quad h(u, v) = \frac{g(u, v)}{g(u, v) + g(1-u, 1-v)}$$

Si la fonction g est associative on a tout naturellement

$$h(u_1 \dots u_n) = \frac{g^{(n)}(u_1, u_2 \dots u_n)}{g^{(n)}(u_1, u_2, \dots, u_n) + g^{(n)}(1-u_1, 1-u_2, \dots, 1-u_n)} \quad (5)$$

où $g^{(n)}$ est défini par (3). (5) suffit à générer les formes n -aires de toutes les sommes symétriques présentées au chapitre 2.

Le tableau ci-contre donne les extensions des principales opérations rencontrées dans ce mémoire.

Il faut remarquer que l'on peut penser à d'autres types d'agrégation symétriques d'objectifs, mettant en jeu plusieurs types d'opérations. Par exemple la spécification : deux des trois objectifs G_1, G_2, G_3 doivent être atteints correspondra à une agrégation de type

$$H(G_1, G_2, G_3) = (G_1 \cap G_2) \cup (G_2 \cap G_3) \cup (G_1 \cap G_3).$$

6.1.3.4 - Identification des opérations

Possédant un large éventail d'opérations exprimant de nombreuses attitudes possibles du décideur, on peut envisager l'identification d'une opération par questionnaire du décideur et recherche de l'opération la plus fidèle dans un catalogue donné.

Considérons pour simplifier le cas de deux objectifs d'égale importance G_1 et G_2 à agréger. La méthode d'identification de l'opération va se fonder sur la donnée d'un sous-ensemble restreint T d'actions ou de conséquences typiques dont la conformité à chaque objectif est établie. Ces actions auront des caractéristiques contrastées, c'est-à-dire que l'une d'elle sera incompatible avec G_1 mais totalement compatible avec G_2 , une autre sera moyennement compatible avec les deux objectifs, etc...

On peut envisager deux étapes dans l'identification.

1) On se donne un catalogue restreint mais bien dispersé d'opérations simples types. On demande au décideur de formuler, à l'aide d'une échelle linguistique de jugement telle que celle qui est suggérée en 6.1.1, un jugement global sur les actions contenues dans T . Dans cette procédure un très petit nombre d'actions bien choisi peut suffire ; voir table ci-contre.

2) La première étape a déterminé une ou plusieurs opérations plausibles.

Si on veut raffiner dans la détermination d'une opération, on peut choisir une famille paramétrée d'opérations dont la sémantique correspond au type d'attitude du décideur, tel qu'on a pu l'observer dans la première étape. On requestionne le décideur à partir d'un échantillon plus étoffé d'actions de A (ou de conséquences dans X_1 x X_2). On peut alors optimiser le paramètre caractérisant l'opération à l'aide d'une méthode du genre "moindres carrés".

Si le décideur est cohérent dans ses choix, on peut alors classer par ordre de préférence les éléments de A , et ce classement doit représenter celui qu'il aurait spontanément fait si A contenait peu d'actions.

T	a ₁	a ₂	a ₃	
Objectif G ₁	I	MC	MC	Opérations sélectionnées
Objectif G ₂	TC	MC	TC	
Réponses possibles du décideur	I	I	MC	$\max(0, u+v-1)$
	I	FC	MC	$u \cdot v$
	I	MC	MC	$\min(u, v)$
	I	MC	PC	$\sqrt{u \cdot v} ; \frac{2uv}{u+v}$
	FC	MC	MC	$\text{med}(u, v, 1/4)$
	MC	MC	MC	$\text{med}(u, v, 1/2) ; \frac{\min(u, v)}{1- u-v }$
	MC	MC	PC	$\frac{x+y}{2} ; \sigma_+ ; \frac{\max(u, v)}{1+ u-v }$
	MC	MC	TC	$\sigma_.$
	PC	MC	PC	$\text{med}(u, v, 3/4)$
	TC	MC	TC	$\max(u, v) ; 1 - \sqrt{(1-x)(1-y)}$
	TC	PC	TC	$u+v-uv$
	TC	TC	TC	$\min(1, u+v)$

Exemple de procédure de sélection des opérateurs définis au § 1.3.3 à l'aide de 3 actions types.

Il semble acquis que l'identification très précise d'une telle opération sera parfois (peut-être souvent) impossible. Cela n'invalide pas l'approche pour autant, et ce pour deux raisons :

- une identification très précise peut être jugée inutile ; seul un certain comportement de l'opération sera souhaité. Cela est à rapprocher du caractère arbitraire des valeurs d'appartenance, et de la traduction numérique des jugements émis par le décideur.

- si on trouve que plusieurs opérations sont des modèles possibles, on pourra malgré tout tenter de classer les éléments de A. On pourra introduire une incertitude sur le résultat de l'agrégation. Par exemple si on n'a pas pu discriminer entre les opérations h_1, h_2, \dots, h_q on peut poser

$$\forall (u,v), h(u,v) \in \left[\min_{i=1,q} h_i(u,v), \max_{i=1,q} h_i(u,v) \right]$$

On peut même obtenir comme résultat de l'agrégation un nombre flou : ce sera le cas si parmi les opérations sélectionnées, on peut considérer que certaines sont plus crédibles que d'autres. Ce sera aussi le cas si on a utilisé pour l'estimation du paramètre une méthode de régression floue du type de celles décrites au chapitre 5. On se retrouvera alors avec un sous-ensemble flou d'opérations, qui, en utilisant le principe d'extension, fournira des évaluations floues des actions.

On notera que le choix des objectifs (c'est-à-dire la détermination des μ_{G_i}) et le choix de l'agrégation interagissent : une conjonction d'objectifs peu exigeants peut être équivalente à un compromis entre deux objectifs très sélectifs. Ce phénomène qui paraît intuitif apparaît nettement dans le modèle. Par exemple on a l'équivalence entre $(\mu_{G_1}, \mu_{G_2}, h = \text{moyenne géométrique})$ et $(\sqrt{\mu_{G_1}} \sqrt{\mu_{G_2}}, h = \text{produit})$. Une recherche systématique de telles équivalences présenterait un certain intérêt. En tout cas elles soulignent que la façon dont on combine des objectifs dépend de la façon dont on spécifie ceux-ci.

6.1.3.5 - Critères d'importance inégale.

Le concept d'importance d'un critère par rapport à un autre a été fort peu élucidé. Le sens qu'on donne à ce mot est très variable selon

les décideurs, ou selon les situations. Sans prétendre éclaircir définitivement ce problème très ardu, on va tenter de dégager quelques interprétations possibles bien distinctes du concept d'importance.

a) Utilisation du seuils de satisfaction

Une première interprétation de l'importance des critères peut être l'existence pour chaque critère i d'un seuil minimum x_i^0 qu'il faut atteindre pour considérer comme acceptable la conséquence $m_i(a)$ d'une action a . On suppose ici implicitement que l'action a est d'autant préférée au sens du critère i que $m_i(a)$ est grand. Il est clair que cette approche est un cas particulier de celle qui a été adoptée ici, à savoir les objectifs flous, puis que la connaissance du seuil x_i^0 est équivalente à la donnée d'un objectif G_i tel que $\mu_{G_i}(x) = 1 \forall x \geq x_i^0$ et 0 sinon. On voit donc que l'agrégation symétrique d'objectifs flous parvient à capturer une des interprétations possibles de l'importance des critères : plus un critère est jugé important, plus l'objectif correspondant sera pointu.

Pour faire face à d'éventuels conflits trop sévères entre les objectifs ($\mathcal{D} = \emptyset$) on peut en modifier certains en élargissant le domaine qu'ils délimitent, dans une optique de type point de mire évolutif (Zeleny /50/ ou Roy /34/). C'est-à-dire que l'on change l'ensemble flou G_i en $G_i' \supseteq G_i$.

b) Pondération des critères

Une méthode très répandue pour exprimer l'importance inégale des critères est de leur affecter un poids à chacun, et d'intégrer ces poids dans l'opération d'agrégation. Dans cette optique la combinaison convexe de critères

$$m_{\vec{W}}(a) = \sum_{i=1}^p m_i(a)W_i \quad ; \quad \sum_{i=1}^p W_i = 1 \quad (6)$$

est la plus répandue des techniques. Sa principale justification réside dans le fait que l'on peut parcourir tout l'ensemble des éléments maximaux de A au sens de p critères en faisant varier la pondération \vec{W} et en optimisant pour chaque \vec{W} le critère $m_{\vec{W}}(a)$. Néanmoins, l'évaluation a priori des poids est problématique, et ce d'autant plus que les m_i renvoient à des grandeurs

qui n'ont rien en commun (la vitesse et le coût par exemple dans le cas où l'action est le choix d'une voiture). D'ailleurs $m_i(a)$ n'est pas forcément un nombre !

Si on admet comme raisonnable la définition d'objectifs flous G_i , $i = 1, p$ pour chaque critère, et qu'on souhaite pondérer les objectifs, alors une quantité telle que

$$\sum_{i=1}^p W_i \mu_{G_i}(x) = \mu_D(x) \quad ; \quad \sum_{i=1}^p W_i = 1 \quad (7)$$

est plus satisfaisante que (6), car on agrège des grandeurs de même nature. Le coefficient W_i mesure la pertinence de l'objectif G_i , à quel point il participe à l'objectif global. On peut voir \vec{W} comme la donnée d'une fonction de probabilité sur l'ensemble $C = \{1, 2, \dots, p\}$ des critères, et W_i s'interprète comme la probabilité pour que le "vrai" (ou le bon) critère soit le $i^{\text{ème}}$. Cette conception de l'importance des critères est clairement différente de celle qui se dégage de l'existence de seuils de satisfaction plus ou moins exigeants.

On remarque que l'ensemble des opérations décrites par (7) sont des opérateurs de compromis non symétriques. On aimerait étendre ce concept de pondération à d'autres types d'agrégation. Ce sera possible quand on possède des représentations additives des opérations. Tout opérateur de compromis de la forme $f^{-1}\left(\frac{f(a)+f(b)}{2}\right)$ se généralise aisément en :

$$h_W^+(u_1 \dots u_p) = f^{-1}\left(\sum_{i=1}^p W_i f(u_i)\right) \quad (8)$$

où les W_i sont les mêmes qu'en (7).

Quand l'opération h est de la forme $f^{-1}(f(a)+f(b))$ on peut adopter le point de vue suivant. Dans l'agrégation symétrique de p objectifs chaque critère compte pour un. Dans la version pondérée on laisse libre le poids de chaque critère, à condition qu'au total, on retrouve un poids égal à p (= le nombre "équivalent" de critères est toujours p). On peut donc poser

$$h_W^+(u_1 \dots u_p) = f^{-1}\left(p \sum_{i=1}^p W_i f(u_i)\right) \quad (9)$$

(9) peut s'appliquer directement à la plupart des opérations de conjonction

disjonction, et à de nombreuses sommes symétriques. Pour ces dernières on appliquera (9) à la fonction g qui les engendre (cf. eqn (5)) lorsque cette fonction est associative.

On remarque que dans le cas où h est une co-norme ou une t-norme stricte (9) revient à modifier les objectifs au sens de (a). On retrouve, pour le produit, l'approche empiriquement proposée par Yager /46/. Un cas important qui "sort de l'épure" est celui des opérateurs min et max ; ce cas est résolu à l'aide de la notion de pondération généralisée suggérée maintenant.

c) Pondération généralisée

Définissons l'ensemble flou F^a des objectifs atteints par l'action a par :

$$\mu_{F^a}(i) = \mu_{G_i}(m_i(a)) \quad \forall i = 1, p \quad (10)$$

Si l'on agrège l'ensemble des objectifs par (7) il est clair que l'évaluation globale de l'action a , soit $\mu_{\mathcal{D}}(a)$, est la probabilité pour que l'action "a" satisfasse le bon critère. C'est, au sens de Zadeh, la probabilité d'un événement flou (cf. chap. 2 § 4.3.1) i.e.

$$\mu_{\mathcal{D}}(a) \triangleq P(F^a) \quad (11)$$

On notera que l'agrégation de type (7) est toujours une opération de compromis ce qui en indique les limitations.

L'intérêt de l'interprétation de (7) en termes de probabilité d'un événement flou n'est pas purement formel. On a vu au chapitre 2 qu'on peut définir d'autres types de mesures d'incertitude que la probabilité pour des événements flous. L'emploi de telles mesures se justifie pleinement dans le contexte de l'agrégation de critères, par le fait que les poids W_i sont donnés de façon subjective par le décideur, et aussi parce que la combinaison convexe n'exprime qu'une attitude possible pour celui-ci : le compromis.

On appellera pondération généralisée toute mesure floue g , au sens de Sugeno, définie sur l'ensemble C des critères. Si S est un sous-ensemble de critères la quantité $g(S)$ évaluera l'importance relative de ce

groupe de critères par rapport aux autres. L'idée d'utiliser une mesure floue pour pondérer les critères est à porter au crédit de Sugeno (voir Terano et Sugeno /42/). Les extensions des mesures d'incertitude à des événements flous, décrites au chapitre 2, section 4.3 vont maintenant être utiles à la définition de $\mu_{\mathcal{D}}(a)$, comme mesure de l'événement flou F_a .

Soit $\{\alpha_1 < \dots < \alpha_q\} = \mu_{F_a}(C)$ l'ensemble des évaluations distinctes de a . Sugeno /41/ utilise son concept d'intégrale floue, qui est une médiane généralisée (cf. chap. 2 § 4.3.4), pour calculer l'évaluation globale.

Soit :

$$\mu_{\mathcal{D}}(a) = \text{med} \{ \alpha_i, i=1, q \} \cup \{ g(F_{\alpha_i}^a) \}, i = 2, q \quad (12)$$

où $F_{\alpha_i}^a$ est l' α_i -coupe de F^a

La quantité

$$\bar{\mu}_{\mathcal{D}}(a) = \text{med} \{ \alpha_i, i=1, q \} \cup \{ \bar{g}(F_{\alpha_i}^a) \}, i = 2, q \quad , \text{ où } \bar{g}(A) = tg(\bar{A}) \quad (13)$$

est aussi à considérer dans cette optique (cf. chap.2 § 4.3.5) (12)-(13) est justifié surtout quand \vec{W} définit une mesure de possibilité, vue comme l'ensemble flou des critères pertinents.

Derrière (12) et (13) on sous-entend que le concept de médiane entre plusieurs évaluations partielles est un bon évaluateur global. Néanmoins, en raison de son insensibilité aux valeurs extrêmes, on peut préférer une agrégation de type moyenne. Dans le cas d'une pondération généralisée qui se définirait par des poids de pertinence répartis de façon probabiliste sur des sous-ensembles de critères, g est alors une fonction de Dempster-Shafer (cf. chap.1 section 4), c'est-à-dire une probabilité supérieure ou inférieure. On ne possède alors plus sur $\mu_{\mathcal{D}}(a)$ que des bornes supérieure et inférieure (cf. chap.2, § 4.3.3).

$$\mu_{\mathcal{D}^+}(a) = E^*(\mu_{F_a}) = \sum_{i=1}^q \alpha_i (P^*(F_{\alpha_i}^a) - P^*(F_{\alpha_{i+1}}^a)) \quad (14)$$

$$\mu_{\mathcal{D}^-}(a) = E_*(\mu_{F_a}) = \sum_{i=1}^q \alpha_i (P_*(F_{\alpha_i}^a) - P_*(F_{\alpha_{i+1}}^a)) \quad (15)$$

(14) et (15) peuvent s'utiliser avec des mesures monotones quelconques (cf. chap.2 § 4.3.5). Un dernier point de vue généralisant le concept de

pondération est le cas où g est une mesure d'incertitude décomposable au sens d'une norme ou d'une co-norme triangulaire. Dans le cas d'une pseudo-probabilité on obtient (cf. chap.2 § 4.3.4)

$$\exists \hat{\varphi} : [0, 1] \rightarrow \mathbb{R}^+, \hat{\varphi}(0) = 0, \hat{\varphi}(1) = 1$$

$$\hat{\mu}_{\mathcal{D}}(a) = \sum_{i=1}^p \mu_{F^a}(i) \hat{\varphi}(g_i) \quad (16)$$

et

$$\sum_{i=1}^n \hat{\varphi}(g_i) = 1$$

Dans ce cas les poids W_i valent $\hat{\varphi}(g_i)$;

En particulier dans le cas d'une mesure g_λ de Sugeno (ch.1 ; § 3.1) telle que $\forall A, B, A \cap B = \emptyset \rightarrow g(A \cup B) = g(A) + g(B) + \lambda g(A)g(B), \lambda > -1$, (16) s'écrit :

$$\lambda \mu_{\mathcal{D}}(a) + 1 = \prod_{i=1}^p (1 + \lambda W_i)^{\mu_{F^a}(i)}$$

où λ joue le rôle de paramètre (subjectif) évaluant la discordance avec l'agrégation linéaire ($\lambda = 0$).

Il n'est pas facile de situer ce type d'agrégation par rapport à la classification définie dans les sections précédentes, en raison de son caractère complexe. Notons néanmoins que les opérateurs \min et \max se généralisent naturellement à l'aide de la médiane de Sugeno qui peut s'interpréter comme la possibilité d'un événement flou au sens de Zadeh (cf. chap.2 § 4.3.2) puisque (12) s'écrit aussi

$$\mu_{\mathcal{D}}(a) = \max_{i=1,p} \min(\mu_{F^a}(i), \pi(i)) = \Pi(F^a)$$

où π définit l'ensemble flou des critères pertinents et (13) vaut

$$\bar{\mu}_{\mathcal{D}}(a) = \min_{i=1,p} \max(\mu_{F^a}(i), 1 - \pi(i)) = N(F^a)$$

c'est-à-dire la nécessité de l'événement flou F^a . On voit donc que la notion de possibilité (resp : nécessité) d'un événement flou au sens de Zadeh généralise la disjonction (resp : conjonction) d'objectifs par le maximum (resp : minimum)

d) Spécification dissymétrique de l'objectif global

Il est des situations où l'objectif global est formulé de façon complexe en fonction des objectifs partiels, ceux-ci formant le niveau bas d'une hiérarchie que l'on peut décrire informellement, selon la terminologie de l'intelligence artificielle comme une arborescence et/ou. Par exemple si l'achat d'une voiture est conditionnée par des exigences telles que "une voiture très puissante, ou une voiture peu puissante mais très bon marché", on aura une structure de l'opérateur d'agrégation définie par

$$D = G_1 \cup (G_2 \cap G_3)$$

où G_1 et G_2 sont définis sur la même échelle. Un exemple d'opérateur d'agrégation pouvant rendre compte d'un tel souhait est :

$$h(u_1, u_2, u_3) = \max(u_1, u_2, u_3)$$

On voit nettement que ce type de dissymétrie

- est profondément différent de ceux évoqués plus haut
- n'existe que quand le nombre d'objectifs élémentaires est supérieur à deux.

De plus c'est principalement dans le cadre d'une hiérarchisation par regroupement des objectifs élémentaires que la distinction entre les 3 attitudes fondamentales conjonction - compromis - redondance prend tout son sens. On aura remarqué que la plupart des opérations rencontrées sont (plus ou moins) homomorphiques à une addition ; des opérateurs de conjonctions ont des analogues qui expriment le compromis, etc... si bien que l'on peut douter de l'intérêt d'agréger des critères avec $\min(a+b, 1)$ plutôt que $\frac{a+b}{2}$ ou $\max(0, a+b-1)$, hormis le fait que le premier ne saura départager les meilleures décisions et le troisième les moins bonnes. En fait l'importance du choix du bon type d'opérateur devient plus évidente quand une évaluation agrégée partielle est à son tour combinée avec d'autres évaluations.

6.1.3.6 - Liens avec d'autres approches

Il existe un certain nombre de points communs entre la théorie de l'utilité multicritère et l'approche décrite ci-dessus. Ce sont :

- . la même structure de préférence transitive sur les ensembles de conséquences.

- . l'utilisation d'une échelle numérique de valeur sur chaque critère
- . la recherche d'une fonction numérique unique pour mesurer la valeur des décisions au sens de plusieurs critères, obtenue par une opération d'agrégation entre les fonctions de valeur partielles.

On peut partager la théorie multidimensionnelle de l'utilité en deux écoles :

- celle qui ne fait pas appel à la théorie des probabilités (Debreu /7/ , Krantz et al /29/)
- celle qui utilise une structure de préférence sur l'ensemble des choix en environnement incertain, modélisés par l'ensemble des mesures de probabilité sur X (voir Keeney et Raiffa /27/, Farquhar /20/).

L'approche par les ensembles flous décrite ci-dessus est dans l'esprit de la première école. Mais des axiomatiques différentes sont adoptées. En théorie de la valeur, on suppose toujours que la détérioration d'une conséquence est contre-balancée par l'amélioration d'une autre conséquence. Si $v(x_1, x_2)$ mesure la valeur de $(x_1, x_2) \in X$, et si on suppose

$$x_1 \succeq x'_1, \quad x_2 \succeq x'_2 \rightarrow v(x_1, x_2) \succeq v(x'_1, x'_2)$$

l'hypothèse de compensation dit que $\forall (x_1, x_2) \in X$

$$\forall a > 0, \exists b > 0 \text{ tel que } v(x_1 - a, x_2 + b) = v(x_1, x_2)$$

Il est facile de voir que v doit être strictement croissante en fonction de chacun de ses arguments, pour pouvoir satisfaire l'hypothèse. L'obtention d'une fonction d'utilité additive sur $X_1 \times X_2 \dots X_n$ repose sur l'hypothèse de compensation, et une autre hypothèse dite des préférences indépendantes : si $\forall I \subseteq \{1, \dots, n\}$, la structure de préférence induite sur X_I (= produit cartésien des $X_i, i \in I$) à partir de la structure de préférence sur X , en fixant le vecteur $x_{\bar{I}} \in X_{\bar{I}}$, ne dépend pas de ce vecteur, alors $\exists v_1 \dots v_n, v$ tel que

$$v(x_1 \dots x_n) = v_1(x_1) + \dots + v_n(x_n) \quad (17)$$

Les fonctions d'agrégation d'ensembles flous qui correspondent à l'existence d'une fonction d'utilité additive sont donc les conjonctions et les disjonctions

strictes, les moyennes bisymétriques et les sommes symétriques associatives strictement croissantes. Par exemple :

Si $\exists f : [0, 1] \rightarrow [0, +\infty)$ $f(0) = 0$, $f(1) = +\infty$ telle que

$$\mu_G(x) = f^{-1}(f(\mu_{G_1}(x_1)) + \dots + f(\mu_{G_n}(x_n)))$$

alors on voit clairement que $f \circ \mu_G$ est une fonction d'utilité additive. Par conséquent l'approche par les ensembles flous indique comment, sur l'échelle de valeur $[0, 1]$, interpréter les utilités additives en termes d'union, d'intersection, ou de moyennes d'ensembles flous. On obtient donc un classement naturel des utilités additives, qui peut en faciliter la recherche en pratique.

Mais il est clair que l'hypothèse de compensation de la théorie de l'utilité n'est pas toujours vérifiée par une fonction d'agrégation issue de la théorie des ensembles flous : c'est le cas avec des conjonctions ou disjonctions nilpotentes, avec certaines sommes symétriques rencontrées, avec la médiane, et les opérateurs min et max.

Il est important de noter que l'on rencontre depuis longtemps les opérations d'agrégation min et max dans la littérature, puisque ce sont celles qu'ont utilisé Von Neuman et Morgenstern pour fonder la théorie des jeux /43/. Ces opérations sont couramment utilisées par l'école "objectiviste" (critères du maximin, critère du maximax).

En fait on peut, suite à cette analyse succincte, constater que l'approche de l'agrégation de critères par les ensembles flous propose un cadre commun aux écoles objectiviste et subjectiviste (utilité), et développe un système d'axiomes permettant de choisir entre diverses façons de construire une fonction de valeur.

NB : Il est intéressant de noter qu'un exemple de pondération généralisée, telle qu'elle est définie au § 1.3.4c), est apparu indépendamment dans la littérature sur la théorie de l'utilité, sous le nom de "forme multiplicative de la fonction d'utilité". L'obtention de cette forme suppose l'utilisation de distributions de probabilité pour modéliser la décision en univers incertain. La fonction d'utilité multiplicative est telle que (cf. Keeney, /26/)

$$\exists \lambda > -1 \quad 1 + \lambda v(\vec{x}) = \prod_{i=1}^n (1 + \lambda g_i \cdot v_i(x_i))$$

où les g_i sont des constantes telles que

$$1 + \lambda = \prod_{i=1}^n (1 + \lambda g_i)$$

En posant

$$\hat{\varphi}(x) = \frac{\text{Log}(1 + \lambda x)}{\text{Log}(1 + \lambda)}, \quad v(x) = \mu_D(x), \quad v_i(x) = \mu_{G_i}(x_i)$$

on trouve

$$\mu_D(\vec{x}) = \hat{\varphi}^{-1} \left(\sum_{i=1}^n \hat{\varphi}(g_i \mu_{G_i}(x_i)) \right)$$

c'est-à-dire que la forme multiplicative de l'utilité peut s'exprimer comme la mesure d'incertitude d'un événement flou (l'ensemble flou des objectifs atteints par \vec{x}) au sens d'une mesure floue g_λ de Sugeno /41/. Au lieu de (16), on retrouve ici l'autre définition possible de $g_\lambda(F)$ quand F est flou (cf. chap.2, NB du § 4.3.4). Ces découvertes du même concept, effectuées indépendamment, tendent à conforter encore l'approche de la pondération des critères fondée sur les mesures généralisées d'incertitude.

6.1.4 - EXEMPLES D'APPLICATION

Dans cette section on passe en revue des études qui ont donné lieu à l'utilisation d'une approche multicritère basée sur les ensembles flous.

Dubois /8/, /11/, a suggéré un outil de communication homme-machine pour la représentation de spécifications vagues d'un objet recherché dans une base de données technologiques. Le problème se pose lorsqu'un concepteur de système mécanique recherche une pièce standard pour remplir une certaine fonction, et qu'il faut classer des objets de formes analogues mais de dimensions variées, par ordre de compatibilité avec un objet idéal souhaité. Une solution possible consiste à sélectionner les objets standards dont la distance à l'objet idéal est minimale. Néanmoins si la distance est un opérateur standard, on a aucun contrôle sur le processus de sélection. Partant de la constatation que les tolérances existant sur l'objet idéal en font un objet imprécis, on peut définir la spécification à l'aide d'intervalles admissibles pour chacune des dimensions. Le risque est alors de ne rien

sélectionner (tolérances trop strictes) ou de sélectionner trop d'objets (tolérances trop larges). Un compromis consiste à spécifier les tolérances par des nombres flous qui seront autant d'objectifs partiels, qu'il faudra agréger pour définir le degré global de compatibilité des objets du catalogue par rapport à la spécification. L'identification des opérations d'agrégation se fait par quelques questions posées au demandeur, pour discriminer parmi un ensemble restreint d'opérations. De cette façon on construit en temps réel une distance entre la spécification et les objets standards, qui correspond mieux aux souhaits du demandeur. Cette étude a été menée avec l'Agence pour le Développement de la Production Automatisée, en vue d'améliorer les performances d'un système d'interrogation de base de données de pièces mécaniques.

L'application évoquée ci-dessus entre dans le cadre plus large de la conception assistée par ordinateur. L'intérêt de la théorie des ensembles flous pour la C.A.O. a été soulignée dans Dubois /9,10/. Elle offre en effet des méthodes pour représenter l'imprécision qui envahit les premiers stades d'un processus de conception (qu'on peut voir comme une suite de décisions à prendre). L'utilisation possible du modèle d'agrégation de critères, décrit dans cette section, pour l'analyse des spécifications fonctionnelles imprécises d'un système à concevoir, en vue de définir à partir d'elles les objets qui le composeront, a été suggérée /10/. Chaque fonction que doit remplir le système est décrite par un ensemble d'attributs. Ces fonctions sont spécifiées au travers des valeurs que peuvent prendre ces attributs ou propriétés. Le problème est alors de regrouper les fonctions selon leur compatibilité, de façon à définir des objets qui pourront remplir simultanément plusieurs fonctions. La compatibilité de deux (ou plusieurs) fonctions se définit comme l'agrégation de leurs compatibilités partielles, au sens de chaque propriété. La compatibilité partielle est calculée au sens d'un indice de comparaison entre les deux ensembles flous caractérisant chacune des propriétés (cf. chap.2 § 4.2).

Une application à un problème plus classique de décision a été menée à bien par l'équipe du Prof. Zimmermann /51,52/. Il s'agissait d'évaluer les demandes de prêts reçues par une compagnie bancaire, pour décider s'il y avait lieu ou non de les accepter. L'objectif global qui est exprimé par le concept "digne de crédit" a été décomposé de façon hiérarchique en objectifs élémentaires définis sur des échelles physiques. La famille d'opérateurs

choisie permet d'explorer une zone de comportement entre la conjonction stricte jusqu'à la disjonction stricte :

$$h(u_1, \dots, u_n) = \left(\prod_{i=1}^m u_i \right)^\gamma \left(1 - \prod_{i=1}^n (1-u_i) \right)^{1-\gamma}, \quad \gamma \in [0,1]$$

γ représente un degré de compensation entre les objectifs agrégés.

Une analyse empirique a permis de déterminer pour chaque niveau de la hiérarchie de concepts des valeurs plausibles de γ permettant des agrégations valides dans le contexte étudié. Nous envisageons une application analogue à un problème d'évaluation de projets de recherche dans une firme de produits pharmaceutiques (voir Boschi /4/ pour un exposé de certains aspects de ce problème).

Enfin, est actuellement à l'étude au CERT, Toulouse, un problème d'ergonomie de pilotage, en vue du développement d'un outil logiciel d'aide à la certification d'avions. Un des aspects cruciaux concerne l'évaluation de la charge mentale d'un pilote, qui suppose un modèle capable de prédire sa capacité à détecter des situations anormales et de réagir en face de ces situations. Les décisions prises par le pilote sont le résultat d'une confrontation entre une situation souhaitée (décrite par des ensembles flous relatifs aux divers paramètres mesurés en vol) et une situation perçue (qui dépend de la visibilité des instruments de mesure, donc de l'arrangement du cockpit ; cet aspect est aussi représenté par des ensembles flous). L'évaluation de la gravité de la situation telle qu'elle est ressentie par le pilote est obtenue en partie à l'aide d'une agrégation d'ensembles flous. Les opérations correspondantes peuvent être obtenues à l'aide de l'approche décrite aux paragraphes précédents. Pour une présentation de ce problème voir Boy /6/.

6.2 - AIDE À LA DÉCISION EN ENVIRONNEMENT INCERTAIN OU IMPRÉCISÉMENT PERÇU.

Le modèle classique pour la prise de décision en environnement incertain a son origine dans les travaux de Von-Neumann et Morgenstern /43/, et a été développé dans les années 40-50 par des chercheurs tels que Savage /38/, puis étendu au cas multicritères par Fishburn /21/, Keeney et Raiffa /27/, etc... Tous ces modèles reposent sur une hypothèse fondamentale, à savoir que la mauvaise connaissance sur les conséquences d'une action potentielle a $\in A$ peut

s'exprimer par une distribution de probabilité sur l'ensemble des conséquences X . La justification de cette approche est fondée sur la notion de pari que l'on peut résumer par les deux conventions suivantes :

- Une personne qui affirme "x est la conséquence de a" accepte de payer 1 franc si son affirmation s'avère être fausse.
- Une personne qui affirme "x est la conséquence de a avec la probabilité p" accepte d'affirmer "x est la conséquence de a" si on lui donne $1-p$ francs, et le contraire si on lui donne p francs.

Si on admet comme raisonnable ces hypothèses, alors le problème de choisir la meilleure action $a \in A$ en environnement incertain se ramène à trouver une relation d'ordre complet dans l'ensemble des distributions de probabilité définies sur X , soit $\mathcal{P}(X)$. Dans ce cas, la fonction d'utilité est une application de $\mathcal{P}(X)$ dans \mathbb{R} qui respecte cet ordre complet. Von-Neuman et Morgenstern /43/ ont, les premiers, définis une série d'axiomes permettant d'affirmer, à une transformation affine près, l'unicité d'une fonction d'utilité.

Le problème majeur rencontré par cette approche, au demeurant rigoureuse, réside dans l'obtention, en pratique de valeurs de probabilités relatives à l'occurrence de conséquences, au cas où telle action serait choisie. On peut séparer la difficulté en deux aspects, qui néanmoins interfèrent.

- . L'appréciation des conséquences par un individu n'est pas forcément bien représentée par une mesure de probabilité, stricto sensu, mais par une mesure d'incertitude monotone à déterminer, qui dépendra de son attitude face au risque.
- . Quand on affirme "x est la conséquence de a avec 0,7 de probabilité", il est clair que la valeur .7 est fournie à titre indicatif, et qu'en réalité on veut dire ".7 avec une certaine marge d'arbitraire."

L'exploration de classes de fonctions d'ensemble généralisant la notion de probabilité, faite au chapitre 1, prend tout son sens dans le problème de décision en environnement incertain puisque on possède ainsi de nouveaux modèles pour représenter une action potentielle incertaine. De plus l'arithmétique des nombres flous peut faciliter la représentation de "proba-

bilités linguistiques" en vue d'une analyse de la sensibilité des choix aux valeurs numériques utilisées dans le modèle. Il est clair que la valeur x donnée comme conséquence de a est aussi une approximation, et qu'une plage de conséquences représente mieux la connaissance que possède un individu.

Dans ce domaine, notre contribution a, dans le passé, concerné surtout le second aspect, à savoir la modélisation et l'utilisation des "probabilités linguistiques" (Dubois et Prade /13,17/). Un modèle possibiliste de décision a été néanmoins esquissé dans /18,19/, et sera évoqué ici.

Mais, avant d'étudier les deux aspects d'une généralisation des modèles classiques de décision en environnement incertain, on va revenir sur le concept d'utilité au sens de Von Neumann-Morgenstern et indiquer, qu'à l'intérieur même du cadre probabiliste, il est possible, en modifiant l'un des axiomes, de prendre en compte l'attitude d'un décideur. Cette section se terminera par un retour sur le multicritère.

6.2.1 - UNE EXTENSION DU CONCEPT D'UTILITE DE VON NEUMANN MORGENSTERN

Soit un ensemble \mathcal{X} de conséquences incertaines telles que $\forall p \in [0,1]$

$$x, y \in \mathcal{X} \rightarrow (x, p, y) \in \mathcal{X}$$

où (x, p, y) signifie que la probabilité d'occurrence de x est p et celle de y est $1-p$. On définit sur \mathcal{X} un préordre \geq complet, $>$ designant la préférence stricte et \sim l'indifférence. On a par définition

$$(x, 1, y) \sim x \quad ; \quad (x, p, y) \sim (y, 1-p, x)$$

On se donne alors les axiomes suivants :

- U1) $x \sim y \rightarrow \forall p \forall z, (x, p, z) \sim (y, p, z)$
- U2) $x > y \rightarrow \forall p \in]0, 1[, x > (x, p, y) > y$
- U3) $x > y > z \rightarrow \exists p \in]0, 1[, y \sim (x, p, z)$
- U4) $((x, p, y), q, y) \sim (x, pq, y)$.

Von Neumann et Morgenstern /43/ ont prouvé l'existence d'une fonction d'utilité $u : \mathcal{X} \rightarrow \mathbb{R}$ telle que

$$x > y \leftrightarrow u(x) > u(y) \quad (18)$$

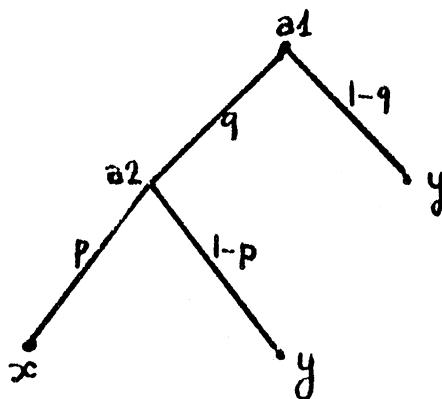
$$x \sim y \leftrightarrow u(x) = u(y) \quad (19)$$

$$u(x, p, y) = pu(x) + (1-p)u(y) \quad (20)$$

et si u, u' sont deux utilités alors $\exists a, b \in \mathbb{R}, u = a+bu'$

Parmi les axiomes U1-4, celui que l'on peut le plus critiquer est U4.

En voyant $((x, p, y), q, y)$ comme un arbre de décision



On voit que la conséquence x est obtenue si en choisissant l'acte a_1 , le décideur est encore en mesure de choisir a_2 (ce qui arrive avec la probabilité q) qui fournit la conséquence espérée (avec probabilité p). U4 suppose clairement que l'événement qui suit a_1 est indépendant de celui qui suit a_2 ; c'est le décideur qui fait cette supposition.

Si on veut assouplir cette hypothèse, on peut simplement supposer que la probabilité d'obtenir x est une certaine combinaison de p et q , soit $p \perp q$ (Fine /20/). Il est naturel de supposer \perp associatif, non décroissant en fonction de ses arguments, commutatif et tel que $1 \perp p = p$; donc \perp est naturellement une norme triangulaire. Soit (A, B, C) trois événements mutuellement "indépendants" au sens de \perp , sur un espace de probabilité (Ω, \mathcal{A}, P) . Imposer $P(A \wedge B \wedge C) = P(A) \perp P(B) \perp P(C)$ et préserver l'associativité de l'union ensembliste conduit à supposer que

$$P(A \cup B) = P(A) + P(B) - P(A) \perp P(B)$$

est associatif. Donc \perp est une t -norme de la famille de Frank (cf. annexe),

c'est-à-dire telle que

$$p \perp_s q = f_s^{-1}(f_s(p) + f_s(q)) ; f_s(p) = \text{Log}_s \left(\frac{s-1}{s^p-1} \right), s > 0 \quad (21)$$

pour $s = 0 : p \perp_s q = \max(0, p+q-1)$

$s = 1 : p \perp_s q = p \cdot q$.

$s = +\infty : p \perp_s q = \min(p, q)$

Le paramètre s permet de synthétiser l'attitude du décideur :

- pessimiste $s = 0$: si x est obtenu par une chaîne d'événements incertains, ces événements ne se produiront ensemble qu'un minimum de fois. En particulier $\forall \{p_n\}_n, p_n < 1, \exists n_0, p_1 \perp_0 \dots \perp_0 p_{n_0} = 0$
- optimiste $s = +\infty$: si le décideur obtient ce qu'il souhaite pour la décision la plus incertaine, alors les autres décisions auront le résultat escompté.
- intermédiaire $s = 1$: les événements qu'entraînent les décisions arrivent indépendamment, au hasard.

La question naturelle est alors : si on substitue une t -norme de Frank au produit dans U_4 , peut-on alors construire une fonction d'utilité ?

Dans la suite on propose quelques éléments de réponse, qui constituent un début d'étude de ce problème. Prenons donc comme axiomes $U1-U3$ et

$$U_4' : ((x, p, y), q, y) \sim (x, p \perp q, y)$$

où \perp est une t -norme de Frank.

Lemme 6.1 : $\forall p, q \in]0, 1[, \forall x, y \in X$

$$x > y, p > q \Rightarrow (x, p, y) > (x, q, y)$$

Preuve

(U2) : $x > y \rightarrow (x, p, y) > y \rightarrow (x, p, y) > ((x, p, y), r, y), \forall r \in]0, 1[$

(U4') : $(x, p, y) > (x, p \perp r, y) \forall r \in]0, 1[$

$\exists r : p \perp r = q$, et c'est (cf. chap.2, § 3.5.2) $r = p \perp q$

qui vaut ici : $r = f_s^{-1}(f_s(q) - f_s(p))$.

si $s = 0$: $r = 1 - p + q$; si $s = 1$: $r = q/p$; si $s = +\infty$: $r = q$

Q.E.D.

Lemme 6.2 : $\forall p, q \in]0, 1[, \forall x \neq y \in \underline{X}$,

$$x > y \text{ et } (x, p, y) \sim (x, q, y) \rightarrow p = q$$

Preuve immédiat par le lemme 1.

Lemme 6.3 : Si $p > r$ alors $\exists \alpha$:

$$((x, p, y), q, (x, r, y)) \sim (x, \alpha, y)$$

Preuve (lemme 6.1) : $x > (x, p, y) > (x, r, y) > y$

$$\rightarrow \exists t, (x, p, y) \sim (x, t, (x, r, y)) \sim (x, t \star r, y)$$

où \star est la co-norme duale de \perp ($s \star r = 1 - (1-s) \perp (1-r) = s + r - s \perp r$) ; ceci se déduit facilement de U_4' .

La solution de l'équation $p = t \star r$ est $t = r \leftarrow p$ (cf. chap.2 § 3.5.2)

qui vaut $\varphi_s^{-1}(\varphi_s(p) - \varphi_s(r))$ où φ_s est le générateur additif de \star ($=f_s(1-.)$)

si $s = 0$: $t = p - r$; si $s = 1$: $t = \frac{p-r}{1-r}$; si $s = +\infty$: $t = p$.

D'où

$$\begin{aligned} ((x, p, y), q, (x, r, y)) &\sim ((x, s, (x, r, y)), q, (x, r, y)) \\ &\sim (x, s \perp q, (x, r, y)) && (U_4') \\ &\sim (x, (s \perp q) \star r, y) && (U_4') \end{aligned}$$

et on a donc

$$\alpha = ((r \leftarrow p) \perp q) \star r \tag{22}$$

Q.E.D.

Si $p < r$, on utilise la relation $(x, p, y) \sim (y, 1-p, x)$ pour déduire

$$\alpha = ((p \leftarrow r) \perp (1-q)) \star p \tag{23}$$

On notera que pour $s = 1$ (le cas étudié classiquement) on a toujours

$$\alpha = pq + (1-q)r \tag{24}$$

c'est-à-dire que le lemme 3 suggère la propriété fondamentale d'une fonction d'utilité $u(x, q, y) = u(x)q + (1-q)u(y)$.

Pour $s = 0$ on trouve $\alpha = \max(r, q+p-1)$ si $p > r$, et sans cette hypothèse

$$\alpha = \min(\max(r, q+p-1), \max(p, r-q)) \quad (25)$$

pour $s = +\infty$ on trouve $\alpha = \text{med}(p, q, r)$ si $p > r$. Il est remarquable de retrouver la médiane quand l'opération de composition dans U_4' est le minimum, alors qu'on trouve la moyenne pondérée quand cette opération est le produit. Sans hypothèses on a

$$\alpha = \min(\text{med}(p, q, r), \text{med}(p, 1-q, r)) \quad (26)$$

notons qu'on a bien $((r \leq p) \perp 0) * r = r$;

$$((r \leq p) \perp 1) * r = (r \leq p) * r = \varphi_s^{-1}(\varphi_s(r) + \varphi_s^0 \varphi_s^{-1}(\varphi_s(p) - \varphi_s(r))) = p$$

de sorte que α varie de r à p quand q passe de 0 à 1.

Avant de formuler le résultat principal, supposons que X comporte un plus grand élément x_1 et un plus petit élément $x_0 < x_1$. Définissons une fonction $u = X \rightarrow [a_0, a_1] = I$ par

$$u(x_1) = a_1 ; \quad u(x_0) = a_0 < a_1 \quad (27)$$

$$\forall x, u(x) = a_1 q + a_0(1-q) \quad \text{où } x \sim (x_0, q, x_1) \quad (28)$$

On a alors le résultat suivant :

Proposition 6.1 : La fonction u vérifie les propriétés suivantes

$$\forall x, y, x > y \rightarrow u(x) > u(y) \quad (29)$$

$$x > y \rightarrow \forall q \in [0, 1] \quad u(x, q, y) = (((u(y) \leftarrow_I u(x)) \perp_I (a_1 q + a_0(1-q))) *_{I_1} u(x)) \quad (30)$$

où $\leftarrow_I, \perp_I, *_{I_1}$ sont définis par

$$\frac{a_0 b - a_0}{a_1 - a_0} = \frac{a - a_0}{a_1 - a_0} \quad \circ \quad \frac{b - a_0}{a_1 - a_0} \quad \text{avec } \circ = \leftarrow, *, \perp \quad (31)$$

Preuve (29) est évident, en posant $x = (x_1, p, x)$ $y = (x_1, r, y)$

on a alors d'après le lemme 3 :

$$\begin{aligned}
u(x,q,y) &= u(x,(r \leftarrow p) \downarrow q) * r, y \\
&= (a_1 - a_0) [(r \leftarrow p) \downarrow q] * r + a_0 \\
&= (a_1 - a_0) \left[\left(\frac{u(x) - a_0}{a_1 - a_0} \leftarrow \frac{u(y) - a_0}{a_1 - a_0} \right) \downarrow q * \frac{u(y) - a_0}{a_1 - a_0} \right] + a_0
\end{aligned}$$

Comme (31) s'écrit aussi $(a_1 - a_0) a_0 b + a_0 = [(a_1 - a_0)a + a_0] \circ_I [(a_1 - a_0)b + a_0]$ le résultat (30) est évident.

Proposition 6.2 : si u et u' vérifient (27) et (28) pour (a_0, a_1) et (a'_0, a'_1) respectivement, alors $\exists \alpha, \beta u' = \alpha u + \beta$

Preuve u et u' sont en bijection croissante i.e. $\exists f : [a'_0, a'_1] \rightarrow [a_0, a_1]$ telle que

$$f(a'_0) = a_0 ; f(a'_1) = a_1 \text{ et } f(u(x,q,y)) = u'(x,q,y).$$

En particulier

$$f(x_1, q, x_0) = f((a_1 - a_0)q + a_0) = [f(a_1) - f(a_0)] \cdot q + f(a_0)$$

donc f est une fonction affine de q et

$$\alpha = \frac{a'_1 - a'_0}{a_1 - a_0} , \quad \beta = \frac{a_1 a'_0 - a_0 a'_1}{a_1 - a_0}$$

on vérifie facilement alors, que

$$\forall x > y, \alpha(u(x,p,y)) + \beta = u'(x,p,y)$$

Q.E.D.

On voit donc que l'on peut définir un concept d'utilité qui ne présuppose pas une règle de composition des probabilités fondée sur l'hypothèse d'indépendance entre événements résultants de décisions successives. Si l'utilité d'une conséquence x est toujours le barycentre des conséquences extrêmes x_1 et x_0 avec des poids p et $1-p$, où (x_1, p, x_0) a pour équivalent certain x , cela n'est en général plus vrai pour n'importe quel x', x'' tels que $x \sim (x', q, x'')$. On notera que lorsque le paramètre s de la t -norme \downarrow prend des valeurs limites ($s = 0, s = 1$) alors le lemme 6.3 contredit l'axiome U2:

$$\begin{aligned}
\underline{s = 0} : & \quad x = (x_1, p, x_0) > y = (x_1, r, x_0) \text{ et } q < \min(p, r) = r \\
& \quad \text{alors } (x, q, y) \sim (x_1, r, x_0) \sim y
\end{aligned}$$

$$\begin{aligned}
\underline{s = 1} : & \quad \text{dès que } r > q + p - 1 \text{ c'est-à-dire } q < 1 + r - p \\
& \quad \text{alors } (x, q, y) \sim (x_1, r, x_0) \sim y
\end{aligned}$$

Autrement dit les propositions 6.1 et 6.2 ne sont plus valides. La raison en est que les t-normes de Frank ne sont strictes que pour $s \in]0, +\infty[$; pour $s = 0$ on obtient une t-norme nilpotente, et $s = +\infty$, le minimum. Ces deux t-normes ne sont pas compatibles avec le concept d'ordre strict c'est-à-dire

$$a > b \not\equiv \forall c, \min(a, c) > \min(a, b)$$

$$a > b \not\equiv \forall c, \max(0, a+c-1) > \max(0, b+c-1)$$

En particulier (29) est peu crédible a priori avec $\perp = \min$ ou $\max(0, a+b-1)$, et (30) s'avère faux.

On peut expliciter plus avant le sens de (30) pour $0 < s < +\infty$. Notons (30) sous la forme d'une fonctionnelle $\Phi(u, v, w)$ avec

$$u(x, q, y) = \Phi(u(x), q, u(y))$$

alors si (z, r, y) est une loterie telle que $z \sim (x, q, y)$, on a

$$\Phi \left[\Phi(u(x), p, u(y)), r, u(y) \right] = \Phi(u(x), p \perp q, u(y))$$

La brève investigation qui précède indique qu'il est sans doute possible de sortir des hypothèses probabilistes traditionnelles pour définir une fonction d'utilité. Parmi les prolongements naturels de la tentative esquissée ici, mentionnons

- a) l'extension des propositions 6.1 et 6.2 au cas où X n'a pas de plus grand ni de plus petit élément.
- b) l'étude de nouveaux axiomes remplaçant U1-U4 quand on admet qu'une conséquence incertaine est décrite par une mesure d'incertitude monotone telle que probabilité supérieure et/ou inférieure, fonction d'ensemble décomposable générale, etc... telles qu'on les évoque au chapitre 1. Néanmoins cela nécessite au préalable une analyse de ce que devient le concept d'indépendance dans ce nouveau cadre.

La section qui suit tente d'aborder des problèmes voisins de b), en admettant le modèle probabiliste classique comme valide, mais en supposant que les valeurs de probabilité d'occurrence des conséquences sont mal connues.

6.2.2 - ENVIRONNEMENT ALEATOIRE IMPRECISEMENT PERCU

Dans cette partie on considère un problème d'évaluation d'actions potentielles dont les conséquences ne sont connues qu'imparfaitement au sens où on ne sait les décrire que macroscopiquement, et que leur probabilité d'occurrence est subjectivement et donc imprécisément estimée.

Une première façon simple de représenter cette situation est de partager l'ensemble X des conséquences en m sous-ensembles $X^1 \dots X^m$, et l'estimée des probabilités $P(X^i)$ est un intervalle $[a_i, b_i] \subseteq [0, 1]$.

Une façon plus élaborée consiste à admettre que les X^i sont des ensembles flous sur X et les $P(X^i)$ des nombres flous à support dans $[0, 1]$ qui sont des représentations quantitatives nuancées de connaissances imprécises exprimées linguistiquement, telles que

"Si on effectue l'action a , il est très probable que nos ventes augmentent beaucoup".

Dans ce cadre de modélisation on peut vouloir classer les actions au vu d'un critère qui sera l'espérance mathématique de l'utilité des actions, ou la probabilité pour que l'utilité des actions dépasse un certain seuil (éventuellement flou !).

Dans le cas d'un modèle par intervalles on retrouve un problème d'analyse de sensibilité évoqué déjà dans la littérature (Roy /32/, Sage /37/ par exemple) mais apparemment non analytiquement résolu dans le cas général. Le modèle fondé sur l'utilisation de nombres flous a été proposé par Zadeh⁽¹⁾, étudié par Watson et al /44/ dans le cas $m = 2$ (2 conséquences possibles) et Dubois et Prade ($m = 3$ dans /13/ puis généralisé à m quelconque dans /17/) Le modèle utilisant des nombres flous coïncide avec l'analyse de sensibilité quand ces nombres flous sont des intervalles. Dans la suite on passe en revue les résultats obtenus dans /17/, basés sur ceux du chapitre 4 dans le présent mémoire.

6.2.2.1 - Conséquences imprécises

On suppose que X est un intervalle réel où l'attribut significatif de l'action prend sa valeur, et $X^1 \dots X^m$ forment des intervalles (non

(1) in /36/ chap.3.

nécessairement adjacents) $[\underline{x}_i, \bar{x}_i]$. On connaît $p_i = P(X^i)$. Dans ce cas l'espérance mathématique de l'évaluation que l'on peut faire sur l'action a correspondante est simplement un intervalle

$$\left[\sum_{i=1}^m \underline{x}_i p_i, \sum_{i=1}^m \bar{x}_i p_i \right] = \bar{x}(a) \quad (32)$$

Si les X^i sont des intervalles flous, alors

$$\bar{x}(a) = p_1 X^1 \oplus \dots \oplus p_m X^m \quad (33)$$

qui est simplement obtenu avec les techniques de l'arithmétique floue, laquelle revient à effectuer le calcul sur les intervalles

$$X_\alpha^i = \{x / \mu_{X^i}(x) > \alpha\}, \text{ pour tous les } \alpha \in]0, 1].$$

Il est facile de voir que $\{(X^i, p_i) / i = 1, m\}$ définit des probabilités supérieures et inférieures, et que l'intervalle $\bar{x}(a)$ est borné par les espérances mathématiques supérieures et inférieures de la fonction identité sur X . Si on a défini une fonction d'utilité sur X , telle que $u(x)$ évalue l'utilité de la conséquence x , l'utilité espérée de a , soit $U(a)$ est donc définie par (cf. chap.1).

$$U(a) = [E_\star(u), E^\star(u)] \quad (34)$$

avec

$$E_\star(u) = \int_{-\infty}^{+\infty} u \, dF_\star(u)$$

$$E^\star(u) = \int_{-\infty}^{+\infty} u \, dF^\star(u)$$

et

$$F_\star(u) = P^\star \{x/u(x) \leq u\} \quad F^\star(u) = P_\star \{x/u(x) \leq u\}$$

Entre dans ce schéma de calcul le cas où l'évaluation souhaitée est la probabilité pour que la conséquence de a soit dans un sous-ensemble S de X (notée $P_a(S)$) alors l'évaluation vaut

$$P_a(S) \in [E_\star(\mu_S), E^\star(\mu_S)] = [P_\star(S), P^\star(S)]$$

$$\text{où } P_\star(S) = \sum_{\substack{X^i \subseteq S \\ X^i \neq \emptyset}} p_i \quad P^\star(S) = \sum_{\substack{X^i \not\subseteq S \\ X^i \neq \emptyset}} p_i$$

Si l'objectif recherché est décrit par un événement flou F (par exemple : "des ventes importantes" définissant un seuil flou) on utilise (34) avec $u = \mu_F$. Le calcul pratique de l'évaluation ne pose pas de problème particulier, si la fonction u est simple à inverser.

Dans le cas de conséquences typiques X^i décrites par des ensembles flous, $\{(X^i, p_i), i=1, m\}$ définit une variable aléatoire floue, et la probabilité de l'événement flou F décrivant l'objectif espéré vaut alors, en suivant Yager (cf. chap.5)

$$P_a(F) = p_1 \tau(F/X^1) \otimes \dots \otimes p_m \tau(F/X^m) \quad (35)$$

où $\tau(F/X_i)$ est la compatibilité de X^i par rapport à F . (35) redonne (34) si les X^i sont des intervalles usuels. Le calcul de (35) est simple, une fois obtenus les $\tau(F/X^i)$. Ces derniers demandent de savoir inverser μ_F pour être analytiquement calculés. Mais en général, une discrétisation du référentiel X et un simple lissage suffiront.

6.2.2.2 - Probabilités imprécises des conséquences

On suppose que la décision a possède m conséquences possibles $\{x_i \in X, i=1, m\}$ dont les probabilités p_i sont bornées par des intervalles $[a_i, b_i]$. Comme le mentionnent Steuer /40/, ou Sage /37/ la détermination de la conséquence espérée $\bar{x}(a) = \sum x_i p_i$ se ramène à un problème de programmation linéaire

$$\min_{p_1 \dots p_m} \sum_i x_i p_i$$

sous les contraintes

$$a_i \leq p_i \leq b_i \quad i = 1, m$$

$$\sum_{i=1}^m p_i = 1$$

qui donne la borne inférieure de $\bar{x}(a)$. Néanmoins le calcul de la borne supérieure est tout aussi significatif.

Dubois et Prade /15/ (voir aussi le chapitre 4) ont montré que si les x_i sont ordonnés de façon à ce que $x_1 \leq \dots \leq x_n$, alors les bornes de

$\bar{x}(a)$ sont chacune atteinte parmi un ensemble de n solutions $\vec{p}^1 \dots \vec{p}^n$ dont seule une est réalisable. Dans le cas de la borne inférieure, \vec{p}^k est défini par

$$\begin{aligned} p_i^k &= b_i & i < k \\ p_i^k &= a_i & i > k \\ p_k^k &= 1 - \sum_{i < k} b_i + \sum_{i > k} a_i \end{aligned}$$

et

$$\inf \bar{x}(a) = \max_{k=1, n} \left(\sum_{i=1}^m x_i p_i^k \right) \quad (36)$$

(voir chap. 4 pour plus de détails).

Il paraît plus naturel de représenter l'imprécision sur les probabilités subjectives p_i par des nombres flous, pour souligner le caractère relativement arbitraire des bornes a_i et b_i . Ces nombres flous sont une représentation plus satisfaisante de catégories linguistiques telles que "très probable", "douteux" etc... utilisées couramment par les décideurs. Le concept de nombre flou paraît plus naturel qu'une distribution de probabilité sur la valeur même de p_i , comme suggère Sage /37/, car on se heurte alors à des problèmes d'interprétation et d'estimation très délicats. Des valeurs floues de probabilité peuvent aussi être construites à partir d'informations floues par l'intermédiaire de calculs de cardinaux d'ensembles flous (cf. Zadeh /49/ qui relie par exemple similarité floue et probabilité).

Si les \tilde{p}_i sont des intervalles flous quelconques, il convient que la distribution de probabilité imprécise ainsi obtenue soit normalisée ; ceci veut dire

$$\mu_{\tilde{p}_1 \oplus \dots \oplus \tilde{p}_n}^{\sim} (1) = 1 \quad (37)$$

(37) signifie, lorsque $\tilde{p}_i = [a_i, b_i] : \exists p_i \in [a_i, b_i], i = 1, m$ tel que $\sum p_i = 1$. pour calculer la conséquence espérée $\bar{x}(a)$, on peut appliquer (36) (et son dual pour la borne supérieure) aux α -coupes des \tilde{p}_i . Si (37) est faux, alors

$$\exists \bar{\alpha}, \forall \alpha > \bar{\alpha}, \bar{x}(a)_{\alpha} = \emptyset$$

c'est-à-dire que $\forall \alpha > \bar{\alpha}$ le programme linéaire formulé ci-dessus n'a pas de solution. Si on peut représenter paramétriquement les \tilde{p}_i , par des nombres

floous de type LL (fonctions d'appartenance de type symétrique) alors on a montré (Dubois et Prade /15/, voir aussi chap.4) que $\bar{x}(a)$ est aussi un nombre flou de type LL, et qu'il suffit de calculer (36) et son dual sur les supports de \tilde{p}_i . Dans ce cas, on est bel et bien ramené à du calcul d'intervalles, mais on véhicule une information plus riche.

6.2.2.3 - Cas mixte

L'action a possède m conséquences possibles décrites macroscopiquement par des ensembles flous, X^i que l'on suppose être des intervalles flous de \mathbb{R} . La probabilité de chaque conséquence imprécise est un intervalle flou représentant une probabilité subjective linguistique \tilde{p}_i .

Le calcul de la conséquence (floue) espérée $\bar{x}(a)$ est facilement obtenu en faisant la synthèse des deux précédents cas. On ne peut travailler que sur les α -coupes car le produit de nombres flous ne laisse pas invariante les formes I.R.. On obtient la borne inférieure de $\bar{x}(a)$ par la procédure suivante (voir aussi le chapitre 4) :

- . ranger les nombres $\underline{x}^i = \inf X^i$ par ordre croissant $\underline{x}^1 \leq \dots \leq \underline{x}^m$
- . appliquer (36) avec $x_k = \underline{x}^k \quad \forall k = 1, m$.

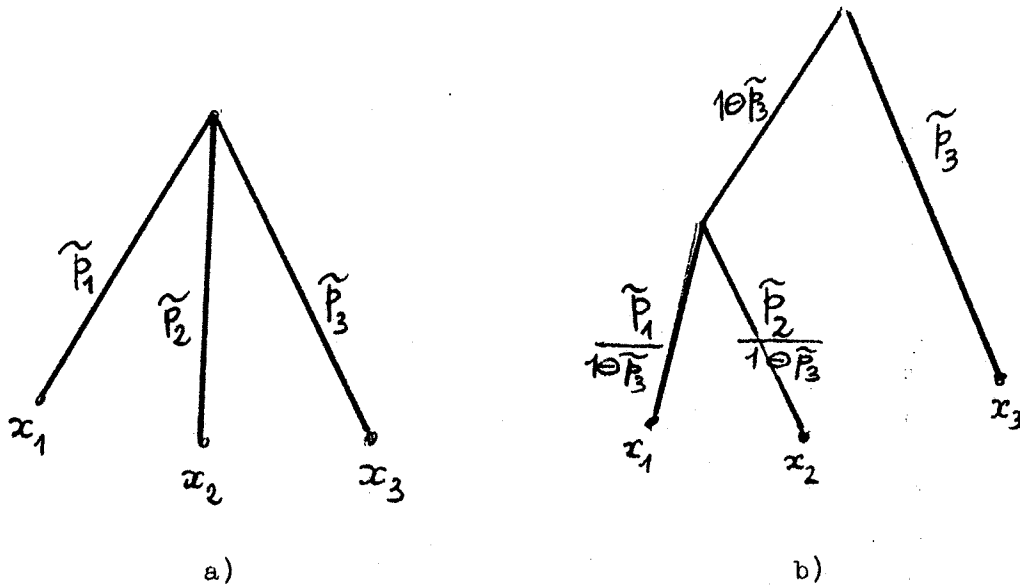
On a la procédure duale pour la borne supérieure. On obtient une bonne idée de $\bar{x}(a)$ en sélectionnant un nombre suffisant d' α -coupes.

Si l'évaluation pertinente est plutôt la probabilité $P_a(F)$ d'un événement F (flou ou non) souhaité comme conséquence de a , on peut réappliquer la même procédure que pour le calcul de $\bar{x}(a)$ ci-dessus, à la formule (35) où l'on remplace les probabilités par des nombres flous \tilde{p}_i .

6.2.2.4 - Anomalies engendrées par l'imprécision des probabilités

Quand les probabilités des conséquences sont imprécises, les techniques usuelles de décomposition d'arborescences en schémas équivalents ne s'appliquent plus. En particulier si d'un noeud part n branches, on ne peut pas développer ce noeud en une arborescence binaire. Considérons par exemple une action ayant trois conséquences possibles x_1, x_2, x_3 , avec des probabilités $\tilde{p}_1, \tilde{p}_2, \tilde{p}_3$ floues. On peut penser outre l'arborescence a) ci-après

représenter cette situation à l'aide de l'arborescence b) par exemple.



Dans la figure b) on utilise les opérateurs d'arithmétique floue définis au chap.4. En b) on est amené à calculer la conséquence espérée de x_1 avec la probabilité $\tilde{p}_1 / 1 - \tilde{p}_3$ et x_2 avec la probabilité $\tilde{p}_2 / 1 - \tilde{p}_3$, puis recombinaison le résultat avec la conséquence x_3 . Notons \bar{x}_a et \bar{x}_b les nombres flous obtenus en a) et b). Si $\forall i p_i \in (\tilde{p}_i)_\alpha$, alors

$$1 - p_3 \in (1 - \tilde{p}_3)_\alpha, \quad \frac{p_i}{1 - p_3} \in \left(\frac{\tilde{p}_i}{1 - \tilde{p}_3} \right)_\alpha \quad \forall i$$

On a supposé ci-dessus que $p_1 + p_2 + p_3 = 1$, et donc on a bien $\frac{p_1 + p_2}{1 - p_3} = 1$.

Donc si $x \in (\bar{x}_a)_\alpha$, alors $x \in (\bar{x}_b)_\alpha$, et donc \bar{x}_b est au moins aussi imprécis que \bar{x}_a . c'est-à-dire qu'en situation imprécise, la décomposition d'une arborescence est une technique encore valide (\bar{x}_b est cohérent avec \bar{x}_a) mais risque de détériorer la précision du résultat. Ce type d'anomalie, liée à l'imprécision a été mentionné par Freeling /22/ et Adamo /1/. Il met en évidence l'intérêt des méthodes simples de calcul direct de l'imprécision sur des arborescences n-aires, présentées dans ce mémoire.

NB : Borisov et Merkuryeva /5/ ont défini des loteries linguistiques avec la même signification que ci-dessus. Néanmoins, ils n'utilisent pas du tout ce concept de la même façon. Si $\{(x^i, \tilde{p}_i) | i = 1, m\}$ est une loterie floue, ils

cherchent à déterminer la probabilité \tilde{p} d'un autre événement flou F. Cette probabilité est définie par

$$\forall u \in [0,1], \mu_{\tilde{p}}(u) = \max_f \min_{i=1,n} \mu_{P_i}^{\sim}(P(X^i))$$

où $P(X^i)$ est la probabilité scalaire de l'événement flou X^i au sens de Zadeh (cf. chap. 2 § 4.3.1) telle que

$$P(X^i) = \int_{\mathbb{R}} \mu_{X^i}(r) f(r) dr \quad \forall i = 1, n$$

f est une densité de probabilité telle que ,

$$P(F) = \int_{\mathbb{R}} \mu_F(r) f(r) dr = u$$

Le calcul de \tilde{p} nécessite le recours à la programmation non linéaire, et s'avère bien plus complexe que l'utilisation de (35) avec des probabilités floues.

6.2.3 - MODELES NON PROBABILISTES D'ENVIRONNEMENT INCERTAIN

Toute opinion quant à l'occurrence d'un événement, suite à une certaine décision n'est pas forcément représentable par un poids de probabilité. Le cadre conceptuel des mesures d'incertitude monotones semble a priori plus naturel, de par le peu d'hypothèses qu'il présuppose, même s'il est trop large pour être opérationnel à ce niveau de généralité, puisqu'il suppose la connaissance d'une fonction d'ensemble g sur l'ensemble des conséquences X.

Pour conserver des outils de calcul pratique raisonnables, sans se limiter au cadre restrictif des probabilités, on peut faire appel au concept de probabilités supérieure et inférieure, où à celui de fonction d'ensemble décomposable, tous deux décrits au chapitre 1. Avant d'avancer dans cette voie, on va mentionner une tentative très récente⁽¹⁾ pour généraliser l'approche Bayésienne pragmatique de la théorie de la décision qui associe les probabilités subjectives à des paris dont l'issue peut coûter ou rapporter une somme d'argent (cf. introduction de la section 2).

Giles /24/ utilise la terminologie possible/impossible pour

(1) dont nous n'avons eu connaissance qu'au cours de la rédaction de ce mémoire.

caractériser l'occurrence d'une conséquence. Pour toute proposition P (= "la conséquence de a est x ", par exemple) il définit la possibilité de P , soit $g(P)$, comme le plus petit nombre $p \in [0,1]$ tel que le décideur acceptera d'affirmer $\neg P$ (= non P) contre un versement préalable d'une somme d'argent proportionnelle à p (avec l'acceptation implicite de devoir payer une somme proportionnelle à 1 si P s'avère être vrai).

Dans ce cadre, l'affirmation " P a une probabilité p " revient à affirmer simultanément $g(P) = p$ et $g(\neg P) = 1 - p$ où $\neg P$ signifie la négation de P .

Formellement la définition ci-dessus amène Giles /24/ à poser les axiomes suivants :

- . g est une mesure d'incertitude monotone sur X
- . $\forall A, B \subseteq X, \max(g(A), g(B)) \leq g(A \cup B) \leq g(A) + g(B)$ (38)

La dernière inégalité s'explique par le fait que, pour au moins $g(A)+g(B)$ francs le décideur acceptera d'affirmer $\neg A$ et $\neg B$ séparément, et donc de payer une somme de 2 francs si A et B s'avèrent vrais. Donc pour cette même somme $g(A)+g(B)$, il acceptera d'affirmer $\neg(A \cup B)$, qui l'amènera, même si A et B s'avèrent vrais, à ne payer qu'un franc.

On remarque que les possibilités de Zadeh et les probabilités sont les 2 cas limites de (38). Un cas particulier important de fonctions d'ensemble satisfaisant (38) est une probabilité supérieure P^* de Shafer-Dempster. On sait par ailleurs que la donnée d'une fonction P^* revient à celle d'un ensemble de poids de probabilité portant sur des sous-ensembles de X . On note donc qu'il y a une claire équivalence entre une situation incertaine modélisée par des probabilités précises de conséquences imprécises (c'est-à-dire

$$\{(p_i, X^i) / i = 1, n, \sum_{i=1}^n p_i = 1, \emptyset \neq X^i \subseteq X\}$$

et une situation incertaine modélisée par un ensemble de conséquences précises, muni d'une mesure d'incertitude de type Dempster-Shafer (c'est-à-dire une probabilité supérieure).

Entrent aussi dans le cadre défini par Giles /24/ les fonctions d'ensemble basées sur des co-normes $*$ telles que

$$\forall a, b \quad a * b \leq a + b$$

C'est le cas des fonctions g_λ de Sugeno /41/ pour $\lambda \in [-1, 0]$, des fonctions d'ensemble basée sur la co-norme de Frank (cf. équ.(21)), ou celle de Yager c'est-à-dire

$$A \cap B = \emptyset \rightarrow g(A \cup B) = \sqrt[q]{g(A)^q + g(B)^q} \quad (39)$$

pour $q \geq 1$

L'intérêt pratique de ces fonctions d'ensemble est de pouvoir être facilement identifiées et de ne nécessiter que peu d'information pour être caractérisées. En effet, grâce à l'interprétation pragmatique de Giles /24/ un décideur est capable de fournir des probabilités subjectives qui ont un sens pour lui (comme dans la théorie classique de l'utilité) mais sans qu'on lui impose une condition stricte sur les poids qu'il fournit, à savoir

$$\sum_{i=1}^m g_i = 1$$

où g_i est la probabilité subjective estimée de la conséquence i . Une fonction décomposable se caractérise par sa densité $g(\{x_i\}) = g_i$, $i = 1, m$ et l'opération de combinaison $*$ dont la découverte peut se ramener à l'estimation d'un paramètre. Une procédure d'élicitation de probabilités subjectives, dans cette optique peut donc être esquissée comme suit (elle ne concerne que des pseudo-probabilités. Cf. chap.1).

1/ fournir des nombres g_i reflétant la possibilité d'occurrence de x_i , $i = 1, m$, sans autre contrainte que $g_i \in [0, 1]$. Il faut néanmoins s'assurer que le décideur interprète bien les g_i en termes de possibilité (ou plausibilité) et non de certitude. En particulier il faut vérifier que l'inégalité $g(A) + g(\bar{A}) \geq 1$ n'est jamais violée (elle est moins forte que la condition de normalisation des probabilités). Cette inégalité se déduit de (38).

2/ Soit φ_λ le générateur additif de la famille de co-normes choisies pour représenter la fonction d'incertitude. Trouver λ tel que

$$\sum_{i=1}^m \varphi_\lambda(g_i) = \varphi_\lambda(1) \quad (40)$$

On remarquera que dès que $\exists i, j \ g_i = 1, g_j > 0$, la résolution de (40) n'a de sens que si $\varphi_\lambda(1) = +\infty$, c'est-à-dire que g n'est plus une probabilité distordue (ou pseudo-probabilité) mais une pseudo-possibilité, c'est-à-dire que la t -norme \star sera stricte. La procédure ci-dessus échouera alors pour la détermination de la fonction d'incertitude g . On peut alors adopter l'une des deux attitudes.

- arrêter la procédure d'élicitation, considérer $\{g_i, i=1, m\}$ comme un ensemble flou, et g comme une mesure de possibilité au sens de Zadeh.

- fournir l'évaluation de $g(A)$ pour l'événement A contenant tous les x_i tels que $g_i < 1$, trouver λ tel que

$$\sum_{i \in A} \varphi_\lambda(g_i) = \varphi_\lambda(g(A)) \quad (41)$$

où dans (41) φ_λ est un générateur de co-norme stricte. La solution de (41) suppose à nouveau que $g(A) > g_i \ \forall i$.

Une telle poursuite de la procédure est relativement délicate à mettre en oeuvre et risque de dérouter le décideur, qui peut considérer "qu'on lui demande trop". En revanche l'utilisation des g_i comme ensemble flou de conséquences possibles présente des avantages certains dans le traitement ultérieur des données ainsi recueillies. En particulier tout l'arsenal des méthodes de calcul fournies par l'arithmétique floue peut alors être appliqué.

Cette utilisation des fonctions d'ensemble décomposables pour évaluer l'incertitude qui entoure les conséquences d'une décision en montre tout l'intérêt. Elle évite de déclarer "irrationnel" un décideur qui fournit des probabilités subjectives non conformes à l'axiomatique de Kolmogorov ! Cette méthodologie tente d'adapter le modèle aux réponses du décideur plutôt que de forcer celui-ci à se plier aux a priori probabilistes (ce que la théorie classique de l'utilité est forcée de préconiser). On rappelle que l'emploi d'une fonction de probabilité pour modéliser l'incertitude nécessite une quantité d'information non nulle ; en particulier on se souviendra que l'ignorance totale est une attitude qui ne peut être modélisée par des probabilités (elle correspond à $g_i = 1 \ \forall i$ dans le modèle précédent, ce qui implique que g est une mesure de possibilité au sens de Zadeh).

Les fonctions d'ensemble décomposables semblent d'un emploi plus aisé que les probabilités supérieures quant à leur élicitation. En effet la donnée de $\{P^*(x_i) | x_i \in X\}$ ne suffit pas à caractériser P^* . De plus, même si on connaît $\{P^*(A) | A \subseteq X\}$, il faut vérifier qu'il existe une probabilité de base m telle que

$$P^*(A) = \sum_{B \wedge A \neq \emptyset} m(B)$$

ce qui peut s'avérer très lourd (malgré l'existence d'une formule d'inversion dite de Moebius). Les probabilités supérieures apparaissent plus naturellement quand on affecte des probabilités à des conséquences imprécises (cf. section 2.2.1).

Pour calculer la conséquence espérée $\bar{x}(a)$ ou l'utilité espérée $\bar{u}(a)$ d'une action a dont les conséquences sont décrites par une mesure d'incertitude g , on peut

- a) soit utiliser les concepts d'espérances mathématiques inférieure et supérieure définis au chapitre 1 et utilisés à la section précédente. Ceci est naturel quand g est une probabilité supérieure (ce qui est le cas pour certaines fonctions d'ensemble décomposables, par exemple les g_λ de Sugeno /41/ les possibilités de Zadeh ...), mais demeure possible pour n'importe quelle mesure monotone (cf. chap.2 § 4.3.5).
- b) soit, si g est une pseudo-probabilité, utiliser l'isomorphisme de structure entre g et une probabilité et construire une espérance mathématique distordue (voir remarque du § 3.5 au chapitre 1 et eqn (95) du chap.2) pour évaluer une action potentielle. Quand g est une pseudo-possibilité (par exemple une norme stricte la génère) une telle espérance mathématique n'existe plus.
- c) soit, si g est une fonction de possibilité de Zadeh utiliser le concept de possibilité d'un événement flou au sens de Zadeh (chap.2, § 4.3.2). Ceci suppose que l'utilité des conséquences x , $u(x)$, prend ses valeurs sur $[0,1]$. C'est la voie suivie par Yager /47/ et Whalen /45/.
- d) si g est une pseudo-possibilité, on a pu au chapitre 2 définir des extensions de g aux événements flous. Si la fonction d'utilité sur les conséquences peut être vue comme la fonction d'appartenance de l'ensemble S des conséquences souhaitables, comme au point précédent, alors on peut évaluer une action par la mesure $g(S)$ définie comme dans l'équation (100) au chapitre 2.

e) si g est une possibilité de Zadeh et les g_i définissent l'ensemble flou des conséquences de a , soit F : on peut manipuler F comme un nombre flou. Dans ce cas l'utilité de a sera floue et égale à $u(F)$ définie à l'aide du principe d'extension.

Discutons l'opportunité de ces diverses évaluations : a) et b) sous-tendent une interprétation probabiliste, qui font que les évaluations correspondantes méritent encore le nom d'espérance mathématique au sens strict, même si elle est imprécise (cas a) ou distordue (cas b). L'approche c) pour les possibilités s'interprète en termes de médiane, si on la replace dans le contexte de l'intégrale de Sugeno (chap.2 § 4.3.5). Elle est justifiée si on admet le concept de médiane comme raisonnable pour représenter un ensemble de nombres, plutôt que le concept de moyenne pondérée⁽¹⁾. L'approche c) se généralise bien sûr à n'importe quelle mesure d'incertitude (c'est l'intégrale de Sugeno) mais se justifie mal quand cette mesure a une interprétation probabiliste (cas a) et b). En revanche, elle est défendable pour une mesure de pseudo-possibilité au moins autant que l'approche d) qui ne propose qu'une extension possible de l'évaluation d'un événement. Enfin, l'approche e, valide seulement si g est une mesure de possibilité de Zadeh, a l'immense avantage de ne pas combiner arithmétiquement les valeurs g_i de possibilité mais de les transporter à l'aide du principe d'extension. On ne refait alors qu'une analyse de sensibilité (nuancée), qui calcule des évaluations floues (et non en moyenne, ou médiane) des actions potentielles.

On a donc mis en évidence dans cette section l'existence d'autres modèles d'évaluation d'actions potentielles en environnement flou, que ceux qui se fondent sur une représentation probabiliste de l'incertain. Un problème ouvert est alors celui de mener à bien une démarche similaire à celle que Von Neumann et Morgenstern ont tentée pour définir axiomatiquement une fonction d'utilité, à savoir de quelle structure ordonnée peut-on munir l'ensemble des mesures d'incertitude monotones d'une famille donnée sur X , quel type de projection de cette structure sur la droite réelle peut-on espérer réaliser ?

(1) L'idée de médiane est particulièrement utile si les évaluations d'incertitude ne sont plus numériques mais forment un treillis (abstrait). Alors le concept de moyenne n'existe plus, mais la médiane (dans sa définition à l'aide de max et min) existe encore. C'est ce qui justifie le travail de Whalen /45/.

6.2.4 - MODELE POSSIBILISTE D'ENVIRONNEMENT INCERTAIN

(Dubois et Prade /18,19/)

Supposons que l'action potentielle "a" ait plusieurs conséquences possibles $x_1 \dots x_n$ dont l'éventualité est mesurée au sens de degrés de possibilité: $\pi_1 \dots \pi_n$, dont l'interprétation est conforme à l'approche de Giles /24/. On sait que les fonctions de possibilité de Zadeh sont un cas limite qu'il est intéressant d'analyser pour faire le pendant avec l'autre cas limite que sont les probabilités.

Soit $\overset{\circ}{F}$ l'ensemble flou des conséquences possibles de a, tel que $\forall x_i, \mu_{\overset{\circ}{F}}(x_i) = \pi_i$. Alors on peut interpréter les niveaux π_i comme suit :

$\forall x_i \notin \overset{\circ}{F}$ (le noyau de $\overset{\circ}{F}$), π_i est la limite supérieure de la probabilité de x_i
 $\forall x_i \in \overset{\circ}{F}$, $\pi_i = 1$ mais le degré de nécessité $N(\overset{\circ}{F})$ vaut

$$N(\overset{\circ}{F}) = \inf_{x_i \notin \overset{\circ}{F}} 1 - \pi_i \leq P(\overset{\circ}{F})$$

c'est-à-dire que la probabilité de l'événement $\overset{\circ}{F}$ est bornée inférieurement.

Dans ce cas la conséquence espérée de a est un intervalle $[E_{\star}(x), E^{\star}(x)]$ borné par les espérances mathématiques supérieures et inférieures de π (cf. chap. 1 § 4.4), et on a :

$$E_{\star}(x) = \sum_{i=1}^k x_i (\max_{j < i} \pi_j - \max_{j < i} \pi_j)$$

$$E^{\star}(x) = \sum_{i=k}^n x_i (\max_{j > i} \pi_j - \max_{j > i} \pi_j)$$

où

$$x_1 < x_2 < \dots < x_n \quad \text{et} \quad x_k \in \overset{\circ}{F}$$

Si on décrit linguistiquement la conséquence de a, on peut l'exprimer par un intervalle flou \tilde{x} et les quantités $E^{\star}(x)$ et $E_{\star}(x)$ sont encore définies, en substituant aux formules ci-dessus leurs contreparties continues (cf. chap. 4)

$$E_{\star}(x) = \int_{-\infty}^{+\infty} x \, d\mu_{[\tilde{x}, +\infty)}$$

$$E^{\star}(x) = \int_{-\infty}^{+\infty} x \, d\mu_{\tilde{x}, +\infty)}$$

On notera qu'au lieu de calculer la conséquence espérée, on peut dans le cadre d'un modèle possibiliste manipuler directement l'intervalle flou des conséquences grâce aux techniques de l'arithmétique floue (voir section suivante).

Si on fait intervenir l'utilité des conséquences, sous la forme $u(x)$, on peut calculer, à l'aide du principe d'extension l'utilité floue de l'action potentielle a ; si la fonction d'utilité est imprécise définie sur l'échelle numérique S (par exemple de l'argent), le problème d'évaluation se pose alors comme suit /19/:

L'action a a pour conséquence possible x (possibilité $\pi(x)$) dont l'utilité est dans un intervalle $U(x)$, $\forall x \in \{x_1 \dots x_n\}$.

La distribution de possibilité π induit sur le domaine S des fonctions de possibilité inférieure et supérieure (cf. chap.1 § 4.5). Si on mesure l'intérêt de l'action a par les degrés de possibilité et de nécessité pour que l'utilité de a soit dans un sous-ensemble E de S on aura

$$\Pi(E) \in [\Pi_{\star}(E), \Pi^{\star}(E)] ; N(E) \in [1 - \Pi^{\star}(\bar{E}), 1 - \Pi_{\star}(\bar{E})]$$

avec

$$\Pi^{\star}(E) = \max_{i=1, n} \{ \pi_i / E \cap U(x_i) \neq \emptyset \}$$

$$\Pi_{\star}(E) = \max_{i=1, n} \{ \pi_i / U(x_i) \subseteq E \}$$

Quand E et $U(x_i) \forall i$ sont des intervalles flous, alors ces formules s'étendent naturellement, dans le cadre des possibilités de Zadeh :

$$\Pi^{\star}(E) = \max_{i=1, n} \min(\pi_i, \sup_s \min(\mu_E(s), \mu_{U(x_i)}(s)))$$

$$\Pi_{\star}(E) = \max_{i=1, n} \min(\pi_i, \inf_s \max(\mu_E(s), 1 - \mu_{U(x_i)}(s)))$$

Notons R la relation floue définie par $\mu_R(x_i, s) = \mu_{U(x_i)}(s)$. Alors si E est un singleton $\{s\}$, et en se rappelant que Π^{\star} est encore une fonction de possibilité, l'expression de $\Pi^{\star}(E)$ devient

$$\Pi^{\star}(\{s\}) = \pi^{\star}(s) = \max_{i=1, n} \min(\pi_i, \mu_R(x_i, s))$$

c'est-à-dire qu'on retrouve la composition sup-min des relations floues. En revanche $\Pi_{\star}(\{s\}) = 0$ en général (sauf si $\exists_i, \mu_i \neq 0$, et $\mu_R(x_i, s') < 1 \forall s' \neq s$); mais Π_{\star} n'est pas une mesure de possibilité.

6.2.5 - EVALUATION MULTICRITERE EN ENVIRONNEMENT INCERTAIN

Dans le cas classique ou l'incertitude sur l'environnement est représentée par une mesure de probabilité, il existe une théorie de l'évaluation multicritère qui est celle de l'utilité multidimensionnelle (Keeney & Raiffa /27/). Quand la méconnaissance des conséquences d'une action potentielle est exprimée par une distribution de possibilité sur X , il est naturel de faire appel à l'arithmétique floue pour évaluer cette action au sens de chaque critère. (Dubois et Prade /12/).

Soient $X_1 \dots X_p$ les échelles numériques où les évaluations $m_1 \dots m_p$ des p critères prennent leurs valeurs. Soit F l'ensemble flou des conséquences d'une action potentielle "a" définies sur $X = X_1 \times \dots \times X_p$. On supposera que $F = F_1 \times \dots \times F_p$ où $F_j = m_j(a)$ est l'ensemble flou des conséquences de a au sens du critère j, et \times dénote le produit cartésien d'ensembles flous. Cette hypothèse revient à supposer que le décideur fournit pour chaque dimension j pertinente son opinion sur les conséquences de a. $F_1 \times \dots \times F_p$ est alors le plus grand F dont les projections sont les F_j , $j = 1, p$ (cf. chap.3).

Si $\{G_1 \dots G_p\}$ est l'ensemble des objectifs de l'action que l'on veut choisir, objectifs définis sous forme de conséquences, on est amené pour chaque action candidate a à évaluer son degré de compatibilité avec les objectifs, compte tenu de l'incertitude sur ses conséquences. Ce degré sera un nombre flou $\mu_{G_j}(m_j(a))$ défini pour chaque critère j par le principe d'extension, et on retrouve ici le concept de compatibilité d'un ensemble flou par rapport à un autre (cf. chap.3). On a

$$\mu_{G_j}(m_j(a)) = \tau(G_j/F_j)$$

tel que

$$\forall u, \mu_{\tau(G_j/F_j)}(u) = \sup_{v=\mu_{G_j}(v)} \mu_{F_j}(v) \quad (42)$$

quand F_j est un nombre précis $m_j(a)$ on obtient bien le degré d'appartenance de ce nombre à G_j .

Si h est l'opérateur d'agrégation trouvé par la méthode décrite en section 1 l'évaluation globale de l'action a sera

$$\mu_{\mathcal{D}}(a) = h(\tau(G_1/F_1), \dots, \tau(G_p/F_p)) \quad (43)$$

où (43) se calcule à l'aide des résultats d'arithmétique floue du chapitre 4, par exemple si $h = \min$ on aurait

$$\begin{aligned} \mu_{\mathcal{D}}(a) &= \min_{j=1,p} \tau(G_j/F_j) \\ &= \tau(G_1 \times G_2 \dots \times G_p / F) \end{aligned}$$

NB : On pourrait bien sûr se contenter d'indices de comparaison scalaires entre F_j et G_j tels que $\Pi_{F_j}(G_j)$, $N_{F_j}(G_j)$ qui sont respectivement la possibilité et la nécessité de l'événement flou G_j , connaissant F_j (chap.2 § 4.3.2). Ces indices, on l'a vu au chapitre 4 (§ 4.4), se retrouvent sur le graphe de la fonction d'appartenance de $\tau(G_j/F_j)$.

6.3 - CLASSEMENT D' ACTIONS POTENTIELLES IMPRÉCISÉMENT ÉVALUÉES

Il est rare, en pratique, que l'évaluation globale d'une action provenant d'un processus d'agrégation de critères, ou de l'intégration de facteurs incertains, soit un nombre précis. En général elle sera naturellement représentée par un intervalle flou qui exprime l'imprécision et/ou l'incertitude des connaissances qui ont rendu possible le processus d'évaluation : imprécision de mesures, données linguistiques, définition incomplète du mode d'agrégation des critères, incertitude sur les conséquences attendues des actions.

Même dans le cas idéal où les actions potentielles seraient évaluées de façon précise, le classement qui en résulte doit être nuancé en vertu de ce que représente, pour le décideur un écart sur l'échelle d'évaluation. Il y a des seuils d'indifférence et de préférence stricte qui font que l'ordre naturel des nombres ne reflète pas exactement l'échelle de jugement du décideur. Ces seuils ont été introduits explicitement par Roy /34/ qui, lorsqu'ils existent de façon distincte, appelle pseudo-critère la fonction d'évaluation. On peut penser que de tels seuils ont en partie leur origine

dans le manque de confiance du décideur, à la fois quant à la pertinence des critères (numériques) utilisés et quant à la précision réelle des évaluations. Dans l'approche proposée ici ce dernier aspect est directement pris en compte, et la notion très affaiblie d'ordre qui existe sur l'ensemble des intervalles flous est supposée traduire les hésitations du décideur à départager certaines actions potentielles.

Dans cette section on passe en revue les méthodes que l'on peut utiliser pour trier parmi un ensemble d'actions imprécisément évaluées. On peut distinguer deux cas :

- chaque action a est évaluée par un nombre flou unique $M(a)$
- chaque action a est évaluée par un vecteur d'évaluations $M_1(a) \dots M_p(a)$ flous.

On suppose que plus $M(a)$ est grand, meilleure est l'évaluation. Par exemple $M(a)$ est le degré de compatibilité de a avec un objectif (section 1).

6.3.1 - COMPARAISON D'EVALUATIONS IMPRECISES

Les méthodes décrites par Dubois et Prade /16/ et rappelées au chapitre 4, section 4 pour comparer deux intervalles flous, et ordonner un ensemble de tels intervalles trouvent ici leur utilisation naturelle.

On rappelle qu'étant donné deux nombres flous $M(a)$ et $M(a')$ évaluant deux actions a et a' on définit les indices suivants : (chap. 4 § 4.2).

- possibilité de sur-classement de a' par a :

$$PS(a, a') = \Pi_{M(a)}([M(a'), +\infty)) \quad (44)$$

- nécessité de surclassement de a' par a

$$NS(a, a') = \mathcal{N}_{M(a)}([M(a'), +\infty)) \quad (45)$$

- possibilité de surclassement strict de a' par a

$$PSS(a, a') = \Pi_{M(a)}(]M(a'), +\infty)) \quad (46)$$

où $\Pi_{M(a)}$ et $\mathcal{N}_{M(a)}$ sont les mesures de possibilité et de nécessité au sens de Zadeh, de distribution $\mu_{M(a)} \cdot [M(a'), +\infty)$ (resp : $]M(a') + \infty)$) est l'ensemble de nombres possiblement (resp : nécessairement) plus grands que $M(a')$.

Chacun de ces indices définit sur A une relation floue de surclassement dont la sémantique est proche de celle qu'a introduite Roy /35/ récemment. Chacune de ces relations floues permet de définir une structure d'ordre plus ou moins strict sur A , et c'est en confrontant les trois classements ainsi obtenus qu'on peut réaliser un choix entre les actions potentielles. Il y a deux façons de traiter les relations floues PS, NS et PSS :

- soit construire des indices globaux de comparaison
- soit traiter comme telles les relations de surclassement.

6.3.1.1 - Définition d'ordre totaux sur A (cf. chap.4, § 4.3.1)

Soit T un sous-ensemble d'actions de A et $a \in A-T$. Si les évaluations étaient précises, le surclassement des actions de T par a s'écrirait

$$\forall a' \in T, M(a) \geq M(a')$$

ou encore

$$M(a) \geq \max \{M(a') / a' \in T\}$$

Dans le cas d'évaluations imprécises, cette approche revient à comparer $M(a)$ à la quantité $\max_{a' \in T} M(a')$ où \max est l'extension floue du maximum, définie au chap.4 § 2.5.5.

On est donc amené à considérer les quantités $PS(a, \forall a')$, $NS(a, \forall a')$, $PSS(a, \forall a')$ qui expriment respectivement la possibilité de surclassement, la nécessité de surclassement et la possibilité de surclassement strict par a , de toutes les autres actions, obtenues en substituant dans (44-46) $\max_{a' \neq a} M(a')$ à $M(a')$. On construit ainsi des ensembles flous d'actions possiblement dominantes, nécessairement dominantes et possiblement strictement dominantes, soit PD, ND, PSD avec

$$\mu_{PD}(a) = PS(a, \forall a') ; \quad \mu_{ND}(a) = ND(a, \forall a')$$

$$\mu_{PSD}(a) = PSS(a, \forall a')$$

On obtient donc trois ordres totaux sur A, induits par les trois ensembles flous. On remarquera que plus la hauteur de F est proche de 1, plus il est licite d'admettre l'existence d'actions dominantes au sens de $F = PD, ND, PSD$.

Ces ordres totaux sont les homologues du classement par "distillation descendante" dans la méthode ELECTRE III /35/. On peut aussi introduire l'homologue d'une "distillation ascendante" en cherchant des sous-ensembles flous d'actions possiblement dominées Pd, nécessairement dominées Nd, et possiblement strictement dominées PSD, en cherchant pour chaque action a, à comparer $M(a)$ avec $\min_{a' \neq a} M(a')$. On définit par exemple

$$\mu_{Pd}(a) = PS(\forall a', a) = \prod_{\substack{a' \neq a \\ \min M(a')}} \gamma(\overline{M(a)}, +\infty)$$

Nd, PSD sont définis en transformant de façon analogue (45) et (46). On notera que plus la hauteur de F est proche de 1, plus il est licite d'admettre que certaines actions sont dominées au sens de $F = Pd, Nd, PSD$. L'ordre induit par ces sous-ensembles flous sera donc globalement inverse de celui induit par PD, ND, PSD. On est donc amené pour mettre en vis-à-vis les deux classements induits pour chaque type d'indice de crédibilité, à comparer les ensembles flous PD et Pd ND et Nd, PSD et PSD.

Pour cette comparaison, on peut faire appel à des indices de similarité scalaire évoqués au chap.2, § 4.2 ; on peut aussi faire une comparaison point par point à l'aide de l'indice flou de compatibilité de Zadeh. On peut chercher les discordances entre chaque classement pour valider la position de chaque action, comme dans Roy /35/. Enfin on peut "compter" le nombre d'actions dominantes ou dominées à l'aide de la notion de cardinal flou discutée au chapitre 3.

6.3.1.2 - Méthodes de traitement de relations de surclassement floues

L'étude des relations floues de préférence a été abordée dans la littérature par plusieurs auteurs, qui indiquent des méthodes pour dégager un classement effectif d'un ensemble d'actions potentielles. Ces méthodes dépendent de la représentation adoptée pour la relation floue et notamment comment sont - éventuellement - exprimées les propriétés de réflexivité, antisymétrie ou complétude, transitivité. On résume ci-dessus les diverses approches existantes, pour pouvoir indiquer celles qui peuvent être utilisées dans

le traitement des relations PS, NS, et PSS issues de la comparaison de quantités imprécises.

Zadeh /48/ considère des relations d'ordre partiel flou dont l'interprétation est la suivante :

$$\mu_R(a, a') \in [0, 1]$$

$$\mu_R(a, a') = 1, \quad \mu_R(a', a) = 0 \quad \leftrightarrow \quad a \text{ surclasse (complètement) } a'$$

$$\mu_R(a, a') = \mu_R(a', a) = 1 \quad : \quad \text{indifférence entre } a \text{ et } a'$$

$$\mu_R(a, a') = \mu_R(a', a) = 0 \quad \text{incomparabilité de } a \text{ et } a'.$$

Zadeh suggère les axiomes suivants :

$$\mu_R(a, a) = 1 \quad (\text{reflexivité})$$

$$a \neq a' \rightarrow \min(\mu_R(a', a), \mu_R(a, a')) = 0 \quad (\text{antisymétrie parfaite})$$

$$R \circ R \subseteq R \quad (\text{transitivité}) \quad \text{où } \circ \text{ est la composition sup-min}$$

Sous ces hypothèses R peut être vu comme un ensemble de graphes flous acycliques et on peut donc trouver des éléments dominants pour R. Kaufmann (1973) affaiblit l'antisymétrie parfaite en

$$\forall (a, a') \quad \mu_R(a, a') > 0 \rightarrow \mu_R(a, a') \neq \mu_R(a', a) \quad (47)$$

(antisymétrie)

Dans ce cas la domination de a par a' est représentée par

$$\mu_R(a, a') > \mu_R(a', a) \quad (48)$$

et on peut donc construire à partir de (48) un ordre partiel au sens usuel. Quand (47) n'est pas valide on a un préordre flou. Le problème est alors d'interpréter l'existence de nombres $\alpha \in]0, 1[$ tels que

$$\mu_R(a, a') = \mu_R(a', a) = \alpha \quad (49)$$

Ovlowski (/31/) assimile l'indifférence à l'incomparabilité pour interpréter (49). Dans ce cas, sur le graphe flou associé à R, il élimine les arcs (a, a') tels que

$$a \neq a' \quad \text{et} \quad \mu_R(a, a') \leq \mu_R(a', a)$$

En étudiant la relation R' antisymétrique parfaite telle que

$$\mu_{R'}(a, a') = \max(0, \mu_{R'}(a, a') - \mu_{R'}(a', a))$$

Orlowski montre ainsi l'existence, pour R , d'éléments non dominés a tels que

$$\forall a' \in A, \mu_R(a, a') \geq \mu_R(a', a)$$

qui sont ceux de R' , au sens des graphes flous acycliques qui sont la représentation de R' . Kaufmann /25/ interprète (49) comme l'indifférence entre a et a' . Il est alors amené à rechercher les ensembles $E(a)$ d'actions indifférentes à a , appelées sous-classes maximales de similitude qui forment une couverture de A . Quand elles forment une partition le préordre est alors dit réductible au sens où on se ramène à un ordre partiel entre classes de similitude. Sinon on ne sait pas conclure, dans le cadre de cette interprétation.

Ce type de relation floue de préférence à un inconvénient : celui de supposer la transitivité au sens "sup-min" qui est une condition très forte. En général, s'il est déjà difficile d'obtenir la transitivité d'une relation d'ordre classique déduite d'un questionnaire, il sera encore plus difficile de l'obtenir d'une relation floue. Par exemple la relation PS de possibilité de dominance entre intervalles flous est compatible avec l'interprétation décrite ci-dessus, mais n'est pas transitive.

Une autre difficulté réside dans la façon de traiter ces relations floues. On n'utilise jamais les degrés de surclassement. Les graphes flous acycliques sont analysés sur leur structure, et non d'après la valuation des arcs.

Shimura /39/ a étudié des relations floues ayant la même interprétation que celles de Zadeh, possédant la réflexivité, et vérifiant un axiome dual de l'antisymétrie parfaite :

$$\forall a, a' \quad \max(\mu_R(a, a'), \mu_R(a', a)) = 1 \quad (50)$$

Il ne suppose pas la transitivité, mais définit un degré de surclassement de a sur les éléments d'un sous-ensemble T de A par

$$\mu_R(a, T) = \min\{\mu_R(a, a') \mid a' \in T\}$$

On peut alors définir le sous-ensemble flou des éléments dominants de A par

$$\forall a, \mu_D(a) = \min\{\mu_R(a, a') / a' \in A\} \quad (51)$$

Si D est normalisé son noyau contient les éléments dominants de A. Sinon, la quantité $\max_{a \in A} \mu_D(a)$ mesure à quel point la relation R permet de conclure à l'existence d'éléments non dominés dans A. Cette existence est assurée par une condition de transitivité de R qui s'écrit :

$$\forall a, a', a'' \in A, \quad \begin{aligned} & \mu_R(a, a') > \mu_R(a', a) \quad \rightarrow \mu_R(a, a'') > \mu_R(a'', a) \\ & \mu_R(a', a'') > \mu_R(a'', a') \end{aligned} \quad (52)$$

on vérifie que (52) $\rightarrow \overset{\circ}{D} \neq \emptyset$, c'est-à-dire que (51) légitime l'existence d'éléments non dominés dans A.

(Pour le voir il suffit de noter que $\mu_R(a, a') > \mu_R(a', a) \rightarrow \mu_R(a, a') = 1$; donc le noyau de R est un préordre dès que (52) est valide).

Roy /35/ considère des relations de surclassement flou dont l'interprétation est proche de celles considérées par Zadeh, mais dégagées complètement d'hypothèses de transitivité et d'antisymétrie, ce qui facilite la mise en oeuvre de cette approche. Le traitement de la relation de surclassement, qui exprime une crédibilité plutôt qu'une intensité de surclassement est fondé sur l'existence d'un seuil de perception s tel que la crédibilité du surclassement de a' par a est suffisamment forte dès que

$$\mu_R(a, a') \geq \mu_R(a', a) + s(\mu_R(a, a'))$$

L'algorithme de classement est double; il considère le nombre d'actions surclassées par a, et le nombre d'actions que surclasse a. On aboutit à deux ordres partiels dont la significations se complète pour permettre un classement final.

Blin /3/ et Dubois-Prade /14/ considèrent des relations d'ordre flou dites "de tournoi" vérifiant la propriété d'antisymétrie.

$$\forall a, a' \quad \mu_R(a, a') + \mu_R(a', a) = 1 \quad (53)$$

Elles s'interprètent alors ainsi :

a surclasse complètement a' : $\mu_R(a, a') = 1$

a est surclassé complètement par a' : $\mu_R(a, a') = 0$

indifférence ou incomparabilité : $\mu_R(a, a') = .5$

Si R est une relation d'ordre classique (53) exclut l'indifférence pour $a \neq a'$. Blin (1974) suppose que le noyau de R est une relation transitive, et forme donc un ordre partiel. Il cherche alors l'ordre total $*$ qui soit le plus possible compatible avec la partie nuancée de R , parmi ceux qui sont compatibles avec l'ordre partiel R . Dubois et Prade (1980) reliant (53) à l'inclusion d'ensemble basée sur l'implication logique multivalente $\max(1-a, b)$, suggèrent de classer les éléments de A au sens de l'inclusion des ensembles flous d'actions surclassées par a , $\forall a \in A$, soit $S(a)$ tels que

$$\mu_{S(a)}(a') = \mu_R(a, a') \quad \forall a'$$

Pour cette inclusion d'ensemble, notée \prec voir Dubois et Prade (/14/), chapitre II.1 et le chapitre 2 du présent mémoire (§ 3.5.1). dans /16/ on montre qu'une condition suffisante pour que \prec définisse sur $\{S(a)/a \in A\}$ un ordre partiel est que

$$\forall a, a' \quad S(a) \prec S(a') \quad \text{ou} \quad S(a') \prec S(a) \quad (54)$$

Dans ce cas, l'ordre partiel \prec induit un ordre partiel sur A .

NB1: Dans sa méthode, Roy /35/ utilise un seuil pour donner un sens au nombre d'actions potentielles surclassées par ou dominant une action a . Il semble que les notions de cardinal d'un ensemble flou (scalaire ou flou) puissent être utilisées avec profit pour éviter le recours à ce seuil. Par exemple, avec les notations ci-dessus, le nombre d'actions dominées par a sera $|S(a)|$ (cf. chap.2 § 4.1.3) ou l'intervalle flou d'entiers $||S(a)||$ (chap.3 § 6.1). L'arithmétique floue permet ensuite le traitement numérique de ces quantités.

NB2: La méthode de Saaty /36/ utilise des relations floues à valeurs sur \mathbb{R} et l'antisymétrie de R est exprimée par

$$\mu_R(a, a') \cdot \mu_R(a', a) = 1 \quad (55)$$

La transitivité de R signifie

$$\mu_R(a, a') \cdot \mu_R(a', a'') = \mu_R(a, a'')$$

le classement des éléments de a s'effectue au vu des valeurs $w(a)$ d'un vecteur propre de la matrice μ_R . Cette méthode n'est donc pas utilisable telle quelle pour classer des nombres flous.

6.3.1.3 - Application aux relations de surclassement d'intervalles flous

On rappelle que les relations PS, NS, PSS définies à partir des indices de comparaison entre intervalles flous vérifient les propriétés (cf. chap.4 § 4.3.2)

$$\max(\text{PS}(a, a'), \text{PS}(a', a)) = 1 \quad (55)$$

$$\text{NS}(a, a') + \text{NS}(a', a) = 1 \quad (56)$$

$$\text{PSS}(a, a') + \text{PSS}(a', a) = 1 \quad (57)$$

En vertu de (55) la relation PS satisfait la condition demandée par la méthode de Shimura (1973). Cette méthode semble donc adaptée au traitement de la relation PS.

Néanmoins Dubois et Prade /16/ ont fait remarquer que le classement obtenu par la méthode Shimura /39/ est le même que l'ordre total défini par l'ensemble flou PD calculé à la section précédente, à partir d'une comparaison entre $M(a)$ et $\max_{a' \neq a} M(a')$ (voir aussi chap.4 § 4.3.2). Autrement dit

$$\mu_{\text{PD}}(a) = \min_{a'} \mu_{\text{PS}}(a, a')$$

On peut montrer de la même façon que l'ensemble flou Pd obtenu en comparant $M(a)$ et $\min_{a' \neq a} M(a')$ est tel que

$$\mu_{\text{Pd}}(a) = \min_{a'} \mu_{\text{PS}}(a', a)$$

De sorte que cette méthode de traitement de la relation de possibilité de surclassement a une interprétation satisfaisante en termes de comparaison de nombres flous.

En revanche les équations (56) et (57) suggèrent l'utilisation, pour traiter les relations de nécessité de surclassement et de possibilité de surclassement strict, de la méthode de comparaison des classes floues d'actions dominées $S(a)$, à l'aide de l'inclusion floue \prec , suggérée plus haut. Dubois et Prade /16/ ont montré que généralement (par exemple avec des évaluations floues des actions de A dont les fonctions d'appartenance sont continues) la condition (54) est vérifiée (cf. Proposition 4.9, chapitre 4), c'est-à-dire que l'inclusion \prec définira toujours un ordre partiel entre les classes $(S(a)/a \in A)$.

Dualement, on peut considérer l'ensemble flou $\Sigma(a)$ d'actions qui surclassent a , et étudier l'ordre partiel défini dans $\{\Sigma(a)/a \in A\}$ par la relation \prec . $\Sigma(a)$ est simplement défini par

$$\forall a' \quad \mu_{\Sigma(a)}(a') = \mu_R(a', a)$$

Le classement des $\Sigma(a)$ fournit un résultat analogue à une "distillation descendante" dans ELECTRE III.

La méthode ELECTRE III /35/ est d'ailleurs une alternative possible aux procédures spécifiques décrites ci-dessus, puisqu'elle ne pré-suppose aucune propriété pour la relation à traiter. De plus les relations PS, NS et PSS sont bien des cas particulier de relation de surclassement flou de Roy /35/ puisqu'elles expriment bien un degré de crédibilité de surclassement, qui est tantôt un degré de possibilité tantôt de nécessité.

Au total si on compare les méthodes de surclassement avec données précises à celles qui utilisent des données imprécises, on constate que dans ce dernier cas, on aboutit à 6 ordres partiels (au lieu de 2) ayant chacun sa sémantique (descendant/ascendant ; possibilité/nécessité ; surclassement faible ou strict).

6.3.2 - SURCLASSEMENT FLOU MULTICRITERE

Supposons que les actions potentielles aient été évaluées selon plusieurs critères, de façon homogène, c'est-à-dire au sens des degrés de compatibilité entre ces actions et divers objectifs. Pour parvenir à un classement final on choisit entre deux procédures :

- agréger les diverses évaluations partielles, selon l'approche de la section 1 et construire des relations de crédibilité de surclassement à l'aide d'indices de comparaison d'intervalles flous. C'est le propos de la section 3.1 qui précède.

- construire pour chaque critère des relations de crédibilité de surclassement et les agréger. C'est ce qu'il nous reste à envisager.

Soit $M_1(a) \dots M_p(a)$ les évaluations de a au sens de p critères, et

$$\{(PS_i, NS_i, PSS_i) / i=1, p\}$$

les relations floues induites par la comparaison par paire des $\{M_i(a) / a \in A\}$. Le problème d'agrégation consiste à trouver trois opérations h_1 , h_2 et h_3 pour construire des relations floues de surclassement global PS, NS, PSS

$$PS = h_1(PS_1, PS_2 \dots PS_p)$$

$$NS = h_2(NS_1, NS_2 \dots NS_p)$$

$$PSS = h_3(PSS_1, PSS_2 \dots PSS_p)$$

Une approche possible est celle proposée par Roy /35/ qui utilise la notion d'indices de concordance et de discordance. L'indice de concordance est la somme pondérée des crédibilités de surclassement de a' par $a \forall a, a'$:

$$C(a, a') = \sum_{i=1}^p p_i \mu_{R_i}(a, a') \quad (58)$$

où μ_R remplace l'un des symboles PS, NS, PSS.

L'indice de discordance $D_j(a, a')$ pour le critère j exprime à quel point a' domine a au sens du critère j et dépend de l'écart entre $M_j(a)$ et $M_j(a')$.

En particulier

$$D_j(a, a') = 0 \quad \text{si} \quad \mu_{R_j}(a, a') > 0$$

$$D_j(a, a') = 1 \quad \text{si} \quad \mu_{R_j}(a, a') = 0 \quad \text{et} \quad M_j(a') \text{ est plus grand que } M_j(a) \text{ plus un seuil de veto.}$$

Le sens du seuil de veto est de "refuser toute crédibilité au surclassement de a' par a même si tous les autres critères sont en concordance avec ce surclassement" /35/ dès que la différence $M_j(a') - M_j(a)$ dépasse ce seuil.

Roy /35/ propose alors de définir h par

$$\mu_R(a, a') = C(a, a') \prod_{i=1}^p \frac{1 - D_j(a, a')}{1 - C(a, a')} \quad (59)$$

Ce type d'agrégation pourrait en fait entrer dans le cadre défini en section 1. Il est clair que la formule (59) tente d'allier les aspects qui confortent l'idée de surclassement (formule 58) et ceux qui en font douter (les D_j).

Cette constatation induit une structure de l'opération d'agrégation h qui se construit en combinant un indice de concordance et un indice de discordance $D(a, a')$ construit sur les D_j . Le cadre conceptuel et le catalogue d'opérations d'agrégations proposés en section 1 pourrait certainement permettre d'étendre (59) en préservant cette structure.

Dans le cadre spécifique défini par les relations PS, NS, et PSS on peut souhaiter agréger en préservant la nature de ces relations.

Pour ce qui est des relations de type PS, on voudrait trouver h tel que PS satisfasse la condition (55). Faute de mieux on propose ici une approche ad hoc: si $PS_i(a, a') = 1$, la quantité $1 - PS_i(a', a)$ reflète à quel point a surclasse effectivement a' . Une évaluation globale du surclassement de a' par a est donc, en version non pondérée :

$$SG(a, a') = \sum_{i=1, p} \max(0, PS_i(a, a') - PS_i(a', a)) \quad (60)$$

Pour retrouver une relation de type PS on peut alors poser (comme le fait Shimura /39/)

$$PS(a, a') = \min\left(1, \frac{SG(a, a')}{SG(a', a)}\right) \quad (61)$$

Pour interpréter (60), notons $\mathcal{C}_{aa'}$, l'ensemble flou des critères pour lesquels a peut surclasser a' , tel que

$$\forall i = 1, p \quad \mu_{\mathcal{C}_{aa'}}(i) = PS_i(a, a')$$

Sous le signe somme de (60) on trouve une différence ensembliste $\mathcal{C}_{a, a'} - \mathcal{C}_{a', a}$, c'est-à-dire qu'on trouve l'ensemble des critères pour lesquels a peut surclasser a' mais sans que a' puisse surclasser a . $SG(a, a')$ est alors une évaluation scalaire de $\mathcal{C}_{aa'} - \mathcal{C}_{a'a}$, soit $g(\mathcal{C}_{a, a'} - \mathcal{C}_{a', a})$ au sens du chapitre 2 § 4.1.4. A l'aide des résultats sur la différence ensembliste floue (chap.2 § 3.5) et des évaluateurs existentiels et/ou universels, on peut imaginer de nombreuses variantes de (60) qui auront les mêmes propriétés. On remarque que $SG(a, a')$, dans (60) n'est que le cardinal scalaire de $\mathcal{C}_{aa'} - \mathcal{C}_{a'a}$ (cf. chap.2) (61) n'est qu'une façon de normaliser.

Dans le cas des relations de type NS ou PSS, il est clair que $NS_i(a, a')$ suffit à caractériser le surclassement de a' par a au sens du critère i . On souhaite que l'agrégation h soit telle que

$$NS(a, a') = 1 - NS(a', a)$$

après agrégation. Le chapitre 2 a déjà fourni une telle opération, à savoir une somme symétrique de Silvert, qui vérifie

$$\sigma(1-a, 1-b) = 1 - \sigma(a, b).$$

Dans ce cas, en prenant pour h la version n -aire d'une somme symétrique on vérifie en effet que

$$\begin{aligned} NS(a, a') &= h(NS_1(a, a'), \dots, NS_p(a, a')) \\ &= h(1-NS_1(a', a), \dots, 1-NS_p(a', a)) \\ &= 1 - NS(a', a). \end{aligned}$$

Les résultats de représentation et d'analyse des sommes symétriques (chap.2 § 3.4.2) peuvent permettre de choisir parmi tout un ensemble d'opérations (éventuellement paramétrées) pour agréger les degrés de crédibilité de surclassement. En particulier on peut choisir pour h autre chose que la moyenne arithmétique (pondérée ou non),

REFERENCES

- /1/ ADAMO J.M. (1980)
Fuzzy decision trees, Fuzzy sets and systems 4, 207-219
- /2/ BELLMAN R.E., ZADEH L.A. (1970)
Decision-making in a fuzzy environment ; Management Science 17, B141-B164.
- /3/ BLIN J.M. (1974)
Fuzzy relations in group-decision theory. J. Cybernetics 4 n°2, 17-22.
- /4/ BOSCHI R.A.A. (1982)
Modelling exploratory research. Europ J. Operations Research 6, n°3 250-259
- /5/ BORISOV A.N., MERKURYEVA G.V. (1982)
Linguistic lotteries-construction and properties. BUSEFAL n°11, 39-46.
- /6/ BOY G. (1983)
Le système MESSAGE : un premier pas vers l'analyse assistée par ordinateur des interactions homme/machine, A paraître dans Travail Humain.
- /7/ DEBREU G. (1960)
Topological methods in cardinal utility theory in Mathematical methods in the Social Sciences (K.J. Arrow, S. Karlin, P. Suppes, eds) Stanford Univ. Press. Stanford, 16-26.
- /8/ DUBOIS D. (1979)
Sélection par spécifications floues, Note interne ADEPA, Montrouge.
- /9/ DUBOIS D. (1980)
Vers l'application de la théorie des ensembles flous à la conception aidée par ordinateur. Bull. AF. Micado, Meylan, Mars 1980. Aussi : BUSEFAL n°2, 79-83.
- /10/ DUBOIS D. (1980)
Ensembles flous et conception assistée par ordinateur : rappels théoriques, perspective et exemple. Rapport de recherche IMAG n°199, Grenoble.

- /11/ DUBOIS D. (1982)
Processing vague queries in man-machine systems : a fuzzy set approach, IFAC Symp. on Theory and Appl. of Digital Control, New Delhi.
- /12/ DUBOIS D., PRADE H. (1979)
Criteria aggregation and ranking of alternatives in the framework of fuzzy set theory, à paraître dans TIMS Studies in the Management Sciences.
- /13/ DUBOIS D., PRADE H. (1979)
Decision-making under fuzziness in Advances in Fuzzy set theory and applications (M.M. Gupta, R.K. Ragade, R.R. Yager, eds) North Holland, 279-303.
- /14/ DUBOIS D., PRADE H. (1980)
Fuzzy Sets and Systems, Theory and Applications, Academic Press, N. Y.
- /15/ DUBOIS D., PRADE H. (1981)
Additions of interactive fuzzy numbers, IEEE Trans. Automatic Control 26, 926-936.
- /16/ DUBOIS D., PRADE H. (1982)
Ranking fuzzy numbers in the setting of possibility theory, Soumis à Infor. Sciences.
- /17/ DUBOIS D., PRADE H. (1982)
The use of fuzzy numbers in decision analysis, in Fuzzy Information and Decision Processes (M.M. Gupta, E. Sanchez, eds) North Holland 309-321
- /18/ DUBOIS D., PRADE H. (1982)
Upper and lower possibilistic expectations and their applications, 4th Int Seminar on Fuzzy Set Theory, J. Kepler Universität Linz, Autriche.
- /19/ DUBOIS D., PRADE H. (1983)
Upper and lower possibilities induced by a multivalued mapping, IFAC Symposium on Fuzzy Information, Knowledge representation and decision analysis, Marseille.
- /20/ FARQUHAR P.H. (1977)
A survey of multiattribute utility theory and its applications. TIMS Studies in the Management Sciences, 6, 59-89.

- /20/ FINE T. (1973)
Theories of probability. Academic Press, N.Y.
- /21/ FISHBURN P.C. (1970)
Utility theory for decision-making. Wiley, New-York.
- /22/ FREELING A.N.S. (1980)
Fuzzy sets and decision analysis. IEEE Trans. Systems Man and Cybernetics 10, 341-354.
- /23/ FUNG L.W., FU K.S. (1975)
An axiomatic approach to rational decision-making in a fuzzy environment. In "Fuzzy Sets and Their Application to cognitive and decision processes (L.A. Zadeh et al, eds). Academic Press 227-256.
- /24/ GILES R. (1982)
Foundations for a theory of possibility in Fuzzy Information and Decision Processes (M.M. Gupta, E. Sanchez, Eds). North Holland, 183-196.
- /25/ KAUFMANN A. (1973)
Introduction à la théorie des sous-ensembles flous, vol.I, Masson, Paris.
- /26/ KEENEY R.L. (1974)
Multiplicative utility functions, Operations Research, 22, 22-34.
- /27/ KEENEY R.L., RAIFFA H. (1976)
Decisions with multiple objectives preferences and value trade-offs. J. Wiley, New York.
- /28/ KENDALL J.N. (1975)
Hard and Soft constraints in linear programming. Omega, 3, 709-715.
- /29/ KRANTZ D.H. LUCE R.D., SUPPES P., TVERSKY A. (1971)
Foundations of measurement, Vol.1, Academic Press N.Y.
- /30/ NORWICH A.M., TURKSEN I.B. (1982)
The fundamental measurement of fuzziness in "Fuzzy set and Possibility theory : Recent developments" (R. Yager, ed.) Pergamon Press 49-60.

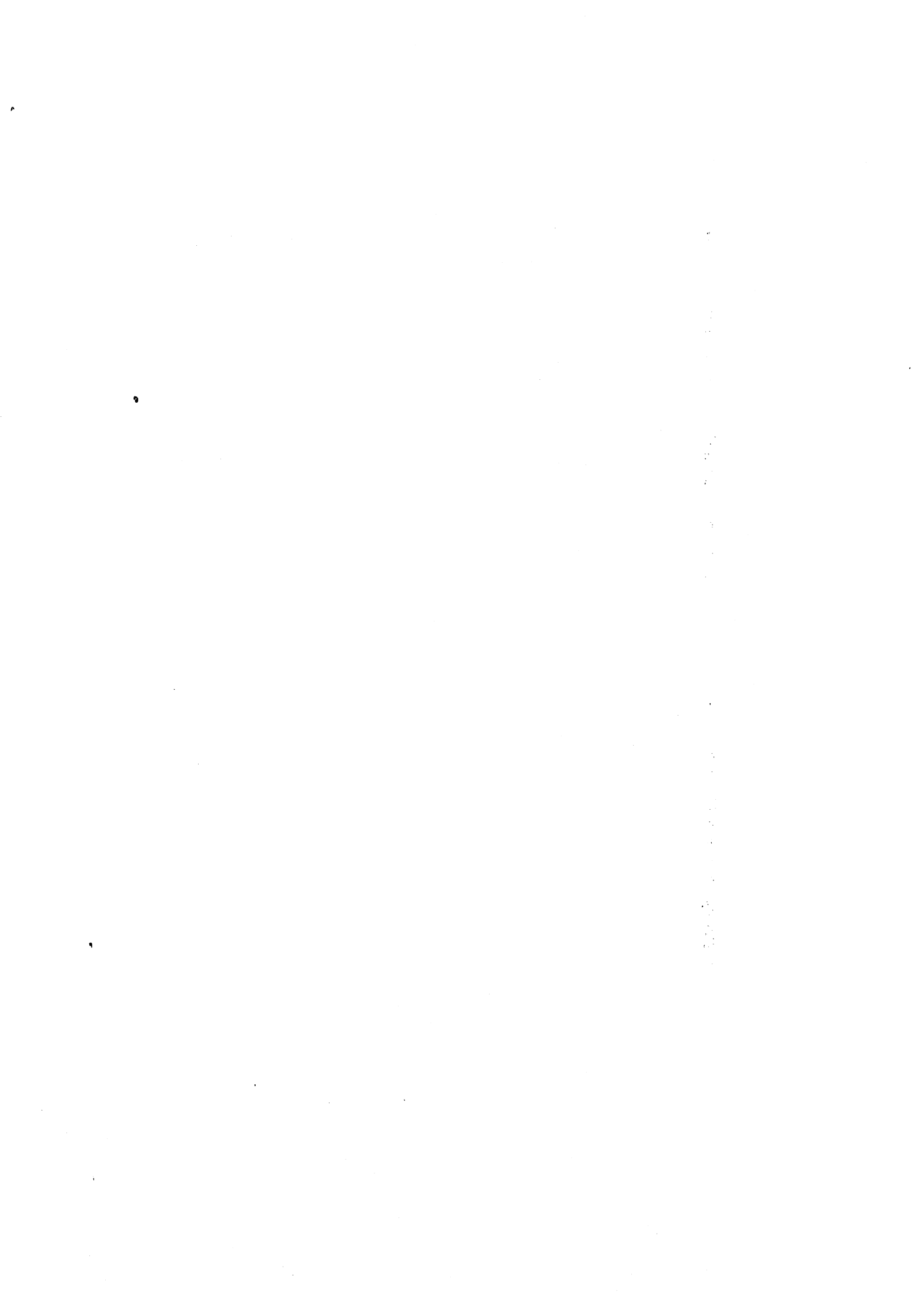
- /31/ ORLOWSKI S.A. (1978)
Decision-making with a fuzzy preference relation, Fuzzy Sets & Syst., 1, 157-167.
- /32/ ROY B. (1972)
Décisions avec critères multiples : problèmes et méthodes Revue METRA, vol.11, n°1, 121-152.
- /33/ ROY B. (1974)
Critères multiples et modélisation des préférences : l'apport des relations de surclassement. Revue d'Economie Politique 1, 1-44.
- /34/ ROY B. (1977)
A conceptual framework for a prescriptive theory of decision-aid. TIMS Studies in the Management Sciences, vol.6, 179-210.
- /35/ ROY B. (1978)
ELECTRE III : Un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples. Cahiers du CERO, 20, 3-24.
- /36/ SAATY T.L. (1978)
Exploring the interfaces between hierarchies, multiple objectives, and fuzzy sets, Fuzzy Sets and Syst. 1, 57-68.
- /37/ SAGE A.P. (1981)
Sensitivity Analysis in Systems for planning and decision support. J. Franklin Institute, 312, 265-291.
- /38/ SAVAGE L.J. (1972)
The foundations of statistics, Dover, New York.
- /39/ SHIMURA M. (1973)
Fuzzy set concepts in the rank-ordering of objects. J. Math. Anal. Appl. 43, 717-733.
- /40/ STEUER R.E. (1977)
An interactive multiple objective linear programming procedure, TIMS Studies in the Management Sciences, 6, 225-240.
- /41/ SUGENO M. (1977)
Fuzzy measures and fuzzy integrals : a survey in Fuzzy Automata and Decision Processes (M.M. Gupta, G.N. Sarichis, B.R. Gaires Eds) North Holland 89-102.

- /42/ TERANO T., SUGENO M. (1975)
Conditional fuzzy measures and their applications in Fuzzy Sets and their applications to Cognitive and Decision Processes
 (L.A. Zadeh et al. Eds) 151-170, Academic Press.
- /43/ VON NEUMANN J., MORGENSTERN O. (1944)
Theory of Games and Economic Behavior. Princeton Univ. Press,
 Princeton N.J.
- /44/ WATSON S.R., WEISS J.J., DONNELL M. (1979)
Fuzzy decision analysis. IEEE Trans. Systems Man and Cybernetics,
 9, 1-9.
- /45/ WHALEN T.H. (1980)
Risk minimization using L-fuzzy sets, IEEE Int. Conf. Cybernetics and Society.
- /46/ YAGER R.R. (1977)
Multiple-objectives decision-making using fuzzy sets. Int. J. Man-Machine Studies 9, 375-382.
- /47/ YAGER R.R. (1979)
Possibilistic decision-making. IEEE Trans. Systems Man Cybernetics 9, 388-393.
- /48/ ZADEH L.A. (1971)
Similarity relations and fuzzy orderings, Information Sciences 3, 177-200.
- /49/ ZADEH L.A. (1982)
Fuzzy probabilities and their role in decision analysis. IFAC Symposium on Theory and Applications of Digital control New-Delhi.
- /50/ ZELENY M. (1977)
Adaptive displacement of preferences in decision-making. TIMS Studies in the Management Sciences 6, 147-157.
- /51/ ZIMMERMANN H.J., ZYSNO (1980)
Latent connectives in human decision-making. Fuzzy Sets & Syst. 4, 37-51.
- /52/ ZYSNO P. (1982)
The integration of concepts within judgmental and evaluative processes, Progress in Cybernetics and Systems Research 8, 509-516, Hemisphere Pub. Corp.



CHAPITRE 7

L'INTÉRÊT DES INTERVALLES FLOUS
EN RECHERCHE OPÉRATIONNELLE.



Dans ce dernier chapitre le concept d'intervalle (ou de nombre) flou est appliqué à quelques problèmes classiques rencontrés en recherche opérationnelle : cheminement dans les graphes, ordonnancement, optimisation, files d'attente

On se rappelle que l'arithmétique floue généralise directement le calcul d'intervalles ; elle permet d'analyser la sensibilité d'un système aux variations d'un paramètre dont la valeur est mal définie ou mal connue ; l'intérêt de modéliser cette mauvaise connaissance par un intervalle flou est de pouvoir d'un seul coup prendre en compte des estimations optimistes et pessimistes du paramètre en question.

Suivant le type de problème, l'arithmétique floue peut fournir deux types d'informations :

- quand il s'agit d'évaluer (des performances par exemple), elle fournit une analyse de sensibilité nuancée. Dans ce cas, on fournit à un algorithme d'évaluation des données vagues sous la forme d'intervalles flous, et on calcule l'influence de cette imprécision sur les résultats de l'algorithme à l'aide des techniques décrites dans le chapitre 4 notamment. Il est alors normal que ces résultats deviennent flous. C'est le cas pour les problèmes de cheminement.

- Quand il s'agit de prendre une décision (précise), on utilise un modèle du type de celui qui a été étudié au chapitre précédent. En substance, on est ramené à confronter une évaluation (imprécise) à un objectif (flou) pour soit préciser la valeurs de paramètres qui conduisent à s'approcher le plus possible de l'objectif, soit pour répondre à des questions du type : est-ce-que le système étudié aura telle ou telle propriété ?

L'autre modèle couramment utilisé d'informations mal connues est le modèle probabiliste. Il convient donc de préciser ce qu'on peut attendre de l'arithmétique floue par comparaison avec une analyse probabiliste.

Rappelons d'abord que la représentation d'un paramètre à l'aide d'une distribution de probabilité suppose qu'on ait pu procéder à un recueil de données statistiques sur les valeurs que prend "habituellement" ce paramètre. On suppose aussi que les fréquences ainsi recueillies, relatives à

l'occurrence de chaque événement possible, sont assez proches des valeurs asymptotiques de probabilité. L'analyse probabiliste concerne donc les phénomènes reproductibles et s'étant suffisamment reproduits.

L'introduction de distributions de probabilité dans les algorithmes d'évaluation a théoriquement pour but de calculer une distribution induite sur les résultats. En pratique cela s'avère généralement trop difficile à mener à bien, et on se contente souvent de calculer la valeur moyenne des résultats et parfois leur variance. Encore ce calcul est-il lui-même parfois très complexe.

Dans un algorithme de décision, on étudie la probabilité pour qu'un système satisfasse à un objectif ou possède une propriété donnée. Cette probabilité s'entend comme le rapport entre le nombre de cas favorables et le nombre d'expériences réalisées avec autant d'exemplaires de ce système.

L'utilisation d'intervalles flous ne présuppose pas nécessairement l'existence de données statistiques objectives. Un ensemble flou peut avoir une interprétation statistique, on l'a vu, dans le cadre d'expériences fournissant des résultats imprécis et cohérents. Mais un intervalle flou modélise aussi une contrainte souple, fournie par exemple par un expert. L'information fournie par l'utilisation d'ensembles flous est moins précise que celle fournie par une analyse probabiliste puisque l'on n'a alors accès qu'à des bornes plus ou moins strictes sur les valeurs de probabilité ; les bornes supérieures évaluent le "possible", les bornes inférieures le "nécessaire", c'est-à-dire par exemple les performances réalisables et celles qu'on est certain de voir réalisées, respectivement. Cette interprétation est cohérente avec celle de l'analyse de sensibilité. Au lieu de raisonner en termes de comportement moyen, l'arithmétique floue manipule les catégories "possible" et "nécessaire" et vise à une analyse de pire et de meilleur cas.

Une évaluation fondée sur l'arithmétique floue a donc un sens même pour des phénomènes non répétables, ou "nouveaux", sur lesquels seule l'expérience humaine peut renseigner. Ce type de situation se rencontre parfois en recherche opérationnelle.

Par ailleurs les calculs que nécessite l'arithmétique floue sont souvent plus légers que ceux qu'une étude probabiliste, fût-elle

grossière, exige. Le chapitre 4 a démontré que la difficulté rencontrée pour travailler avec des intervalles flous est comparable, voire parfois strictement équivalente, à la mise en oeuvre d'une analyse de sensibilité à l'aide d'intervalles d'incertitude.

7.1 - PROBLÈMES DE CHEMINEMENT SUR DES GRAPHS IMPRÉCISÉMENT VALUÉS

Un premier exemple d'utilisation de l'arithmétique floue concerne l'extension des algorithmes de plus court chemin quand la longueur des arcs sur un graphe est imprécise. Cet exemple est particulièrement riche car il est l'occasion de bien comprendre les utilisations respectives de l'arithmétique floue optimiste et de l'arithmétique floue pessimiste. C'est aussi un cas où l'approche probabiliste est très difficile à mettre en oeuvre dans le cas général, et où les résultats de l'analyse floue peuvent être significatifs (pas trop flous).

7.1.1 - ARITHMETIQUE FLOUE OPTIMISTE OU PESSIMISTE

Considérons un problème élémentaire d'occupation du temps : un homme désire prendre l'avion de 10 h. du matin, et souhaite, compte tenu de son heure de lever, et connaissant la suite de n tâches qu'il doit accomplir avant de monter dans l'avion, savoir s'il y parviendra, un certain jour. Soit

t_L l'heure du lever (date au plus tôt de début de la 1ère des n tâches)

d_i la durée de la tâche i

t_A l'heure d'arrivée à l'aéroport

L'homme connaît approximativement la durée des tâches qu'il a coutume d'accomplir chaque jour, Il exprime ces durées à l'aide d'intervalles flous D_i (spécifiques du jour concerné). La durée totale des tâches est donc

$$D = D_1 \oplus D_2 \dots \oplus D_n$$

A partir de là on peut souhaiter répondre à plusieurs questions :

- Connaissant l'heure du lever t_L , limitée par le nombre flou T_L , que peut-on

dire sur la date d'arrivée à l'aéroport ?

t_A est limitée par l'ensemble flou $T_L \ominus D$ (addition pessimiste)

- souhaitant arriver à l'aéroport à l'heure t_A avec des tolérances (pas trop tôt, ou pas trop tard) exprimées par l'intervalle flou T_A , l'heure du lever t_L devra être dans l'intervalle $T_A \ominus D$ (où \ominus désigne la soustraction optimiste).

On remarquera que lorsqu'on veut prédire à partir d'un événement donné, on doit utiliser, pour évaluer la date des événements futurs, l'arithmétique floue pessimiste qui permet de balayer toute l'étendue du possible. En revanche si on veut anticiper à partir d'un événement futur dont la date est contrainte par un intervalle flou, on est amené, en remontant le temps à contraindre de plus en plus la date d'occurrence des événements antérieurs. En particulier si la connaissance sur la durée des activités $i = 1, n$ est trop imprécise, on peut très bien trouver $T_A \ominus D = \emptyset$, ce qui signifie que compte tenu de l'imprécision sur la durée des tâches, on ne peut être sûr d'arriver à l'aéroport dans l'intervalle de temps défini par T_A . Par contraste, $T_A \ominus D$ balaye l'ensemble des heures de lever qui rendent possible l'arrivée dans l'intervalle T_A .

On peut ensuite envisager les questions suivantes:

- arrivera-t-on à l'aéroport avant \bar{t}_A si l'heure du lever est limitée par T_L ?
le degré de possibilité d'un tel événement est

$$\Pi(t_A < \bar{t}_A) = \sup_{u < \bar{t}_A} \mu_{T_L \ominus D}(u)$$

le degré de nécessité est

$$\mathcal{N}(t_A < \bar{t}_A) = \inf_{u > \bar{t}_A} 1 - \mu_{T_L \ominus D}(u)$$

- arrivera-t-on à l'aéroport avant T_A où T_A est un nombre flou ? on est amené à déterminer si T_A est plus grand que $T_L \ominus D$; on dispose de 4 indications :
(cf. chap. 4)

Possibilité d'arriver à l'heure ou avant $\Pi_{T_A}([T_L \ominus D, +\infty))$

Possibilité d'arriver avant l'heure $\Pi_{T_A}(\lceil T_L \ominus D, +\infty))$

Nécessité d'arriver à l'heure ou avant $\mathcal{N}_{T_A}^{\downarrow}([T_L \oplus D, +\infty))$

Nécessité d'arriver avant l'heure $\mathcal{N}_{T_A}^{\downarrow}(\downarrow T_L \oplus D, +\infty))$

- arrivera-t-on à l'heure exactement ?

Possibilité : $\prod_{T_L \oplus D}^{\uparrow}(T_A) = \sup_u \min(\mu_{T_L \oplus D}(u), \mu_{T_A}(u))$

Nécessité : $\mathcal{N}_{T_L \oplus D}^{\uparrow}(T_A) = \inf_u \max(1 - \mu_{T_L \oplus D}(u), \mu_{T_A}(u))$

On rappelle que l'évaluation de possibilité considère l'existence d'une intersection non vide entre $T_L \oplus D$ et T_A , et l'évaluation de nécessité considère l'inclusion de $T_L \oplus D$ dans T_A , au sens de \prec (cf. chap.2). L'inclusion au sens de Zadeh sera étudiée par l'opération

$$\inf_u \mu_{T_L \oplus D}(u) \text{ d } \int \mu_{T_A}(u)$$

qui vaut 1 si et seulement si l'inclusion tient ($\mu_{T_L \oplus D} \leq \mu_{T_A}$)

On peut poser les mêmes questions pour savoir si $T_A \succ (D)$ sera compatible avec une certaine heure de lever T_L .

Enfin on peut évaluer plus globalement ce type de situation en utilisant le concept de compatibilité (cf. chap.3), c'est-à-dire quel est le degré de vérité de la proposition "la date d'arrivée à l'aéroport sera T_A " sachant que, compte tenu de l'heure de lever et de la durée des tâches, elle sera en réalité $T_L \oplus D$. Le degré de vérité sera

$$\tau(T_A / T_L \oplus D) \stackrel{\Delta}{=} \mu_{T_A}^{\downarrow}(T_L \oplus D),$$

un ensemble flou sur $[0, 1]$. Cet ensemble flou évalue à la fois le caractère crédible de l'affirmation " t_A est T_A " (valeurs modales de τ) et l'imprécision de la connaissance (largeur de l'intervalle).

7.1.2 - CALCUL DE CHEMINS EXTREMAUX DANS UN GRAPHE IMPRECISEMENT VALUE

On considère un graphe $G = (\mathcal{N}, \mathcal{A})$ contenant n sommets formant l'ensemble \mathcal{N} et dont les arcs orientés sont contenus dans \mathcal{A} . Chaque arc (i, j) est associé à un poids $W_{i,j}$ (distance, temps...) connu imprécisément, et

défini comme un intervalle flou.

A chaque graphe de cette sorte on associe une matrice A dont les coefficients sont flous. On définit

$$\begin{aligned} A_{ij} &= W_{ij} & \text{si } (i,j) \in \mathcal{A} \\ &= -\infty & \text{si } (i,j) \notin \mathcal{A} \end{aligned} \quad (1)$$

On rappelle que l'ensemble des intervalles flous $\tilde{\mathcal{I}}(\mathbb{R} \cup \{-\infty\})$, muni des opérations internes $\tilde{\max}$ et \oplus forme un dioïde (Gondran-Minoux /20/) dont les éléments 0-réguliers sont les intervalles flous $M \leq 0$.

On définit le maximum de deux matrices à coefficients flous par

$$\tilde{\max}(A,B)_{ij} = \tilde{\max}(A_{ij}, B_{ij}) \quad (2)$$

(c'est la "somme" du dioïde).

Un produit matriciel est défini par

$$(A.B)_{ij} = \tilde{\max}_{k=1,n} A_{ik} \oplus B_{kj} \quad (3)$$

on en vérifie facilement l'associativité.

Les puissances successives de A sont alors définies par $A^0 = E$, tel que $E_{ii} = 0$, $E_{ij} = -\infty$, $i \neq j$ et

$$A^m = A^{m-1}.A = A.A^{m-1}, \quad m \geq 1 \quad (4)$$

Soit P un chemin de i à j dans G de longueur k (= contenant k arcs), et \mathcal{P}_{ij}^k l'ensemble de ces chemins. Le poids de $P = i_0 i_1 \dots i_k$, où $i_0 = i$, $i_k = j$, est :

$$W(P) = W_{i_0 i_1} \oplus W_{i_1 i_2} \oplus \dots \oplus W_{i_{k-1} i_k} \quad (5)$$

Et on voit facilement que le poids maximal des chemins de i à j ayant k arcs sera limité par le domaine de la variable x_{ij} définie par

$$x_{ij}^k = \max_{P \in \mathcal{P}_{ij}^k} x_{i_0 i_1} + x_{i_1 i_2} + \dots + x_{i_{k-1} i_k} \quad (6)$$

où $x_{i_j i_{j+1}}$ est la variable associée à l'arc (i_j, i_{j+1}) , dont le domaine de

variation est W_{ij} . Le domaine de x_{ij}^k est donc obtenu en appliquant le principe d'extension à (6).

Proposition 7.1 : le poids maximal des chemins de i à j en k arcs a pour domaine de variation A_{ij}^k .

Preuve : on sait que

$$\begin{aligned} x_{ij}^k &= \max_1 (x_{i1}^{k-1} + x_{1j}) \\ &= \max_{i_{k-1}} (\max_{P \in \mathcal{P}_{ij}^{k-1}} (x_{i o_1} + \dots + x_{i_{k-2} i_{k-1}}) + x_{i_{k-1} i_k}) \end{aligned} \quad (7)$$

(6) et (7) font intervenir les mêmes variables, associées aux arcs des chemins de i à j au plus k arcs. (6) et (7) sont isotones en fonction de ces variables, et sont donc deux représentations propres de la même fonction (cf. chap.4 § 2.6). Par récurrence, on voit donc facilement que x_{ij}^k a le même domaine de variation selon (6) ou selon (7) qui est bien A_{ij}^k .

Q.E.D.

Proposition 7.2 : Si tous les cycles de G ont un poids flou négatif ou nul, alors la quantité

$$A^* = \max_{k > 0} \widetilde{A}^k \quad (8)$$

existe et vaut :

$$A^* = \max_{k \leq n-1} \widetilde{A}^k$$

Preuve : Soit un chemin P de i à j contenant $k > n-1$ arcs. Il est formé de l'union d'un chemin P' sans cycle de $m \leq n-1$ arcs et de cycles $\gamma_1 \dots \gamma_l$ de poids négatifs ; donc

$$W(P) = W(P') \ominus \left(\sum_{i=1}^l -W(\gamma_i) \right)$$

et on a $\max(\widetilde{W}(P), W(P')) = W(P)$ (facile à voir sur les α -coupes strictes par exemple). On a donc montré qu'entre i et j

$$\forall P, \exists P' \in \mathcal{P}_{ij}^m, m \leq n-1, \tilde{\max}(W(P), W(P')) = W(P')$$

or

$$\forall k \geq n, A_{ij}^k = \tilde{\max}_{P \in \mathcal{P}_{ij}^k} W(P) \quad (\text{Proposition 1})$$

$$\begin{aligned} \text{Donc } \tilde{\max}(A_{ij}^k, \tilde{\max}_{m \leq n-1} A_{ij}^m) &= \tilde{\max}(\tilde{\max}_{P \in \mathcal{P}_{ij}^k} W(P), \tilde{\max}_{m \leq n-1} \tilde{\max}_{P' \in \mathcal{P}_{ij}^m} W(P')) \\ &= \tilde{\max} \{ \tilde{\max}(W(P), W(P')) / P \in \mathcal{P}_{ij}^k, P' \in \mathcal{P}_{ij}^m, m \leq n-1 \} \\ &= \tilde{\max} \{ W(P') / P' \in \mathcal{P}_{ij}^m, m \leq n-1 \} \end{aligned}$$

Q.E.D.

Proposition 7.3 : A^* contient comme coefficients le domaine de variation des poids maximaux des chemins de G.

Preuve : Evident d'après les propositions 7.1 et 7.2.

On voit donc que le calcul de chemins extrémaux dans un graphe imprécisément valué est relativement simple à mener, quand on se souvient que les opérations d'addition et de maximum flous sont simples à mettre en oeuvre. Les résultats détaillés ici ont été déjà suggérés dans le livre de Dubois et Prade /13/. En particulier, un algorithme tel que celui de Floyd /19/ est directement utilisable avec des valeurs floues. (Voir aussi Dubois /8/, Prade /37/).

La notion de chemin de poids maximal disparaît parce que la quantité $\tilde{\max}(M, N)$ n'est pas forcément M ni N. Néanmoins il existe nécessairement entre deux sommets quelconques i et j un sous-ensemble \mathcal{P}_{ij}^* , qui peut être très restreint, de chemins dans

$$\mathcal{P}_{ij}^* = \bigcup_{k=1, n-1} \mathcal{P}_{ij}^k,$$

tel que

$$A_{ij}^* = \tilde{\max}(W(P), P \in \mathcal{P}_{ij}^*) \quad (9)$$

$P \in \mathcal{P}_{ij}^*$ est dit chemin maximal de i à j.

Pour trouver \mathcal{P}_{ij}^* il faut mémoriser la façon dont est atteinte la valeur A_{ij}^* . Supposons qu'on utilise l'algorithme de Floyd

$$i) \quad A^{(1)} = A$$

ii) pour $k = 1, n$

pour $j = 1, n, j \neq k$

$$\text{pour } i = 1, n : A_{ij}^{(k+1)} = \max(A_{ij}^{(k)}, A_{ik}^{(k)} \oplus A_{kj}^{(k)})$$

où $A_{ij}^{(m)}$ est le résultat obtenu quand l'indice k vaut m (c'est le poids maximal des chemins de i à j de sommets intermédiaires dans $\{1, \dots, m\}$).

Soit S_{ij}^k l'ensemble des sommets l tels que

$$A_{ij}^{(k)} = \max_{l \in S_{ij}^k} W_{il} \oplus A_{lj}^{(k-1)}$$

S_{ij}^k est, dans le cas classique non flou l'ensemble des premiers sommets rencontrés sur un chemin de poids maximal de i à j , de longueur inférieure ou égale à k . Dans le cas flou, c'est le même concept pour les chemins maximaux.

On peut trouver l'ensemble S_{ij}^* des premiers sommets rencontrés sur un chemin maximal de i à j en utilisant l'algorithme de Floyd.

$$i) \quad S_{ij}^{(1)} = \{j\} \quad \text{si } (i, j) \in \mathcal{R} \\ = \emptyset \quad \text{sinon}$$

ii) pour $k = 1, n$

pour $j = 1, n ; j \neq k$

$$\begin{aligned} \text{pour } i = 1, n : S_{ij}^{(k+1)} &= S_{ij}^{(k)} \quad \text{si } A_{ij}^{(k+1)} = A_{ij}^{(k)} \\ &= S_{ik}^{(k)} \quad \text{si } A_{ij}^{(k+1)} = A_{ik}^{(k)} \oplus A_{kj}^{(k)} \\ &= S_{ij}^{(k)} \cup S_{ik}^{(k)} \quad \text{sinon - (ce troisième cas est typique} \\ &\quad \text{de données imprécises).} \end{aligned}$$

Un chemin de \mathcal{P}_{ij}^* peut donc être obtenu sous la forme $i-i_1 \dots i_{k-1}-j$ avec $i_1 \in S_{i_1-1}^*$, $\forall l = 0, k-1$

Les chemins maximaux font partie des chemins de poids maximal potentiels obtenus si les données devenaient précises. Mais on aurait tort de se restreindre uniquement à ces chemins là, car on se souvient que l'intervalle flou M peut coïncider presque avec N , tout en étant tel que $\tilde{\max}(M, N) = N$.

Soit P un chemin de i à j , de poids $W(P)$. On peut évaluer les degrés de possibilité et de nécessité pour que ce chemin soit de poids maximal, respectivement

$$\Pi_{ij}^*(P) = \sup_u \min(\mu_{A_{ij}^*}(u), \mu_{W(P)}(u)) \quad (10)$$

$$\begin{aligned} N_{ij}^*(P) &= \mathcal{N}_{W(P)}([A_{ij}^*, +\infty)) \quad (\text{cf. chap.2 } \S 4.2) \\ &= \inf_u \sup_{u > v} \max(1 - \mu_{W(P)}(u), \mu_{[A_{ij}^*, +\infty)}(v)) \end{aligned} \quad (11)$$

l'index $\Pi_{ij}^*(P)$ a été suggéré par Dubois et Prade /13, p.256/. Ces deux évaluations permettent de définir les ensembles flous de chemins possiblement et nécessairement extrémaux⁽¹⁾, respectivement. L'utilisation de l'intervalle flou $[A_{ij}^*, +\infty)$ dans (11) est justifié par le fait que l'imprécision de A_{ij}^* n'est pas liée à celle de $W(P)$.

D'autres évaluations sont possibles, en particulier :

$$E_{ij}^*(P) = \mu_{W(P)}(a_{ij}) \quad (\text{Dubois et Prade /12/}) \quad (12)$$

où a_{ij} est une valeur modale de A_{ij}^* . Ces trois indices scalaires se retrouvent sur le graphe des compatibilités $\tau(A_{ij}^*/W(P))$ ou $\tau(W(P)/A_{ij}^*)$.

NB : Si on remplace $\tilde{\max}$ par $\tilde{\min}$, et $-\infty$ par $+\infty$, on a accès aux chemins de poids minimal. Tous les résultats obtenus ici sont transposables.

Ces techniques ont été utilisées sur des problèmes d'évaluation de performance dans des réseaux de transport en commun (Dubois /6,7,8/) pour déterminer quelles étaient les rues les plus propices à la circulation des véhicules de transport, et quel devait être un tracé raisonnable de lignes de bus.

(1) au sens où leur longueur est nécessairement l'une des longueurs possibles d'un plus long chemin.

7.1.3 - APPLICATION AU PROBLEME CENTRAL DE L'ORDONNANCEMENT

Ce type d'application a été considéré par Dubois et Prade /12/, en 1978, et plus récemment par Chanas et Kamburowski /3/, /23/.

On considère un graphe orienté, sans cycles dont les sommets représentent les n tâches d'un projet, de durées floues $\tilde{p}_i \geq 0$; On définit les poids W_{ij} des arcs $(i,j) \in \mathcal{A}$ par $W_{ij} = \tilde{p}_i$. C'est la formulation de type potentiel - tâches (Roy, 1960) adoptée dans /12/. Chanas et Kamburowski /3/ supposent un graphe potentiel-étapes (PERT classique). On note, conformément à /20/ α et ω les tâches fictives de début et de fin, de durée nulle. On cherche alors, dans l'environnement imprécis dû à la connaissance approximative des durées des tâches, à trouver un ordonnancement de celles-ci qui minimise la date de fin de la tâche ω . Si on se donne une date au plus tôt de démarrage des premières tâches, soit t_α alors $\forall i \in \mathcal{N}$, la date au plus tôt T_i de début d'une tâche i sera : /12/.

$$t_i = \max_{j \in \Gamma_i^{-1}} (t_j \oplus \tilde{p}_j) \quad (13)$$

où Γ_i^{-1} est l'ensemble des tâches qui doivent précéder i .

De même si on connaît une date au plus tard T_ω de fin des tâches finales du graphe, on peut en déduire les dates au plus tard de débuts des tâches par

$$T_i = \min_{j \in \Gamma_i} (T_j \ominus \tilde{p}_j) \quad i = 1, n \quad (14)$$

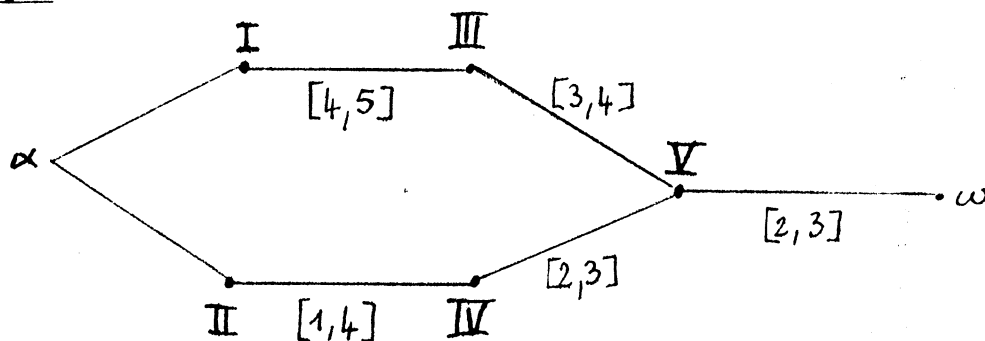
où Γ_i est l'ensemble des successeurs de i .

On peut être tenté dans (14) d'utiliser une soustraction optimiste)-(comme au paragraphe 1.1. Néanmoins le caractère non linéaire du graphe rend les estimations qu'on obtient par des formules du type

$$T_i = \min_{j \in \Gamma_i} (T_j) - (\tilde{p}_j) \quad i = 1, n \quad (15)$$

trop étroites. Par ailleurs, les estimations obtenues par (14) sont relativement plus imprécises que celles sur les dates au plus tôt.

Exemple :



Tâches	t_i	T_i (par θ)	T_i (par)-()	T_i (par θ)	T_i (par)-()
α	0	[0,3]	0	[-1,3]	?
I	0	[0,3]	0	[-1,3]	0
II	0	[2,7]	[0,2]	[0,7]	?
III	[4,5]	[5,7]	[0,5]	[3,7]	[4,5]
IV	[1,4]	[6,8]	[0,6]	[4,8]	[5,6]
V	[7,9]	[9,10]	[0,9]	[7,10]	[7,9]
ω	[9,12]	12	[0,12]	[9,12]	[9,12]

On voit que si on retarde le début de la tâche 1 jusque $t = 3$, on peut atteindre $t_\omega = 15$ ce qui viole la date au plus tard $T_\omega = 12$. Donc l'estimation $T_I = [0,3]$ obtenue par la soustraction pessimiste n'est pas une marge de sécurité.

Pour pouvoir utiliser)-(avec la date au plus tard précise 12 on a posé $T_\omega = [0,12]$ et retenu seulement les parties positives des résultats. On trouve 6 comme borne supérieure de T_{IV} . Néanmoins si les durées des tâches sont respectivement 5,4,4;2,3, il est facile de voir que si $T_\omega (=t_\omega) = 12$ alors $T_{IV} = 7$.

Donc si (13) est directement utilisable, ni (14) ni (15) ne sont entièrement satisfaisants. En particulier la définition de marges et l'étude de criticité des tâches ne peut se faire à partir des formules (14) ou (15) initialisées avec $T_\omega = t_\omega$, date de fin au plus tôt obtenue par (13).

En effet, si on utilise (14), on obtient des intervalles trop

larges dûs à des phénomènes de dépendance entre variables qui sont négligés, en particulier on a des cas où $\max(t_i, T_i) \neq T_i$, comme l'a signalé Kamburowski /23/ (tâches I,II,III dans l'exemple). En fait dans le calcul de T_i on compte deux fois l'incertitude sur la durée de la tâche i , si on pose $T_\omega = t_\omega$. Mais on a toujours $\max((-\infty, t_i], (-\infty, T_i]) = (-\infty, T_i]$.

Si on utilise (15), on est pas sûr que le calcul ait un sens car t_ω n'est pas défini pour $i = II$. Même quand T_i est calculable, l'estimation fournie est trop étroite. En témoignent les deux réalisations du graphe imprécis de l'exemple :

$$\begin{aligned} (P_I, P_{II}, P_{III}, P_{IV}, P_V) &= (4, 1, 3, 3, 2) \rightarrow T_{IV} = 4 \\ &= (5, 4, 4, 2, 3) \rightarrow T_{IV} = 7 \end{aligned}$$

alors que par (15) et $T_\omega = t_\omega$ on a trouvé $T_{IV} = [5, 6]$.

Il est clair que la criticité des tâches ne peut se définir à l'aide de l'approche traditionnelle des graphes potentiel-tâches (Roy /45/). On va la définir directement à l'aide des chemins, dont on connaît exactement la longueur, et la marge d'erreur floue.

. La durée du projet sera la longueur maximale du projet, c'est-à-dire

$$D = \max(W(P) / P \in \mathcal{P}_{\alpha, \omega}) \quad (16)$$

. Un chemin possiblement critique sera un chemin P de α à ω tel que $W(P) \cap D \neq \emptyset$. Son degré de possible criticité sera

$$C_\pi(P) = \Pi_D(W(P)) = \sup_u \min(\mu_{W(P)}(u), \mu_D(u)) \quad (17)$$

. Un chemin crédiblement critique sera un chemin P de α à ω tel que $W(P)$ sera inclus dans l'intervalle $[D, +\infty)$. Si on utilise l'inclusion faible \prec (cf. chap.2), alors le degré de nécessaire criticité sera

$$\begin{aligned} C_N(P) &= \mathcal{U}_{W(P)}([D, +\infty)) \\ &= \inf_u \sup_{u \geq v} \max(1 - \mu_{W(P)}(u), \mu_D(v)) \end{aligned} \quad (18)$$

Ces indices ont été introduits précédemment (10-11). C_N permet de discriminer parmi les chemins possiblement critiques puisque

$$C_N(P) \leq C_\pi(P)$$

En particulier si $[D, +\infty) = [W(P), +\infty)$ alors $C_N(P) \geq 0.5$. La borne inférieure est atteinte dès que les fonctions d'appartenance sont continues. Enfin si $C_N(P) = 1$, $W(P)$ est contenu dans le noyau de $[D, +\infty)$. Avec cet indice on peut considérer comme critiques les chemins tels que $C_N(P) \geq 0.5$.

Si on choisit l'inclusion de Zadeh, alors on peut utiliser une opération \bar{d} telle que $a \bar{d} b = 1 \leftrightarrow a \leq b$ (cf. chap.2) et poser

$$C_N(P) = \inf_u \mu_{W(P)}(u) \bar{d} \mu_{[D, +\infty)}(u)$$

Dans ce cas $C_N(P) = 1 \leftrightarrow W(P) \subseteq [D, +\infty)$ au sens de Zadeh.

. La criticité d'une tâche i sera définie de la même façon en se restreignant aux chemins qui passent par i .

Soit

$$W_i^- = \tilde{\max}(W(P)/P \in \mathcal{P}_{\alpha, i})$$

$$W_i^+ = \tilde{\max}(W(P)/P \in \mathcal{P}_{i, \omega})$$

alors une tâche sera dite non critique si $t_\alpha \oplus W_i^- \oplus W_i^+ \cap t_\omega = \emptyset$.

Le degré de possible criticité de la tâche i sera

$$C_\pi(i) = \Pi_{t_\omega} (t_\alpha \oplus W_i^- \oplus W_i^+) \quad (19)$$

Cette définition est similaire à celle proposée par Kamburowski /23/ dans les graphes potentiels-étapes.

Une tâche sera dite possiblement critique si $C_\pi(i) = 1$

Le degré de crédible criticité de la tâche i sera, au sens de l'inclusion

$$C_N(i) = \mathcal{N}_{t_\alpha \oplus W_i^- \oplus W_i^+}([t_\omega, +\infty)) \quad (20)$$

* Une fois de plus, on utilise l'intervalle flou $[t_\omega, +\infty)$ pour pallier au fait que t_ω peut être précis sans que $W_i^- \oplus W_i^+$ le soit.

Quand $[t_\omega, +\infty) = [t_\alpha \oplus W_1^- \oplus W_1^+, +\infty)$ alors $C_N(i) \geq 0.5$ et cela signifie que la borne inférieure sur le poids des chemins $\alpha \rightarrow i \rightarrow \omega$ sera au moins aussi grande que la plus petite durée possible du projet.

Dans l'exemple on voit que les deux chemins sont tels que $C_N(P) = 1$, et toutes les tâches sont possiblement critiques. En revanche on a

$$\begin{aligned} C_N(P) = 1 & \text{ pour } P_1 = \alpha, I, III, V, \omega & C_N(i) = 1 & \text{ pour } i \in P_1 \\ = 0 & \text{ pour } P_2 = \alpha, II, IV, V, \omega & C_N(i) = 0 & \text{ i = II, IV} \end{aligned}$$

et il est clair que les tâches II et IV auront moins tendance à être critique que les autres tâches.

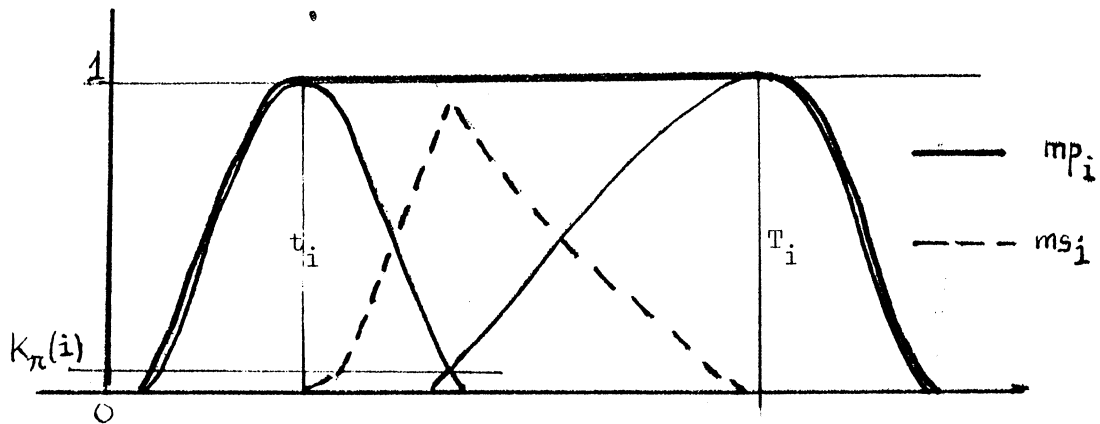
Même si (14) n'est pas très intéressant pour l'étude de la criticité, on peut quand même l'utiliser quand la date de fin au plus tard T_ω est une donnée indépendante. Elle permet de déterminer une notion de criticité des tâches non plus "en soi" mais au sens de l'intervalle limité par t_α et T_ω . Si tel est le problème on peut redéfinir la criticité d'une tâche à l'aide de (19) et (20) en substituant T_ω à t_ω . La donnée des deux contraintes t_α et T_ω permet de définir à quel point, compte tenu de l'imprécision le projet est réalisable. Utilisant la durée D du projet défini par (16), cette réalisabilité peut s'étudier en comparant les deux nombres flous $t_\alpha \oplus D$ et T_ω , à l'aide des 4 indices proposés au chapitre 4 section 4.

En particulier $\rightarrow t_\alpha \oplus D \cap [T_\omega, +\infty) = \emptyset \rightarrow$ le projet est nécessairement réalisable.

$\rightarrow t_\alpha \oplus D \cap (-\infty, T_\omega] = \emptyset \rightarrow$ le projet est nécessairement irréalisable.

Les quatre indices de comparaison permettent de nuancer entre ces cas extrêmes.

Si on utilise (13) et (14) à partir de données exogènes t_α et T_ω respectivement, on peut définir pour chaque tâche une marge, sous la forme d'un domaine temporel limité par les quantités t_i et T_i ainsi obtenues.



Ce domaine flou se réduit à la marge classique quand les données sont précises : on définit alors, à l'aide des notions introduites au chapitre 4 (§ 4.5).

- la marge sûre $ms_i =]t_i, T_i[$. C'est l'intervalle (flou) où on peut commencer la tâche i sans craindre de perturber le projet, en dépit de l'imprécision (c'est-à-dire quelles que seront les durées effectives des tâches).

- le degré de possible criticité de la tâche $K_\pi(i)$. Ce degré est nul si on est certain que la tâche ne sera pas critique, c'est-à-dire si $]t_i, T_i[\neq \emptyset$; on pose

$$K_\pi(i) = 1 - \sup_u \mu_{]t_i, T_i[}(u) = \sup_u \min(\mu_{t_i}(u), \mu_{T_i}(u)).$$

En particulier si $K_\pi(i) = 1$ cela signifie qu'il existe une configuration du projet (durées effectives) pour laquelle la tâche i sera critique.

- la marge possible $mp_i = [t_i, T_i]$; on notera que cette marque tient compte du fait que la durée des tâches est imprécise ; elle somme la marge due aux contraintes de précédence et celle engendrée par cette imprécision. C'est la plus grande marge qu'on peut espérer obtenir pour la tâche i . Par exemple on sait que si on démarre la tâche en un point du temps postérieur à l'intervalle mp_i on est certain de violer la date au plus tard T_ω . En d'autres termes $\mu_{mp_i}(t)$ mesure à quel point il est possible de commencer la tâche i au temps t pour espérer respecter les dates au plus tôt et au plus tard t_α et T_ω .

- le degré de crédible criticité de la tâche i $K_N(i)$. On le définira comme un degré d'inclusion de $(-\infty, T_i]$ dans l'intervalle $(-\infty, t_i]$

soit

$$K_N(i) = \inf_u \max(1 - \mu_{(-\infty, T_i]}(u), \mu_{(-\infty, t_i]}(u))$$

$$= \mathcal{N}_{(-\infty, T_i]}^{(-\infty, t_i]} \quad (\text{cf. chap. 4 section 4})$$

On vérifie que $K_\pi(i) < 1 \rightarrow K_N(i) = 0$ (la criticité doit être possible avant d'être crédible)

$$T_i = t_i \rightarrow K_N(i) \geq 0.5$$

(la limite inférieure est atteinte si μ_{t_i} et μ_{T_i} sont continues).

$$K_N(i) = 1 \leftrightarrow T_i \text{ est contenu dans le noyau de } (-\infty, t_i]$$

On peut considérer comme critiques toutes les tâches telles que $K_N(i) \geq .5$, en particulier dès que $T_i = t_i$ (égalité d'intervalles flous).

On rappelle que K_N et K_π sont des indices de criticité relatifs à des dates t_α et T_ω indépendantes, contrairement à C_π et C_N qui sous-entendent $T_\omega = t_\omega$ et sont des indices de criticité en soi.

Dans le petit exemple étudié plus haut on obtient pour $T_\omega = 12$,

$$t_\alpha = 0$$

Tâches	ms_i	mp_i	$K_\pi(i)$	$K_N(i)$
α	\emptyset	$[0, 3]$	1	0
I	\emptyset	$[0, 3]$	1	0
II	$[0, 2]$	$[0, 7]$	0	0
III	\emptyset	$[4, 7]$	1	0
IV	$[4, 6]$	$[1, 8]$	0	0
V	\emptyset	$[7, 10]$	1	0
ω	\emptyset	$[9, 12]$	1	1

Donc avec $T_\omega = 12$ les tâches II et IV ne seront jamais critiques. Aucune tâche ne sera nécessairement critique. De plus $t_\alpha \oplus D = [9, 12] \subseteq (-\infty, T_\omega]$. Donc on sait que malgré l'imprécision, le projet sera toujours réalisable.

Fixer $T_{\omega} = 10$ détruirait toutes les marges dures, réduirait les marges possibles, et rendrait les tâches III et V critiques avant les autres, c'est-à-dire qu'en agissant sur elles on peut se prémunir contre une violation de la date au plus tard de fin du projet. C'est le sens de la "crédible criticité" introduite ici.

Prade /33,34,35/ a utilisé ce type de méthode sur des problèmes d'emploi du temps.

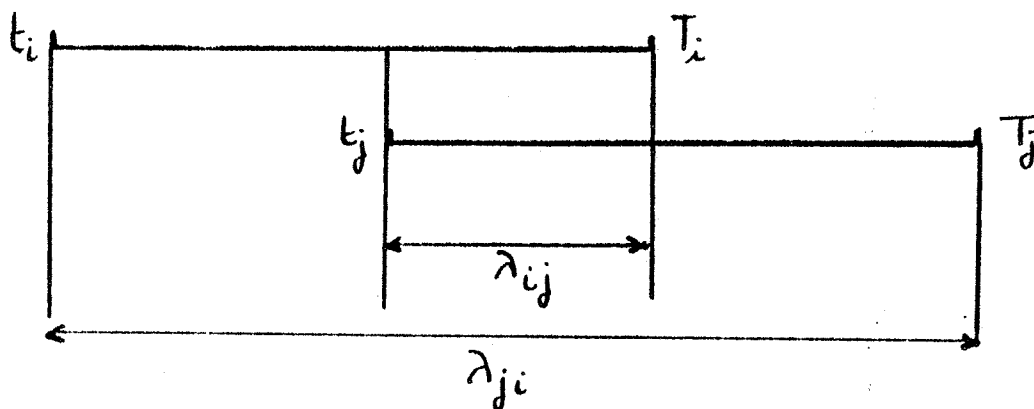
7.1.4 - APPLICATION A DES PROBLEMES D'ORDONNANCEMENT A CONTRAINTES CUMULATIVES

On suppose que l'exécution de chaque tâche nécessite l'utilisation d'une ressource d'un type donné (par exemple, une machine) disponible en un seul exemplaire. On note \mathcal{M} l'ensemble des machines, et m_i la machine qui doit accomplir la tâche i . On suppose qu'une tâche n'est pas interruptible et qu'une machine ne peut effectuer qu'une tâche à la fois.

En plus des contraintes de précédence entre les tâches, on a donc des contraintes disjonctives qu'il faut arbitrer en déterminant un séquençement des tâches affectées à chaque machine, afin de minimiser la durée d'exécution de l'ensemble des tâches. Dans le cas où les durées d'exécution des tâches sont connues, Ercshler et al /17/ et Prade /40/ ont suggéré une méthode pour induire, à partir des contraintes de précédence et des contraintes temporelles, des contraintes d'ordre entre les tâches relatives à chaque machine. Si on calcule pour chaque tâche la marge induite par les contraintes globales de réalisation (dates au plus tard et au plus tôt T_{ω} et t_{α}) il est possible de décider qu'une tâche passera avant l'autre sur telle machine en examinant la position relative de ces marges. Cette analyse du chevauchement des marges peut être étendue au cas où les durées d'exécution sont floues (cf. Prade /33/ et Dubois-Prade /15/).

Rappelons d'abord la méthode d'analyse des chevauchements. Si $[t_i, T_i]$ est la marge, déduite des contraintes de précédence, sur la tâche i pour $i \in \mathcal{J}(m)$, où $\mathcal{J}(m)$ est l'ensemble des tâches affectées à la machine m , on caractérise la position relative de $[t_i, T_i]$ et $[t_j, T_j]$ par λ_{ij} , λ_{ji} et σ_{ij} , définis pour $t_i \leq t_j$ par :

$$\lambda_{ij} = T_i - t_j ; \quad \lambda_{ji} = T_j - t_i ; \quad \sigma_{ij} = p_i + p_j$$



Il est clair que si $\lambda_{ij} \leq 0$ alors i précède j sur m .

Si $\lambda_{ij} > 0$ plusieurs cas peuvent se présenter.

i) $\sigma_{ij} > \max(\lambda_{ij}, \lambda_{ji})$. Alors l'intervalle de temps $[t_i, \max(T_i, T_j)]$ alloué pour la réalisation successive des 2 tâches est trop court. Il n'y a pas d'ordonnement réalisable.

ii) $\lambda_{ji} \geq \sigma_{ij} > \lambda_{ij}$. Dans ce cas la tâche i doit précéder la tâche j et on peut modifier les marges :

$$\begin{aligned} T_i &\text{ devient } \min(T_i, T_j - p_j) \\ t_j &\text{ devient } \max(t_i + p_i, t_j) \end{aligned}$$

iii) si $\lambda_{ji} < \sigma_{ij} \leq \lambda_{ij}$ c'est le cas dual de ii). j doit précéder i et

$$\begin{aligned} T_j &\text{ devient } \min(T_j, T_i - p_i) \\ t_i &\text{ devient } \max(t_j + p_j, t_i) \end{aligned}$$

Sinon on n'a pas de nouvelle information.

La procédure de détermination de séquences partielles de tâches sur les machines est donc basée sur le principe suivant : on utilise la procédure ci-dessus sur toutes les paires $(i, j) \in \mathcal{J}(m)$, $\forall m$; si les marges de i et j ont été modifiées, on rajoute dans l'ensemble des paires à examiner,

les paires (i,k) et (j,l) déjà examinées auparavant. Quand aucune modification n'est possible, la procédure s'arrête. Quand t_i et T_i sont des intervalles flous, la procédure définie ci-dessus peut être adaptée. On pose maintenant

$$\lambda_{ij} = T_i \ominus t_j \quad \lambda_{ji} = T_j \ominus t_i \quad \sigma_{ij} = p_i \oplus p_j$$

Les indices de comparaison définis au chapitre 4 section 4 vont à nouveau être utiles pour interpréter les tests i), ii) et iii). On est en effet amené

- . à comparer σ_{ij} et $\tilde{\max}(\lambda_{ij}, \lambda_{ji})$; ce qu'on sait déjà faire.
- . à placer σ_{ij} entre λ_{ji} et λ_{ij}

Ces tests étant complémentaires on va chercher des relations d'inclusion entre σ_{ij} et $\tilde{\max}(\lambda_{ij}, \lambda_{ji}), +\infty)$, $\lambda_{ji}, \lambda_{ij}[$, $\lambda_{ij}, \lambda_{ji}[$ d'une part ou si cela échoue, définir si σ_{ij} est inclus dans $\tilde{\max}(\lambda_{ij}, \lambda_{ji}), +\infty)$, $[\lambda_{ji}, \lambda_{ij}]$ et $[\lambda_{ji}, \lambda_{ij}]$. Ces intervalles sont définis conformément à la section 4 du chapitre 4. On examine ces indices dans l'ordre décroissant de sévérité :

A - d'abord

$$\mathcal{N}_{\sigma_{ij}}(\tilde{\max}(\lambda_{ij}, \lambda_{ji}), +\infty), \mathcal{N}_{\sigma_{ij}}(\lambda_{ji}, \lambda_{ij}[), \mathcal{N}_{\sigma_{ij}}(\lambda_{ij}, \lambda_{ji}[).$$

On vérifie qu'au plus l'un d'entre eux est strictement positif, car au plus un des ensembles $\lambda_{ji}, \lambda_{ij}[$, $\lambda_{ij}, \lambda_{ji}[$ est non vide, et son support est disjoint de celui du troisième ensemble $\tilde{\max}(\lambda_{ij}, \lambda_{ji}), +\infty)$.

B - Si tous sont nuls on examine

$$\mathcal{N}_{\sigma_{ij}}(\tilde{\max}(\lambda_{ij}, \lambda_{ji}), +\infty), \mathcal{N}_{\sigma_{ij}}([\lambda_{ji}, \lambda_{ij}]) \\ \mathcal{N}_{\sigma_{ij}}([\lambda_{ij}, \lambda_{ji}]) \quad \text{en vue de parvenir à discriminer ;}$$

on remarquera que les indices deux de ceux utilisés ci-dessus (les valeurs de possibilité) contiennent une information redondante, si on admet que dès que

$$\Pi_{\alpha_{ij}}((-\infty, \tilde{\min}(\lambda_{ij}, \lambda_{ji})]) = 1$$

on ne peut plus améliorer les marges sur les deux tâches i et j.

On remarquera que si, à l'issue des tests on se retrouve dans le cas i)

- si la conclusion est fournie par le test A, on est certain qu'il n'y a pas d'ordonnancement réalisable, quelles que soient les durées effectives des tâches.

- si c'est le test B qui donne cette conclusion, il est possible que l'on viole les contraintes de date au plus tard, si les tâches ont des durées effectives pessimistes - situation que l'on peut juger inacceptable.

Dans les cas ii) et iii) on modifie les dates au plus tôt et au plus tard avec les opérations étendues $\overset{\sim}{\min}$, $\overset{\sim}{\max}$, Θ , θ .

Dans Prade /40/, Dubois-Prade /15/, l'analyse floue des chevauchements est effectuée avec une partie seulement de l'information disponible : à savoir les indices

$$\Pi_{\sigma_{ij}}(\overset{\sim}{[\max(\lambda_{ij}, \lambda_{ji}), +\infty)})$$

$$\Pi_{\sigma_{ij}}([\lambda_{ij}, \lambda_{ji}])$$

$$\Pi_{\sigma_{ij}}([\lambda_{ij}, \lambda_{ji}])$$

Il est clair que lorsque deux de ces indices valent 1, on ne sait pas conclure. En particulier le test A ne donne rien. En revanche il se peut très bien que le test B permette de discriminer. Des exemples d'utilisation sont fournis dans /33,15/ qui montrent qu'en environnement imprécis, on peut encore définir des séquences partielles de tâches sur chaque machine. L'emploi du test B permettrait d'améliorer encore les résultats obtenus.

7.1.5 - COMPARAISON AVEC LES APPROCHES PROBABILISTES

Si on considère le problème du graphe PERT dans son énoncé probabiliste, où les x_{ij} sont des variables aléatoires définies par une distribution, on peut mesurer la différence de difficulté dans les calculs à mener.

- le poids d'un chemin est obtenu par convolution de variables aléatoires plus difficile à mettre en oeuvre que l'addition d'intervalles flous. Néanmoins,

si le chemin est long, on peut invoquer la tendance vers la loi normale.

- le poids maximal des chemins est très complexe à calculer en raison de la dépendance des chemins non disjoints.

En pratique on est amené

- soit à supposer que le chemin critique est unique (c'est le chemin "critique en moyenne"), domine toujours les autres chemins, et à lui appliquer le théorème central-limite (Malcolm et al /28/). On sera en général optimiste quant à l'évaluation de la date de fin du projet.

- soit à simuler le graphe avec une méthode (lourde) de type Monte-Carlo.

- soit à supposer que les poids des chemins sont tous indépendants ; on peut alors utiliser la formule (13) dans sa version probabiliste (Sculli /40/) ou évaluer par convolution la longueur de tous les chemins (Ventura /48/). Les résultats seront très douteux avec un graphe fortement maillé.

Vois Sculli /40/ et Sigal, Pritsker, Solberg /42/ pour une bibliographie sur le PERT Stochastique. Ces derniers proposent des indices de criticité des chemins dont le calcul est loin d'être trivial.

Un autre problème avec le PERT stochastique concerne l'hypothèse d'indépendance entre les durées des tâches. Elle peut être mise en doute si par exemple elles seront accomplies successivement par les mêmes personnes. Enfin le PERT stochastique perd complètement sa signification s'il s'agit d'évaluer un projet de type nouveau, et qui ne sera pas (ou peu) recommencé. Les estimations des durées des tâches, fournies par des experts, pourront difficilement être assimilées à des distributions de probabilité, stricto sensu.

En revanche l'analyse d'un graphe potentiel-tâche par l'arithmétique floue peut être très utile justement là où l'approche probabiliste est jugée mal adaptée spécialement quand elle s'avère trop lourde (simulation), trop optimiste (méthodes approchées) ou peu réaliste (absence de statistiques).

En effet

- Les calculs nécessités par l'approche floue sont relativement légers. Ce sera d'autant plus vrai, si les durées des tâches peuvent être modélisées à l'aide d'intervalles flous de type L-R.
- La sémantique des intervalles flous convient tout à fait à la prise en compte de durées de tâches connues approximativement par l'intermédiaire d'un expert. Mais on peut aussi construire des intervalles flous à partir de statistiques (cf. Chap.1, section 6).

Les résultats obtenus par l'approche floue seront moins précis que ceux qu'on peut attendre d'une investigation par les probabilités. Que l'on considère par exemple le cas particulier de distributions uniformes, qui se ramène dans un cas à de l'analyse de sensibilité ; mais la distribution de probabilité de la date de fin aura un écart-type qui sera plus petit que la demi-largeur de l'intervalle contenant la date de fin réelle. Dans certains contextes néanmoins, l'attitude prudente que traduit l'approche floue est celle qui se justifie le plus.

7.2 - PROBLÈMES D'OPTIMISATION EN ENVIRONNEMENT IMPRÉCIS

Cette section fait écho à la section 1 du chapitre 6, et traite comme elle du choix d'une décision optimale en présence de contraintes et de critères définis de façon imprécise. Mais dans le chapitre 6 on supposait que le nombre des décisions possibles était limité et repérable, de sorte qu'il était envisageable de les évaluer une par une. Ici on ne suppose plus cette hypothèse et la recherche de la meilleure décision passera par l'utilisation de techniques d'optimisation.

On considère ici principalement le cas de la programmation linéaire et de diverses extensions de ce modèle en environnement imprécis. Le point essentiel consiste à montrer comment, à partir de contraintes et d'objectifs flous, on peut se ramener à une formulation équivalente qui soit de nouveau un problème d'optimisation traditionnel, autant que possible linéaire. Le problème ainsi obtenu, bien que déjà connu sur le plan mathématique, tiendra compte de l'imprécision des données originelles, et sera donc

plus riche de sens que celui qu'on poserait directement en supposant ces données connues précisément. Le dernier paragraphe décrit un problème réel d'optimisation de paramètres de réglage sur une chaîne d'usinage, traité antérieurement par l'auteur.

7.2.1 - FORMULATION CLASSIQUE D'UN PROBLEME D'OPTIMISATION EN ENVIRONNEMENT IMPRECIS

Soit X un ensemble d'alternatives qui contient la solution d'un problème d'optimisation multicritère. On suppose $X = \mathbb{R}^n$, et $\forall x \in \mathbb{R}^n$ x est un vecteur décrivant les caractéristiques d'une décision. Selon Bellman et Zadeh /1/, contraintes et critères peuvent être vus comme des ensembles flous sur X (voir chap.6, section 1). En effet une contrainte définit un sous-ensemble admissible de X dont les frontières ne sont pas forcément bien définies. Un critère vu comme une fonction d'évaluation de x , $\forall x \in \mathbb{R}^n$, est, à une normalisation près, une fonction d'appartenance d'ensemble flou, à condition d'y adjoindre des seuils d'indifférence, afin de définir au sens du chapitre 6 section 1, un objectif. Ici on examine de plus près la modélisation d'une contrainte floue.

En général les contraintes auront une expression analytique sous la forme :

$$f(x,a) = g(x,b) \quad (21)$$

ou

$$f(x,a) \leq g(x,b) \quad (22)$$

Dans les expressions (21) et (22), a et b sont des vecteurs de paramètres dont les valeurs sont des données. Le caractère flou de la contrainte proviendra d'une tolérance admise sur l'égalité et l'inégalité et d'une spécification imprécise des paramètres a et b .

Si les valeurs a_i et b_j des paramètres sont des quantités floues alors $f(x,a)$ et $g(x,b)$ seront, pour x fixé des quantités floues. On notera qu'en environnement imprécis, on ne peut changer de membre les termes d'une égalité sous peine d'en modifier le sens. C'est pourquoi, dans (21-22) on a fait apparaître deux fonctions f , et g qui doivent posséder une interpré-

tation claire dans le contexte du problème.

La contrainte égalité, pourra être exprimée de façon plus ou moins rigide :

- égalité stricte des quantités floues, que l'on peut voir comme l'égalité des fonctions d'appartenance $\mu_{f(x,a)}$ et $\mu_{g(x,b)}$
- inclusion de l'une des quantités dans l'autre, au sens de Zadeh.
 $f(x,a) \subseteq g(x,b)$ par exemple.
- égalité au sens faible, par exemple au sens de la théorie des possibilités c'est-à-dire construite à partir d'un indice de possibilité d'égalité (intersection $f(x,a) \cap g(x,b)$) ou de nécessité d'égalité (inclusion au sens de \prec)

La première interprétation est beaucoup trop rigide pour être acceptable, en général, parce que

- les fonctions d'appartenance ne sont pas connues précisément
- on risque d'aboutir très souvent à des problèmes sans solution en multipliant ce type de contraintes.

Dans l'interprétation de type inclusion au sens de Zadeh, si le risque de problème impossible est encore présent, on exprime que l'incertitude sur l'un des membres de l'égalité doit être limité par celle de l'autre membre. Ce type de contrainte sur l'incertitude se justifie dans certains contextes, et possède en tout cas une résonance intuitive.

Le plus souvent, par souci de réalisabilité, et en accord avec l'existence de tolérances, c'est l'égalité au sens faible qui sera le plus souvent adoptée pour exprimer la contrainte floue.

L'inégalité (22) peut s'entendre au sens d'une inclusion (stricte ou faible) de $g(x,b)$ dans $[f(x,a), +\infty)$ ou $]f(x,a), +\infty)$ intervalles flous définis au chapitre 4.

Dans tous les cas on peut, à partir de la contrainte floue se

ramener à une ou plusieurs contraintes classiques équivalentes. Par exemple
 . inclusion au sens de Zadeh

$$(21) \leftrightarrow \mu_{f(x,a)} \leq \mu_{g(x,b)}$$

. possibilité ou nécessité d'égalité

$$(21) \leftrightarrow \Pi_{f(x,a)}(g(x,b)) \geq \pi$$

ou

$$\min(\mathcal{N}_{f(x,a)}(g(x,b)), \mathcal{N}_{g(x,b)}(f(x,a))) \geq \nu$$

où π et ν sont des seuils que l'on explicitera ou non.

On construit donc des évaluations du degré de satisfaction de la contrainte, qui définissent sur X un domaine flou d'admissibilité.

Si on doit respecter plusieurs contraintes $C_1 \dots C_p$ simultanément, les C_i désignant ces domaines flous admissibles, on devra chercher la solution dans l'intersection $C_1 \wedge \dots \wedge C_p$ des domaines flous. Voir Tanaka, Okuda, Asai /46/ et Flachs-Pollatshek /18/ pour l'étude générale de ce type de problème.

Si on parvient pour chaque critère à définir des objectifs $G_1 \dots G_q$ au sens du chapitre 6 section 1, le problème d'optimisation en environnement imprécis se ramènera à celui de trouver x qui maximise le degré d'appartenance au domaine flou défini par l'agrégation du domaine flou d'admissibilité ($C_1 \wedge \dots \wedge C_p$) et de l'objectif ($G = G_1 * G_2 \dots * G_q$).

Dans la suite on examine divers cas où le passage de la formulation imprécise à la formulation classique équivalente est simple, et aboutit à un problème d'optimisation résoluble par des méthodes connues.

7.2.2 - PROGRAMMATION LINEAIRE AVEC CONTRAINTES SOUPLES (Zimmermann /52/)

Considérons le problème d'optimisation linéaire à m variables

$$x_1 \dots x_m$$

minimiser $Z = cx$

avec les n contraintes $Ax \leq b, x \geq 0$ (23)

où b, c sont des vecteurs de coefficients, A une matrice réelle. Lorsque l'on est amené à formuler un tel modèle, on n'est pas certain qu'une solution du problème existe, c'est-à-dire si le domaine défini par $Ax \leq b$ est non vide. En cas de problème impossible, on peut modifier la valeur de b de manière à relaxer les contraintes et obtenir une solution réalisable. Néanmoins on ne sait pas a priori de combien il faut modifier ces coefficients pour qu'une solution existe.

En pratique, les coefficients b_i ont une interprétation physique, et celui qui a formulé le problème doit posséder une certaine connaissance sur les valeurs préférables et limites des b_i . Plutôt que de procéder par tâtonnements, on peut substituer aux valeurs des b_i un seuil flou (cf. chap.3, § 7.2) et à (23) substituer : une contrainte souple.

$$\forall i = 1, n \quad A_i x \lesssim b_i, \quad x \geq 0 \quad (24)$$

où \lesssim exprime une certaine tolérance par rapport aux valeurs préférables b_i du seuil. Cette tolérance peut se traduire par des seuils flous (cf. chap.3) B_i tels que

$$\forall u \leq b_i : \mu_{B_i}(u) = 1 \quad (25)$$

$$\exists d_i \quad \forall u \geq b_i + d_i : \mu_{B_i}(u) = 0 \quad (26)$$

$b_i + d_i$ est la limite supérieure admissible pour le coefficient $n^o i$. Sur $[b_i, b_i + d_i]$, μ_{B_i} est décroissant. L'ensemble C_i des points de \mathbb{R}^m satisfaisant la contrainte $n^o i$ a pour fonction caractéristique généralisée

$$\mu_{C_i}(x) = \mu_{B_i}(A_i x) \quad (27)$$

Le domaine admissible flou est obtenu en agrégeant les contraintes floues. Cette agrégation est du type intersection puisque l'on a affaire à des contraintes entre lesquelles aucune compensation n'est envisageable, en général. Donc le domaine flou admissible est C tel que

$$\mu_C(x) = \bigcap_{i=1, n} \mu_{C_i}(x) \quad (28)$$

On notera qu'une contrainte souple de type égalité peut s'exprimer comme deux contraintes souples de type inégalité, puisque

$$A_i x \sim b_i$$

s'exprime à l'aide de B_i tel que $\mu_{B_i}(b_i) = 1$ et

$$\exists d_i^+ , \forall u \geq b_i + d_i^+ , \mu_{B_i}(u) = 0 \quad (29)$$

$$\exists d_i^- , \forall u \leq b_i - d_i^- , \mu_{B_i}(u) = 0 \quad (30)$$

Ces deux contraintes utilisent $[B_i, +\infty)$ et $(-\infty, B_i]$ comme seuils flous.

Zimmermann /52/ a remarqué que le critère peut se mettre sous la forme d'une contrainte souple de type

$$c x \lesssim z_0$$

où z_0 est par exemple la solution de (23) en substituant $b_i + d_i^+$ à b_i (la meilleure valeur accessible du critère), et le seuil flou B_0 est défini comme en (25), (26) avec $b_0 = z_0$. Zimmermann /52/ a proposé de remplacer (23) par :

$$\text{maximiser } \mu_D(x) = \underset{i=0,n}{\perp} \mu_{C_i}(x) \quad (31)$$

avec \perp = minimum ou produit, et des nombres flous à fonction d'appartenance linéaire pour les B_i . La contrainte d_0 peut se calculer avec la valeur du critère au point le plus à l'intérieur du domaine C , défini en (28), soit

$$z_0 + d_0 = c \bar{x} \quad \text{avec} \quad \mu_C(\bar{x}) = \sup \mu_C$$

Dans le cas où \perp = minimum, (31) est équivalent à

$$\begin{aligned} &\text{maximiser } \lambda \\ &\text{sous les contraintes } \lambda \leq \mu_{C_i}(x) \quad i = 0, n \\ & \quad \quad \quad x \geq 0 \end{aligned} \quad (32)$$

Il s'agit alors d'un modèle analogue à celui où la satisfaction des contraintes et du critère seraient évalués par des fonctions d'utilité auxquelles on appliquerait les règles de la théorie des jeux, comme l'a fait remarquer Dyson /16/. Zimmermann /52/ suppose que les B_i ont des fonctions d'appartenance linéaire, et ramène alors (32) à un programme linéaire. Leberling /25/

suppose les μ_{B_i} représentées à partir de fonctions hyperboliques et montre qu'on se ramène encore à un programme linéaire. En fait l'équivalence de (32) avec un programme linéaire est vraie dès qu'on choisit pour les B_i des représentations de type L-L avec (cf. chap.4).

$$B_i = (b_i, -\infty, d_i)_{LL} \quad \text{où} \quad L(1) = 0 ; L(0) = 1$$

alors

$$\lambda \leq \mu_{C_i}(x) \leftrightarrow d_i L^{-1}(\lambda) + b_i \geq A_i x$$

Et en posant $x_{m+1} = L^{-1}(\lambda)$ (32) équivaut à

minimiser x_{m+1}

$$\text{sous les contraintes } A_i x - d_i x_{m+1} \leq b_i ; i = 0, n \quad (33)$$

$$0 \leq x_{m+1} \leq 1 ; x \geq 0$$

Donc le résultat de (32) est indépendant de la forme des fonctions d'appartenance dès que les μ_{B_i} se déduisent les uns des autres par affinité et translation. On a aussi montré par là que le modèle proposé par Zimmermann /52/ est équivalent dans de nombreux cas à un paramétrage du second membre (voir aussi Hamacher et al /21/ pour les liens avec l'analyse de sensibilité et la dualité). $L^{-1}(x_{m+1})$ mesure le degré d'admissibilité de la solution obtenue. Oh'Eigartaigh /30/ applique cette approche au problème de transport et caractérise la solution du problème à contrainte souple en fonction de celle du problème original.

7.2.3 - APPROCHE FLOUE DES PROBLEMES LINEAIRES MULTICRITERES

La possibilité de transformer le critère en contrainte souple rend néanmoins cette approche utilisable pour les problèmes multicritères linéaires (dits du vecteur-maximum) (Zimmermann /52/). Néanmoins, dans ce cas, l'opération minimum ne suffit plus pour agréger les contraintes souples. Luhandjula /26/ a proposé une opération paramétrée qui est une combinaison convexe entre le minimum et la somme bornée, ce qui lui permet de ramener le problème (28) à de la programmation linéaire.

L'approche par contrainte souple est analogue à la méthode

dite "du point de mire" (Goal Programming /5/), qui définit pour les critères des valeurs à atteindre, et cherche à minimiser une distance entre le vecteur caractérisant une solution admissible et le point de mire ainsi défini. Il y a néanmoins quelques différences entre la méthode du point de mire et l'approche par contraintes souples : (voir Kabbara /22/ pour une discussion plus détaillée).

- La première donne toujours une solution, qui ne sera pas toujours utilisable puisque l'on ne fait exprimer dans le modèle que les valeurs préférables du critère. La seconde approche prend explicitement en compte les tolérances admises, et ne fournit donc pas de solution à tous coups.

- La distance entre le point de mire et une solution courante est construite à l'aide de pondérations qui sont difficiles à interpréter quantitativement. En revanche dans la méthode floue, les d_j jouent le rôle de pondération et s'interprètent en termes de zone admissible.

La solution obtenue par l'approche floue n'est pas forcément non dominée au sens de Pareto (= efficace) mais on a le résultat suivant.

Proposition 7.4 : Parmi les solutions optimales obtenues par agrégation d'objectif flous dans un problème de programmation linéaire multicritère, il en est toujours une qui est maximale au sens de Pareto, dès que de telles solutions existent.

Preuve : soit Z_1, \dots, Z_p les p critères à minimiser, et $Ax \leq b, x \geq 0$ les contraintes du problème. On suppose $Z_j = C^j x$. Les objectifs flous sont G_j tels que

$$\forall j \quad \mu_{G_j}(x) = \mu_{B_j}(C^j x)$$

où B_j est défini par

$$\begin{aligned} \mu_{B_j}(u) &= 1 & \forall u \leq Z_j^0 \\ &> 0 & \forall u \in [Z_j^0, Z_j^0 + d_j^0[\quad (\text{et décroissant}) \\ &= 0 & \forall u \geq Z_j^0 + d_j^0 \end{aligned}$$

Soit h un opérateur d'agrégation conforme aux axiomes définis au chapitre 6,

section 1 et tel que l'ensemble flou des bonnes décisions soit

$$\mu_D(x) = h(\mu_{G_1}(x), \mu_{G_2}(x), \dots, \mu_{G_p}(x)) \quad \forall x : Ax \leq b, x \geq 0$$

h est non décroissant en fonction de chaque argument, $h(1,1,1,\dots,1) = 1$, $h(0,0,0,\dots,0) = 0$. Soit \bar{x} admissible tel que $\mu_D(x)$ soit maximum. Supposons \bar{x} non efficace. Alors $\exists x^*$ efficace tel que

$$C_j x^* \geq C_j \bar{x} \quad \forall j = 1, p \quad \text{et} \quad \exists k, C_k x^* > C_k \bar{x}$$

et de par les propriétés de h ,

$$\mu_D(x^*) = \mu_D(\bar{x})$$

Q.E.D.

La proposition 7.4 indique que pour trouver une solution efficace il suffit d'explorer les solutions optimales du problème dont le critère est $\mu_D(x)$. Ce résultat a été montré par plusieurs auteurs dans des cas particuliers (Hannan /22/ pour $h = \text{produit}$, Leberling /25/ pour $h = \text{min}$, Luhanjula /26/ pour son opération paramétrée). Buckley /2/ a récemment montré qu'avec $h = \text{minimum}$, on pouvait, en modulant les coefficients d_i générer toutes les solutions efficaces (résultat déjà connu pour les combinaisons convexes de Z_j). Enfin on notera que le caractère linéaire, ou même différentiable des Z_j n'intervient pas dans la preuve de la proposition 7.4. Un exemple d'application de cette approche est décrit dans Wiedey et Zimmermann /50/.

7.2.4 - CONTRAINTES LINEAIRES FLOUES :

APPROCHE PAR L'INCLUSION DE ZADEH

On suppose maintenant que les coefficients de la matrice A des contraintes sont mal connus, tout comme les coefficients b_j des seconds membres. Ce type de problème a été étudié par Soyster /43/ sous la forme

$$\forall i = 1, n \quad \sum_{j=1}^m A_{ij} x_j \leq B_i \quad (34)$$

où les A_{ij} et les B_i sont des intervalles fermés et Σ est l'addition d'intervalles. En fait la forme générale de telles contraintes est plutôt, conformément à (21)

$$\forall i = 1, n \quad \sum_{j=1}^m A_{ij} x_j \oplus A_{i,m+1} \subset \sum_{j=1}^m B_{ij} x_j \oplus B_{i,m+1} \quad (35)$$

qui n'est pas équivalent à (34). En notant $\bar{a}_{ij} = \sup A_{ij}$, $\underline{a}_{ij} = \inf A_{ij}$, etc...
 (35) équivaut à (si $x > 0$) :

$$\forall i = 1, n \quad \sum_{j=1}^m (\underline{a}_{ij} - \underline{b}_{ij}) x_j + \underline{a}_{i,m+1} - \underline{b}_{i,m+1} \geq 0 \quad (36)$$

$$\sum_{j=1}^m (\bar{a}_{ij} - \bar{b}_{ij}) x_j + \bar{a}_{i,m+1} - \bar{b}_{i,m+1} \leq 0 \quad (37)$$

NB : Si on change de membre les termes dans (35) pour revenir à la forme (34) alors cela revient à changer \underline{b}_{ij} et $\underline{a}_{i,m+1}$ en \bar{b}_{ij} et $\bar{a}_{i,m+1}$ en (36) et inversement en (37).

Negoita /29/ a étendu les égalités de type (34) en égalités floues en substituant des intervalles flous aux intervalles. En utilisant la commutativité avec les α -coupes fortes, il a pu montrer que, si dans (34) les A_{ij} et B_i sont des intervalles flous alors (34) est équivalent aux contraintes:

$$\forall \alpha \in]0, 1] \quad \sum_{j=1}^m (A_{ij})_{\alpha}^{-} x_j \leq (B_i)_{\alpha}^{-} \quad (38)$$

qui sont du type étudié par Soyster /50/. On se ramène à un nombre fini de contraintes par une discrétisation finie de $]0, 1]$. Néanmoins une contrainte floue peut être équivalente à de nombreuses contraintes linéaires, ce qui limite l'applicabilité de cette approche.

Pour en réduire la complexité, Dubois et Prade /14/ ont utilisé des nombres flous de type L-R pour représenter les coefficients flous.

Si

$$A_{ij} = (a_{ij}, \alpha_{ij}, \bar{a}_{ij})_{LR}$$

et

$$B_i = (b_i, \beta_i, \bar{b}_i)_{LR}$$

alors (34) est équivalent à : $(x_j > 0 \forall j)$

$$\sum_{j=1}^m a_{ij} x_j = b_i \quad (39)$$

$$\sum_{j=1}^m \alpha_{ij} x_j \leq \beta_i \quad ; \quad \sum_{j=1}^m \bar{a}_{ij} x_j \leq \bar{\beta}_i \quad (40)$$

Mais (39)-(40) peuvent entraîner l'absence de solutions admissibles pour (34). On est donc amené à éclater a_{ij} et b_i en deux valeurs $(\underline{a}_{ij}, \bar{a}_{ij})$ et $(\underline{b}_i, \bar{b}_i)$. Dans ce cas (34) équivaut à

$$\sum_{j=1}^m \underline{a}_{ij} x_j \geq \underline{b}_i \quad ; \quad \sum_{j=1}^m \bar{a}_{ij} x_j \leq \bar{b}_i \quad (41)$$

$$\sum_{j=1}^m (\underline{a}_{ij} - \alpha_{ij}) x_j \geq \underline{b}_i - \beta_i \quad ; \quad \sum_{j=1}^m (\bar{a}_{ij} + \bar{\alpha}_{ij}) x_j \leq \bar{b}_i + \bar{\beta}_i \quad (42)$$

(41) et (42) expriment respectivement l'inclusion des noyaux et des supports. Quand les coefficients imprécis sont modélisés par des nombres flous de type L-R cette double inclusion équivaut à celle des intervalles flous (34). On retrouve dans ce cas l'invariance de l'ensemble des solutions admissibles de (34) à la forme des fonctions d'appartenance. Seule change le degré d'admissibilité des x satisfaisant (42). Il est clair que tout ce qui vient d'être dit se transpose trivialement à la forme générale d'égalité (35).

On peut traiter similairement les contraintes inégalités.

Dans ce cas B_i est un seuil flou au sens du chap.3 § 7.2, c'est-à-dire limite une contrainte souple comme au paragraphe 2.2. Par exemple $B_i = (b_i, -\infty, d_i)$ dans (34). Si les A_{ij} sont des intervalles et B_i est non flou, $B_i = (-\infty, b_i]$ et (34) équivaut à

$$\sum_{j=1}^m \bar{a}_{ij} x_j \geq \bar{b}_i \quad (42)$$

Si seul B_i est flou alors (34) est encore équivalent à (42), c'est-à-dire n'est pas une contrainte souple au sens de Zimmermann /52/. Si les A_{ij} sont flous alors (34) ne dépend pas des fonctions d'appartenance de A_{ij} dans les intervalles $(-\infty, \bar{a}_{ij}]$ où $\bar{a}_{ij} = \sup A_{ij}$. En particulier (34) équivaut à

$$\forall \alpha \in]0, 1], \quad \sum_{j=1}^m \sup(A_{ij})_{\alpha} x_j \leq \sup(B_i)_{\alpha}$$

On peut encore définir des contraintes inégalité si B_i est vu comme la borne floue d'un domaine, de support $S(B_i)$ borné. On peut alors définir des contraintes inégalité au sens de (34) de deux types :

- inégalité faible :

$$\sum_{j=1}^m A_{ij} x_j \subseteq (-\infty, B_i]$$

- inégalité forte :

$$\sum_{j=1}^m A_{ij} x_j \subseteq (-\infty, B_i[$$

Grâce aux notions d'intervalles flous ouverts et fermés introduits au chapitre 4. On sait donc résoudre des programmes linéaires avec des contraintes floues de type (34) ou (35). Le cas où la fonction critère elle-même a des coefficients flous ne semble pas avoir été abordé dans la littérature. Mais les résultats existants sur la paramétrisation de la fonction objectif, sur la dualité, et le récent article de Steuer /45/, qui considère les c_i comme des intervalles peuvent sans doute permettre d'avancer sur ce point.

7.2.5 - CONTRAINTES LINEAIRES FLOUES : APPROCHE PAR LES POSSIBILITES ET LES NECESSITES

Ce qui suit prolonge et complète un travail esquissé dans /14/ par Dubois et Prade.

On a vu que l'approche par Negoita /29/ des contraintes linéaires floues ne correspondait pas à l'approche par contrainte souple de Zimmermann /52/. La raison en est que cette dernière tend à relaxer des contraintes dures tandis que la première peut très bien les durcir encore plus. Si on veut placer les deux approches dans la même perspective, il faut utiliser les notions de possibilité et de nécessité.

Considérons d'abord des contraintes de type égalité, où le second membre est un intervalle flou au sens de (29) et (30), et les A_{ij} sont des intervalles flous. Considérons la contrainte

$$A_i x = \sum_{j=1}^m A_{ij} x_j \tilde{\sim} B_i \quad (43)$$

où $S(B_i) = [\underline{b}_i - \underline{\beta}_i, \bar{b}_i + \bar{\beta}_i]$. (43) peut avoir deux interprétations :

. L'ensemble des x qui peuvent satisfaire l'égalité $\sum_{ij} a_{ij} x_j = b_i$ pour un choix de a_{ij} et b_i compatible avec l'imprécision.

Le degré de possibilité pour que x satisfasse la contrainte i est : (Dubois, Prade /14/)

$$\Pi_i(x) = \sup_{u \in \mathbb{R}} \min(\mu_{A_i x}(u), \mu_{B_i}(u)) \quad (44)$$

Il est facile de vérifier que si les A_{ij} sont des nombres connus alors $\Pi_i(x) = \mu_{C_i}(x)$, au sens de Zimmermann /59/ (équation (27)).

. L'ensemble des x tels que pour tout choix de coefficients a_{ij} compatibles avec les A_{ij} , il existe b_i compatible avec B_i tel que

$$\sum a_{ij} x_j = b_i$$

Dualement à (44), l'ensemble flou des x qui satisfont cette condition a pour fonction d'appartenance :

$$\mathcal{N}_i(x) = \inf_{u \in \mathbb{R}} \max(1 - \mu_{A_i x}(u), \mu_{B_i}(u)) \quad (45)$$

on notera que (45) revient à évaluer le degré d'inclusion au sens de

$$A_i x \prec B_i \quad (46)$$

où \prec est l'inclusion faible définie au chapitre 2.

Au lieu de (45), choisi par dualité avec (44), on peut choisir un index de nécessité cohérent avec l'approche du paragraphe précédent :

$$N_i(x) = \inf_{u \in \mathbb{R}} \mu_{A_i x}(u) \mathcal{I} \mu_{B_i}(u) \quad (47)$$

où \mathcal{I} est une implication définie aussi au chapitre 2. On a

$$N_i(x) = 1 \leftrightarrow A_i x \subseteq B_i$$

Ce qui est faux avec (45) :

$$\begin{aligned} \mathcal{N}_i(x) = 1 &\leftrightarrow A_i(x) \subseteq \overset{\circ}{B}_i \quad (\text{le noyau}) \\ A_i(x) \subseteq B_i &\rightarrow \mathcal{N}_i(x) \geq 0.5 \end{aligned}$$

Néanmoins (46) et (47) sont équivalents si A_i est non flou, et renvoient aux contraintes d'intervalles de Soyster /50/.

On peut encore considérer, à côté de (43)

. des contraintes inégalité exprimées à partir des indices de comparaison (44) (45) ou (47), avec B_i seuil flou au sens de (25-26).

. des contraintes inégalités avec une borne floue B_i de support borné.

On peut distinguer

- l'inégalité au sens large : on substitue $(-\infty, B_i]$ à B_i dans (44-47).
- l'inégalité au sens strict : on substitue $(-\infty, B_i[$ à B_i dans (44-47).

Dans tous les cas, l'ensemble flou C_i des décisions admissibles pour la contrainte i sera défini par

$$\mu_{C_i}(x) = \Pi_i(x), \quad \text{ou } \mathcal{N}_i(x), \quad \text{ou } N_i(x)$$

On peut alors transformer le problème de trouver une solution admissible d'un ensemble de contraintes floues du type ci-dessus en un problème classique par l'un des deux moyens suivants :

- se fixer des seuils a priori, s_i pour chaque contrainte, et chercher x tel que

$$\forall i = 1, n ; \mu_{C_i}(x) \geq s_i \quad (48)$$

- maximiser $\mu_D(x)$ où D est obtenu en agrégeant les C_i (cf. équation 31).

En particulier, si l'agrégation utilise le minimum, on se ramène à la forme (32). C'est ce qu'on suppose dans la suite, comme alternative à (48).

7.2.6 - INTERPRETATION DETERMINISTE DES CONTRAINTES FLOUES

Examinons maintenant dans quel cas ces formulations mènent à des problèmes simples. Pour cela on suppose que les A_{ij} sont représentables à l'aide d'intervalles flous de type $L_j R_j$, et que les B_i sont de type LR.

A Contrainte souple (Possibilité d'égalité)

L'approche qui suit est équivalente, bien que présentée différemment à celle utilisée par Oh'Eigeartaigh /31/.

Rappelons que, pour tout intervalle flou M , $\mu_M = \min(\mu_{(-\infty, M]}, \mu_{[M, +\infty)})$ où $[M, +\infty)$ (resp $(-\infty, M]$) est l'ensemble des nombres qui peuvent être plus grands que M (resp plus petit). On vérifie (cf. Chap.4) que, \cap étant aussi interprétée par le minimum:

$$\begin{aligned} \Pi_i(x) &= \sup \mu_{B_i} \cap \mu_{A_i x} \\ &= \min(\sup \mu_{(-\infty, B_i]} \cap \mu_{[A_i x, +\infty)}, \sup \mu_{(-\infty, A_i x]} \cap \mu_{[B_i, +\infty)}) = \pi_i^* \end{aligned}$$

Donc la contrainte $\Pi_i(x) \geq \pi_i$ se traduit par les contraintes :

$$\mu_{[B_i, +\infty)}^{-1}(\pi_i) \leq \mu_{(-\infty, A_i x]}^{-1}(\pi_i) \quad (49)$$

et

$$\mu_{[A_i x, +\infty)}^{-1}(\pi_i) \leq \mu_{(-\infty, B_i]}^{-1}(\pi_i) \quad (50)$$

où f^{-1} est la fonction réciproque de f .

En revanche $\forall \pi_i \in]\pi_i^*, 1]$ l'une des deux contraintes est violée. Ceci est dû au fait que les nombres flous considérés sont convexes et que par conséquent π_i^* est atteint pour des valeurs situées dans l'intervalle limité par les noyaux de $A_i x$ et B_i . Par conséquent si on s'intéresse à la possibilité pour que la contrainte égalité (43) soit vérifiée, imposer un seuil minimal π_i à ce degré de possibilité est équivalent à (49-50).

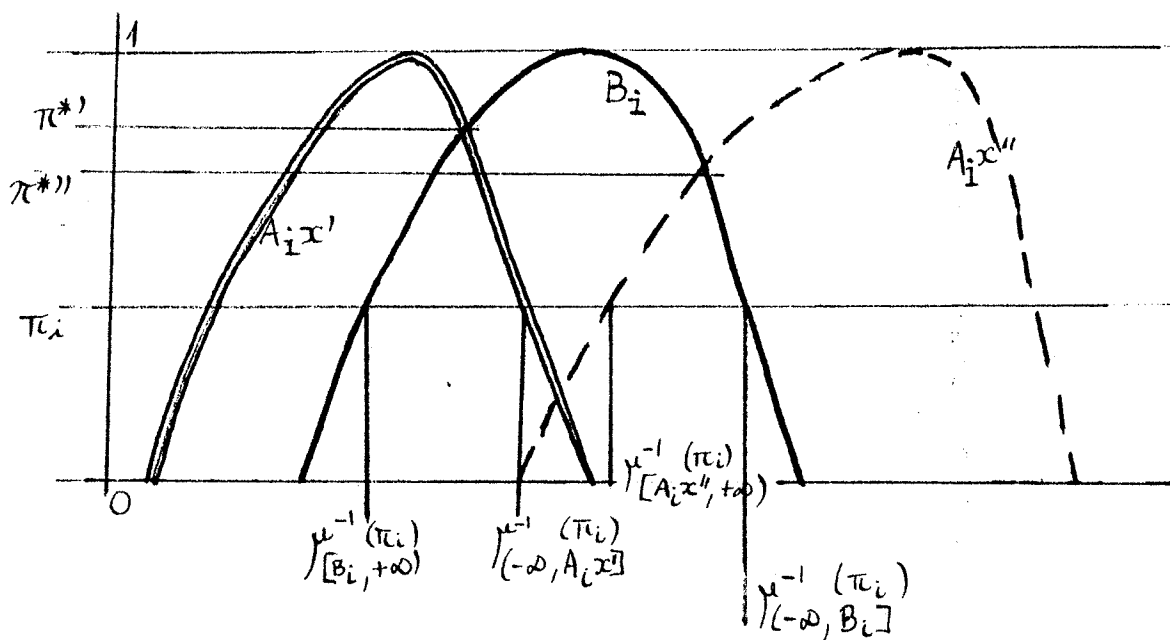
Dans le cas d'une contrainte inégalité avec un seuil flou $B_i = (-\infty, B_i]$ alors (49) est toujours vrai, et peut être supprimé. On peut toujours poser :

$$A_{ij} = (\underline{a}_{ij}, \bar{a}_{ij}, \underline{\alpha}_{ij}, \bar{\alpha}_{ij})_{L_j R_j}, \quad B_i = (\underline{b}_{ij}, \bar{b}_{ij}, \underline{\beta}_{ij}, \bar{\beta}_{ij})_{LR}$$

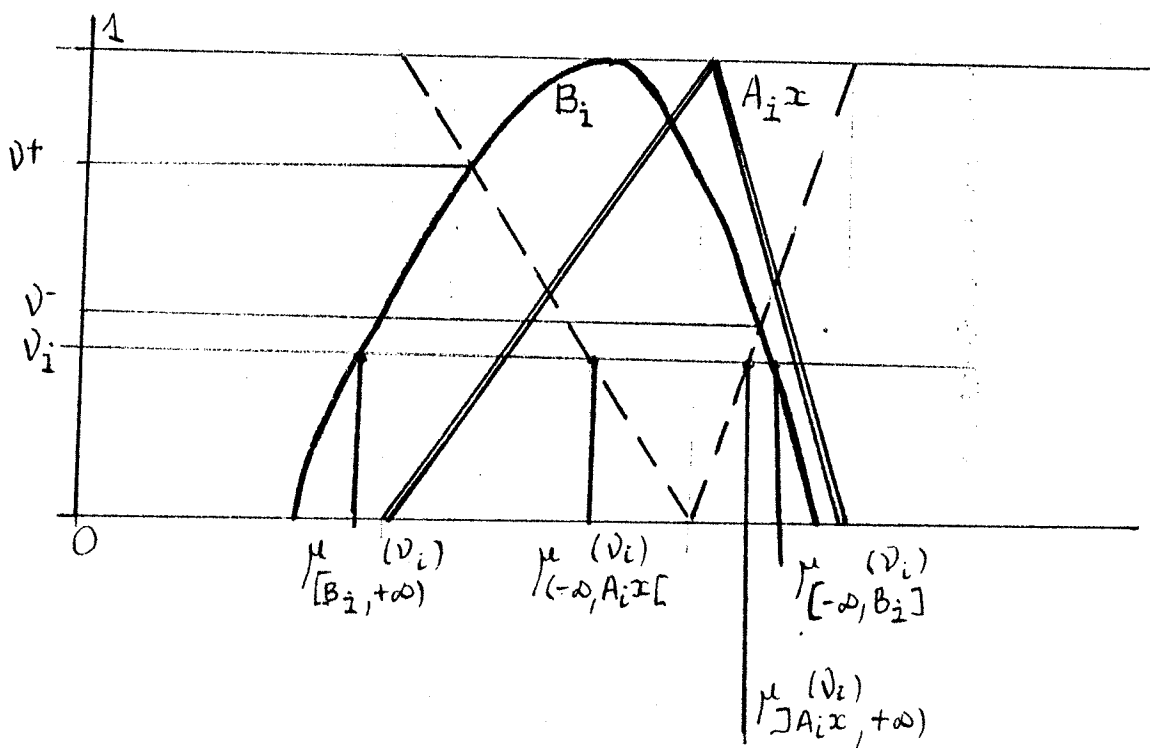
où $L_j, R_j, \forall j, L, R$, sont des fonctions de référence au sens du chapitre 4. Alors (49) et (50) deviennent ($x > 0$)

$$\underline{b}_i - \underline{\beta}_i \cdot L^{-1}(\pi_i) \leq \sum_j (\bar{a}_{ij} + \bar{\alpha}_{ij} R_j^{-1}(\pi_i)) x_{ij} \quad (51)$$

$$\sum_j (\underline{a}_{ij} - \underline{\alpha}_{ij} \cdot L^{-1}(\pi_i)) x_{ij} \leq \bar{b}_i + \bar{\beta}_i R_j^{-1}(\pi_i) \quad (52)$$



CAS A



CAS B

C'est-à-dire que si on se fixe le seuil π_i , le problème (48) conserve la linéarité des contraintes. Néanmoins si π_i est une variable supplémentaire comme dans le problème (32), alors la contrainte n'est pas linéaire.

Notons enfin que si on fait $\alpha_{ij} = 0$ dans (51), $\bar{\beta}_i = d_i$, $x_{m+1} = L^{-1}(\pi_i)$ on retrouve les contraintes de type (33) obtenues par l'approche de Zimmermann.

NB : Dubois /9/ a interprété l'égalité $A_i x \underline{\simeq} B$ comme la comparaison de $A_i x \ominus B$ avec 0. On peut montrer que

$$\Pi_i(x) \geq \pi_i \leftrightarrow \mu_{A_i x \ominus B}(0) \geq \pi_i$$

c'est-à-dire que l'on retrouve les mêmes contraintes linéaires (49-50) ou (51-52) qu'avec l'approche ci-dessus. Tanaka et Asai /47/ ont fait de même, mais avec des fonctions d'appartenance triangulaires.

B. Contrainte dure (Nécessité d'égalité)

La contrainte $\mathcal{N}_i(x) \geq v_i$ s'explique en notant que

$$\mathcal{N}_i(x) = \min(v^+, v^-)$$

avec

$$v^+ = \inf \mu_{(-\infty, A_i x] \cup [B_i, +\infty)}$$

$$v^- = \inf \mu_{(-\infty, B_i] \cup]A_i x, +\infty)}$$

Donc la contrainte $\mathcal{N}_i(x) \geq v_i$ se traduit par les deux inégalités

$$\mu_{[B_i, +\infty)}^{-1}(v_i) \leq \mu_{(-\infty, A_i x]}^{-1}(v_i)$$

$$\mu_{]A_i x, +\infty)}^{-1}(v_i) \leq \mu_{(-\infty, B_i]}^{-1}(v_i)$$

soit en notant que $\overline{]A_i x, +\infty)} = [A_i x, +\infty)$

$$\mu_{[B_i, +\infty)}^{-1}(v_i) \leq \mu_{[A_i x, +\infty)}^{-1}(1-v_i) \quad (53)$$

$$\mu_{(-\infty, A_i x]}^{-1}(1-v_i) \leq \mu_{(-\infty, B_i]}^{-1}(v_i) \quad (54)$$

Dans le cas d'une contrainte d'inégalité, avec un seuil flou B_i tel que $B_i = (-\infty, B_i]$ seule demeure la contrainte (54).

Avec des intervalles de type LR pour la représentation de A_{ij} et B_i , comme précédemment on peut écrire (53) et (54) sous la forme

$$\underline{b}_i - \underline{\beta}_i L^{-1}(v_i) \leq \sum_j (\underline{a}_{ij} - \underline{\alpha}_{ij} L_j^{-1}(1-v_i)) x_{ij} \quad (55)$$

$$\sum_j (\bar{a}_{ij} + \bar{\alpha}_{ij} R_j(1-v_i)) x_{ij} \leq \bar{b}_i + \bar{\beta}_i R(v_i) \quad (56)$$

Quand les coefficients A_{ij} ne sont pas flous ($\bar{\alpha}_{ij} = \underline{\alpha}_{ij} = 0$) mais imprécis (55) et (56) veulent dire $[\underline{a}_{ij}, \bar{a}_{ij}] \subseteq (B_i)_{v_i}$.

Une fois de plus, en fixant le seuil v_i on retrouve une formulation linéaire au sens classique. On notera que l'on ne peut ramener la contrainte dure à une comparaison entre $A_i x \in B$ à 0 comme précédemment.

Pour résumer, le problème (48), avec $\mu_{C_i} = \Pi_i$ ou \mathcal{N}_i peut toujours se ramener à un ensemble de contraintes linéaires, si on fixe des seuils s_i bornant par valeur inférieure les degrés d'appartenance $\mu_{C_i}(x)$. Une contrainte floue d'égalité équivaut à deux contraintes linéaires de type inégalité. Une contrainte floue inégalité équivaut à une contrainte linéaire de type inégalité.

On peut donc en se fixant des seuils de possibilité et de nécessité qui définissent l'ensemble des solutions admissibles, ramener un problème de programmation linéaire avec contraintes floues à un programme linéaire un peu plus grand.

Si on estime impossible, pour des questions d'interprétabilité, de fixer les seuils s_i a priori on est amené à résoudre le problème (31) sous la forme

maximiser x_{m+1}

$$\Pi_i(x) \geq x_{m+1} \quad i = 1, q \quad (57)$$

$$\mathcal{N}_i(x) \geq x_{m+1} \quad i = q+1, n \quad (58)$$

qui généralise la formulation de Zimmermann /52/, donnée en (32) et (33) ; les q premières contraintes correspondent aux contraintes souples, les $n-q$ restantes aux contraintes dures. On peut à l'aide d'une technique de dichotomie ramener ce problème à une succession de problèmes de type (48) qui sont linéaires. La procédure peut être résumée comme suit

1. Soit $\underline{\alpha}$ et $\bar{\alpha}$ des bornes supérieures et inférieures de la valeur optimale de x_{m+1} dans le problème ci-dessus. On fixe $\underline{\alpha} = 0$, $\bar{\alpha} = 1$
2. fixer $x_{m+1} = \frac{1}{2} (\underline{\alpha} + \bar{\alpha})$
3. Résoudre le problème (linéaire) (48) avec $s_i = x_{m+1}$ fixé précédemment
4. Si on a trouvé une solution x^* = aller en 5 sinon on pose $\bar{\alpha} = x_{m+1}^*$ retour en 2.
5. Poser $\underline{\alpha} = \min(\min_{i=1,q} H_i(x^*), \min_{i=q+1,n} N_i(x^*))$. Retour en 2.

On s'arrête dès que la différence $|\bar{\alpha} - \underline{\alpha}|$ est en dessous d'un seuil donné.

7.2.7 - COMPARAISON AVEC LA PROGRAMMATION STOCHASTIQUE

On se contente ici d'envisager un aspect de la programmation stochastique, à titre d'exemple : celui de la programmation linéaire à contraintes aléatoires (Charnes et Cooper /4/). La discussion qui suit est d'après Dubois /9/.

Considérons le problème linéaire

$$\min C x$$

$$\text{sous les contraintes } A_i x \leq B_i \quad x \geq 0, \quad i = 1, n$$

où les A_{ij} et les B_i sont des variables aléatoires de valeurs moyennes \bar{A}_{ij} et \bar{B}_i et d'écart-types $\sigma_{A_{ij}}$ et σ_{B_i} . La satisfaction de la contrainte i peut s'exprimer en probabilité, c'est-à-dire qu'on peut se donner un seuil p_i de probabilité et demander

$$\text{Prob}(A_i x \leq B_i) \geq p_i \tag{59}$$

Cette équation est la contrepartie probabiliste de (48).

Si les A_{ij} et B_i sont des variables aléatoires quelconques il est clair que (59) ne peut s'exprimer simplement en fonction des variables x . Plaçons nous dans le cas favorable où ces variables aléatoires sont gaussiennes.

Dans ce cas $Y_i = A_i x - B_i$ est une variable aléatoire normale telle que

. Valeur moyenne $\bar{Y}_i = \bar{A}_i x - \bar{B}_i$

. Variance $\sigma_{Y_i}^2 = \sum_{j=1}^m \sigma_{A_{ij}}^2 x_j^2 + \sigma_{B_i}^2$

(59) s'exprime dans ce cas par

$$\begin{aligned} \text{Prob}(Y_i > 0) &= \int_{-\frac{\bar{Y}_i}{\sigma_{Y_i}}}^{+\infty} e^{-u^2/2} du \\ &\geq p_i \stackrel{\Delta}{=} \int_{\psi(p_i)}^{+\infty} e^{-u^2/2} du \end{aligned}$$

soit

$$\bar{Y}_i + \psi(p_i) \sigma_{Y_i} \geq 0$$

qui se développe en

$$\bar{A}_i x - \bar{B}_i + \psi(p_i) \sqrt{\sum_{j=1}^m \sigma_{A_{ij}}^2 x_j^2 + \sigma_{B_i}^2} \geq 0 \quad (60)$$

(60) est la contrepartie probabiliste de (51-52) ou (55-56).

On voit donc qu'on se ramène à une contrainte quadratique en x . Donc on saisit clairement ici la différence de complexité entre l'approche floue et l'approche probabiliste de la programmation linéaire à coefficients mal connus :

- . Dans un cas, le problème équivalent est à contrainte quadratique pour des distributions gaussiennes ; on ne sait pas faire dans le cas général.
- . Dans l'autre, le problème équivalent est toujours linéaire, avec plus de contraintes, pour des intervalles flous quelconques.

Cela ne discrédite pas pour autant l'approche stochastique, qui si elle peut être menée à bien donne des résultats plus précis, que la programmation linéaire floue, laquelle n'est qu'une généralisation de l'analyse de sensibilité. On veut simplement faire remarquer que cette dernière peut se substituer à la programmation stochastique soit que celle-ci s'avère trop complexe à mettre en oeuvre, soit que par manque de données statistiques valables, elle se justifie mal. Luhanjula /27/ dans un récent article a combiné les deux approches en utilisant la probabilité d'un événement flou au sens de Zadeh (chap.2 § 4.3.1).

7.2.8 - A PROBLEME D'OPTIMISATION FLOU, SOLUTION FLOUE

Dans les paragraphes de la section 2 qui précèdent on a ramené les problèmes d'optimisation en environnement flou à la découverte d'une solution dont le degré d'appartenance à l'ensemble flou formé par les contraintes et les objectifs soit maximum. On peut souhaiter plus, car cette solution n'est pas la seule qui soit digne d'être retenue. Il est naturel de penser que l'imprécision du problème va de pair avec celle de sa solution, ou tout au moins avec sa non-unicité. Contrastant avec l'approche de la solution unique étudiée dans le cas général par Tanaka et al /46/ et Flachs-Pollatschek /18/, Orlovski /32/ a défini ce que pouvait être une solution floue d'un problème d'optimisation sur un domaine flou. En substance si $X(\alpha)$ est l'ensemble des solutions optimales sur l' α -coupe D_α du domaine des solutions admissibles, la solution floue du problème est l'ensemble flou N tel que

$$\forall x, \mu_N(x) = \sup\{\alpha/x \in D_\alpha\}$$

Cette approche a été appliquée à la programmation linéaire par Verdegay /49/. En pratique, elle semble difficilement utilisable.

Une autre idée consiste à rechercher une solution floue sous la forme d'une relation floue séparable $X_1 \times X_2 \dots \times X_m$ sur \mathbb{R}^m , où X_i est un intervalle flou et x est le produit cartésien flou. Par exemple on peut rechercher les intervalles flous X_i tels que

$$\sum_{j=1}^m A_{ij} \otimes X_j \approx B_i \quad (\text{au sens de (43)})$$

Tanaka et Asai /47/ ont traité ce type de contrainte floues pour

A_{ij} précis = a_{ij} , B_i intervalle flou de fonction d'appartenance linéaire et symétrique, et les X_j sont recherchés de la même façon. On cherche les valeurs modales \bar{x}_j et étalements ϵ_j tels que la possibilité $\Pi_i(X)$ de vérifier la contrainte soit maximale. En fait si on pose $X_j = (\bar{x}_j, \epsilon_j, \eta_j)_{LR}$, avec les a_{ij} positifs, on peut écrire l'homologue de (49) et (50) qui seront des contraintes linéaires en $\bar{x}_j, \epsilon_j, \eta_j$. On peut aussi définir l'homologue de (53-54) pour la contrainte dure. On notera qu'on obtient un système d'équations linéaires à 3 m inconnues.

Quand les A_{ij} sont flous, positifs et de type L-R, les contraintes déterministes équivalentes aux contraintes floues vont être non-linéaires. Au vu de la formule du produit de deux nombres flous positifs (cf. chap.4 § 2.5.3), il est clair qu'elles seront quadratiques en $\bar{x}_j, \epsilon_j, \eta_j$. Ce sera aussi le cas des contraintes de type

$$\sum_{j=1}^m A_{ij} \otimes X_j \subseteq B_i \quad (\text{au sens de Zadeh})$$

étudiées en 2.4. Dubois et Prade /14/ ont utilisé les approximations du 1er ordre de type L-R des quantités $A_{ij} \otimes X_j$ (chapitre 4 § 2.5.4) pour se ramener à des contraintes déterministes linéaires.

L'intérêt des solutions floues de problèmes d'optimisation en environnement imprécis est qu'elles laissent encore une marge de choix pour la définition d'une solution définitive. Cette démarche est en accord avec la philosophie des techniques modernes de l'aide à la décision qui préfèrent l'idée d'un ensemble de bonnes solutions à celui d'une (pseudo) solution optimale. (L'auteur a d'ailleurs tiré parti de cette méthodologie dans /6/). Quand le problème d'optimisation est posé en tenant compte de l'imprécision sur les données, cela se justifie d'autant plus.

Un exemple d'utilisation de solution floue d'un problème d'optimisation est fourni à propos de l'application industrielle évoquée ci-après.

NB : Le caractère positif des A_{ij} dans les contraintes floues n'est pas une hypothèse nécessaire. On peut traiter aussi le cas de coefficients flous négatifs, si on utilise une représentation L-L (symétrique). Plus délicat est le cas où les A_{ij} sont "à cheval" sur 0, car les formules de calcul du produit

de nombres flous sont moins simples.

7.2.9 - UN PROBLEME DE REGLAGE OPTIMAL D'UN SYSTEME DE PRODUCTION

Dans cette section on présente une application de l'arithmétique floue à un problème de détermination de paramètres de coupe sur des machines-outil précédemment abordé par l'auteur. On se reportera pour plus de détails à /10/, et surtout /9/.

Une opération d'usinage est généralement définie par le passage d'un outil coupant ou perçant la surface d'une pièce métallique. Les principaux paramètres de réglage sont la vitesse et l'avance de coupe qui conditionnent directement la qualité du résultat de l'opération. Les principales contraintes pesant sur ces paramètres sont relatives à la façon d'utiliser l'outil, à la dynamique de la machine-outil, et à l'état de surface souhaité pour la pièce. On peut poser le problème de réglage d'une machine-outil pour une opération d'usinage donnée comme un problème d'optimisation dont le critère peut être par exemple le coût de l'opération, la quantité de métal enlevée, etc... (voir Dubois /9/ pour un recensement des critères et méthodes). Néanmoins, un "régleur" est capable, pour une opération donnée de définir très vite des plages de vitesse et d'avance de coupe qui permettront d'atteindre une certaine qualité d'exécution.

Le problème devient plus complexe si on envisage de régler un ensemble de machines disposées en une suite de poste de travail synchronisés. On a alors les deux contraintes supplémentaires qui couplent les divers réglages à effectuer.

. Les calculs de productivité imposent une certaine cadence qui sera la même pour chaque poste.

. On souhaite harmoniser l'usure moyenne des outils de façon à planifier leur remplacement.

Le régleur est alors obligé de procéder par tâtonnement jusqu'à ce qu'il trouve des réglages satisfaisants, respectant la contrainte de cadence. En général, il ne sait pas bien prendre en compte les problèmes d'usure.

C'est pourquoi une méthode a été proposée, qui peut être vue comme un outil interactif d'aide au réglage. Elle a les caractéristiques suivantes

- elle tient compte du savoir du régleur pour le respect de la qualité de coupe
- elle limite le tâtonnement, et tente de tenir compte des usures.
- elle est décentralisée et procède de l'analyse individuelle de chaque poste.

Elle utilise comme données des contraintes sur l'avance et la vitesse de coupe, fournies pour chaque opération par le régleur, sous la forme d'intervalles flous. Celui-ci fixe pour chaque opération et chaque paramètre de réglage des plages préférée et permise qui définissent le noyau et le support d'un intervalle flou. La forme de la fonction d'appartenance peut être précisée par un test sur l'opportunité de s'éloigner de la plage préférée.

Soit \hat{t} la cadence imposée pour chaque poste i . L'opération effectuée par ce poste est donc caractérisée par une vitesse de coupe floue V_i et une avance de coupe floue A_i . La durée de l'opération est fonction de ces deux paramètres sous la forme

$$t_i = f(a_i, v_i) = \frac{K}{a_i v_i}$$

où K est une constante déduite de la géométrie de la pièce. La contrainte de cadence s'écrit donc :

$$t_i = f(a_i, v_i) = \hat{t}$$

Pour le poste i , on peut donc définir le problème d'optimisation des paramètres de coupe, conformément à l'approche de cette section.

$$\text{maximiser } (\min(\mu_{V_i}(v_i), \mu_{A_i}(a_i)) / f(a_i, v_i) = \hat{t})$$

La solution (a_i^*, v_i^*) est facile à trouver car on sait qu'elle vérifie

$$\mu_{V_i}(v_i^*) = \mu_{A_i}(a_i^*) = \mu_{f(A_i, V_i)}(\hat{t}) \quad (\hat{=} \alpha^*)$$

La quantité $f(A_i, V_i)$ est aisément obtenue à l'aide des formules de calcul

des intervalles flous ; on en déduit le degré optimal α^* de satisfaction des contraintes et, par report on en déduit v_i^* et a_i^* , qui vérifient la contrainte de cadence. Le calcul est mené de façon légèrement différente pour un poste multiopérations (voir Dubois /9/).

Néanmoins si on veut harmoniser l'usure moyenne des outils, on va être amené à modifier ces valeurs optimales. C'est ici que l'idée de solution floue d'un problème d'optimisation prend son intérêt.

L'ensemble flou des vitesses de coupe qui satisferont la contrainte de cadence tout en respectant les spécifications de qualité de coupe sera $V_i^!$ tel que

$$\mu_{V_i^!}(v_i) = \min(\mu_{V_i}(v_i), \mu_{A_i}(\frac{K}{tv_i}))$$

on a clairement $V_i^! \subseteq V_i$; on a donc diminué l'imprécision initiale. De plus $\sup \mu_{V_i^!} = \mu_{V_i}(v_i^*)$, par définition

Des formules empiriques permettent de calculer la durée de vie d'un outil. On peut l'exprimer comme le nombre de pièces (d'un même type) produites par un même outil de caractéristiques données, soit N_i . On sait que l'outil pourra usiner N_i pièces avec une certaine probabilité, suffisamment proche de 1. N_i dépend des conditions de coupe et on peut poser

$$N_i = g(a_i, v_i) = g(\frac{K}{t \cdot v_i}, v_i) = h(v_i)$$

Connaissant la cadence \hat{t} et la solution floue $V_i^!$, on peut, par le principe d'extension, définir un nombre flou \hat{N}_i de pièces usinées par outil sur le poste i.

$$\hat{N}_i = h(V_i^!)$$

Le problème d'harmonisation des remplacements d'outils sur les n-postes peut donc être formulé comme suit (cas d'un outil par poste) :
trouver l'entier m (nombre de pièces produites entre deux arrêts de la chaîne)
les entiers k_i , $i = 1, n$ (nombre d'arrêts de la chaîne entre deux changements de l'outil i).

tel que $\beta = \min_{i=1, n} \mu_{\hat{N}_i}^{\wedge}(k_i, m)$ soit maximum

Ce problème peut être résolu par une technique d'énumération, ou même "visuellement" par un opérateur à qui on soumettrait les fonctions d'appartenance des \tilde{N}_i .

NB : Si la valeur optimale de β est trop faible, on peut légèrement élargir certains intervalles \tilde{N}_i en remarquant qu'il est possible de changer certains outils avant qu'ils atteignent leur seuil d'usure (mais pas trop avant pour rester rentable).

Connaissant m et $k_i \forall i$, on en déduit les réglages définitifs (\bar{v}_i, \bar{a}_i) , $i = 1, n$, vérifiant

$$\forall i, \bar{v}_i = h^{-1}(k_i, m), \quad \bar{a}_i = \frac{K}{\hat{t}\bar{v}_i}$$

La procédure ainsi décrite est simple à mettre en oeuvre (voir néanmoins /9/ pour les postes multioutils), et semble réaliste, car utilisant directement le savoir des régleurs. Elle se prête bien à une mise en oeuvre sur calculateur, sur un mode interactif, et peut donner naissance à un outil informatique d'aide au réglage.

Il est possible de poser le problème d'optimisation des conditions de coupes sur les n postes dans son ensemble (voir Dubois /9/). On peut prendre comme critère la minimisation de la fréquence d'arrêt systématique de la chaîne. On obtient un programme non linéaire à variables mixtes qui peut être délicat à résoudre, et dans lequel on n'intègre pas aussi nettement le savoir du régleur.

Une autre application des modèles de décision en environnement imprécis à l'optimisation de processus d'usinage est décrit dans Sakai, Tanaka et Ido /46/.

7.3 - L'IMPRÉCISION DANS LES MODÈLES DE FILES D'ATTENTE

Dans les sections 1 et 2 de ce chapitre on a mis en évidence deux exemples d'utilisation de l'arithmétique floue en recherche opérationnelle où l'approche proposée contrastait avec sa contrepartie probabiliste par la

simplicité des calculs à mener et par son interprétation sémantique, tout en fournissant des résultats exploitables, bien que moins précis que ceux que l'on peut espérer par une analyse probabiliste. Dans cette section, on cite un exemple de modèle flou qui apporte beaucoup trop peu d'informations par rapport au modèle probabiliste homologue : c'est le cas d'une file d'attente. Ce type d'applications est d'ailleurs l'une des plus convaincantes et les plus utiles de la théorie des probabilités.

L'exposé qui suit s'inspire d'un article de Prade /36/, puis indique d'autres façons d'introduire l'imprécision dans un processus stochastique.

7.3.1 - MODELE POSSIBILISTE DE FILE D'ATTENTE

Soit un guichet monoserveur où doivent être servis des clients qui arrivent de façon plus ou moins imprévisible. En général on définit, à partir d'observations des temps de service, et des durées interarrivées, des lois de probabilité sur ces temps, ce qui permet d'évaluer les performances du système.

Ici on supposera que l'on ne connaît les temps de service s et les temps entre deux arrivées de clients y que sous la forme d'intervalles flous, S et Y respectivement. Ce sera le cas si on recueille cette information auprès d'un "habitué" qui dira par exemple, pour définir S : "il faut 10 mn pour servir un client, en général. En tout cas, il faut au moins 5mn. Mais il est arrivé qu'un client reste 20 mn".

Si X_n est la date d'arrivée de $n^{\text{ième}}$ client

w_n est le temps passé par le $n^{\text{ième}}$ client dans la file

s_n est le temps passé par le $n^{\text{ième}}$ client au guichet

alors on a la relation bien connue (e.g. Kleinrock /24/) :

$$w_{n+1} = \max(0, X_n + s_n + w_n - X_{n+1}) \quad (61)$$

exprimant que le client $n^{\circ} n+1$ n'attendra que s'il arrive avant que le $n^{\text{ième}}$ client ait eu le temps de se faire servir. En particulier $w_1 = 0$. On suppose la discipline de service premier arrivé - premier servi.

Les quantités s_n et $X_n - X_{n+1}$ sont connues par les intervalles flous S et Y respectivement. Donc les quantités w_n seront connues sous forme d'intervalles flous W_n , ensembles flous de valeurs possibles de w_n , par

$$W_{n+1} = \widetilde{\max}(0, S\ominus Y \oplus W_n) \quad (62)$$

où \ominus et \oplus sont les soustraction et addition d'intervalles flous, supposés non interactifs dans un premier temps,

Il est facile de développer (61) en

$$w_{n+1} = \max_{i=1, n} \left(0, \sum_{j=1}^{n-i+1} (s_j - y_j) \right) \quad (63)$$

Avec l'hypothèse de non-interaction entre les variables s_j et y_j , on vérifie que $(S\ominus Y) \oplus (S\ominus Y) = 2(S\ominus Y)$. Donc (62) s'écrit encore

$$W_{n+1} = \widetilde{\max}_{i=1, n} (0, i(S\ominus Y)) \quad (64)$$

NB : Prade /36/ a exprimé (62) sous la forme $W_{n+1} = R \circ W_n$ où R est une relation floue construite sur $S\ominus Y$. Néanmoins, l'utilisation des symboles de l'arithmétique floue rend plus visibles les manipulations à effectuer pour parvenir au résultat final.

La question qui se pose est alors de trouver pour W_n une valeur limite W , qui sera d'ailleurs l'"ensemble flou propre" de la relation R c'est-à-dire $W = R \circ W$ au sens de Sanchez (cf. fin du chap.3). Remarquons d'abord que

$$\forall i, \mu_i(S\ominus Y)^{(0)} = \mu(S\ominus Y)^{(0)}$$

On a alors 3 cas à considérer.

A. $S\ominus Y \geq 0$ (nombre flou positif) alors $W_{n+1} = n(S\ominus Y)$.

et n 'a donc pas de limite finie. C'est le cas, où quelles que soient les dates d'années des clients, le service ne peut jamais faire face à la demande

B. Le noyau de $S\ominus Y$ est dans la zone négative.

Alors il est facile de voir que

$$\mu_{W_{n+1}}(0) = 1 \quad \forall n$$

$$\mu_{W_{n+1}}(u) = \mu_{n(S\Theta Y)}(u) = \mu_{(S\Theta Y)}\left(\frac{u}{n}\right) \geq \mu_{W_n}(u)$$

pour $u > 0$

et que $W = \lim_{n \rightarrow +\infty} W_{n+1}$ est défini par

$$\mu_W(0) = 1$$

$$\mu_W(u) = \mu_{S\Theta Y}(0) \stackrel{\Delta}{=} \Pi_a, \quad \forall u > 0$$

c'est-à-dire qu'il est plus plausible de ne pas avoir de client en attente que d'en avoir ($\mu_{S\Theta Y}(0) < 1$).

C. Si $0 \in \overset{0}{S\Theta Y}$, alors $\forall u \geq 0 \quad \mu_W(u) = 1$, c'est-à-dire qu'on n'a aucune information.

Donc un tel modèle est asymptotiquement très peu informatif.

En revanche les distributions de possibilité successives $W_n, n > 0$, fournissent une étude de pire cas du comportement transitoire du système.

Si on introduit une interaction entre s_i et y_i , alors l'information asymptotique qu'on peut recueillir sera beaucoup plus précise. La plus forte interaction qu'on puisse admettre, si on définit toujours $S\Theta Y$ à partir du principe d'extension et du produit cartésien $S \times Y$ consiste à bâtir ce dernier avec la t-norme T_W (voir annexe, et chap.4 § 3.2).

On rappelle que si M est un intervalle flou de type L-R, $M \oplus M \neq 2M$. Plus précisément, si $S = (\bar{s}, \alpha, \beta)_{LR}$ $Y = (\bar{y}, \gamma, \delta)_{LR}$ alors, au sens de T_W

$$S\Theta Y = (\bar{s}-\bar{y}, \max(\alpha, \delta), \max(\beta, \gamma))_{LR}$$

$$\sum_{i=1}^n S\Theta Y = (n(\bar{s}-\bar{y}), \max(\alpha, \delta), \max(\beta, \gamma))_{LR} \neq n(S\Theta Y)$$

L'hypothèse qui sous-tend T_W est que le temps de service et le temps entre deux arrivées ne s'éloignent jamais simultanément de leurs valeurs modales \bar{s}

et \bar{y} . Seul l'un des deux écarts se produit au plus, pour chaque client. L'extension de (63) n'est plus (64). Néanmoins, on peut vérifier que les cas A et C étudiés plus haut conduisent aux mêmes conclusions qu'avec des opérations arithmétiques floues non-interactives.

En revanche dans le cas B, on vérifie que, au sens de T_W ,

$$\tilde{\max}_{i=1}^n (\sum_{i=1}^n S\Theta Y, \sum_{i=1}^{n-1} S\Theta Y, 0) = \tilde{\max}_{i=1}^{n-1} (\sum_{i=1}^n S\Theta Y, 0)$$

de sorte que, au lieu de (64) on obtient

$$W_n(T_W) = \tilde{\max}(0, S\Theta Y) \quad \forall n \quad (65)$$

En particulier la possibilité d'attendre au moins u unités de temps est plus faible ($= \mu_{W(T_W)}(u)$) que dans le cas non interactif ($\mu_{W(\min)}(u)$)

Exemple : (Prade /36/)

$$S = (2, 2, 5)_{LL} ; \quad Y = (5, 5, 5)_{LL} \quad \text{où } L \text{ est linéaire}$$

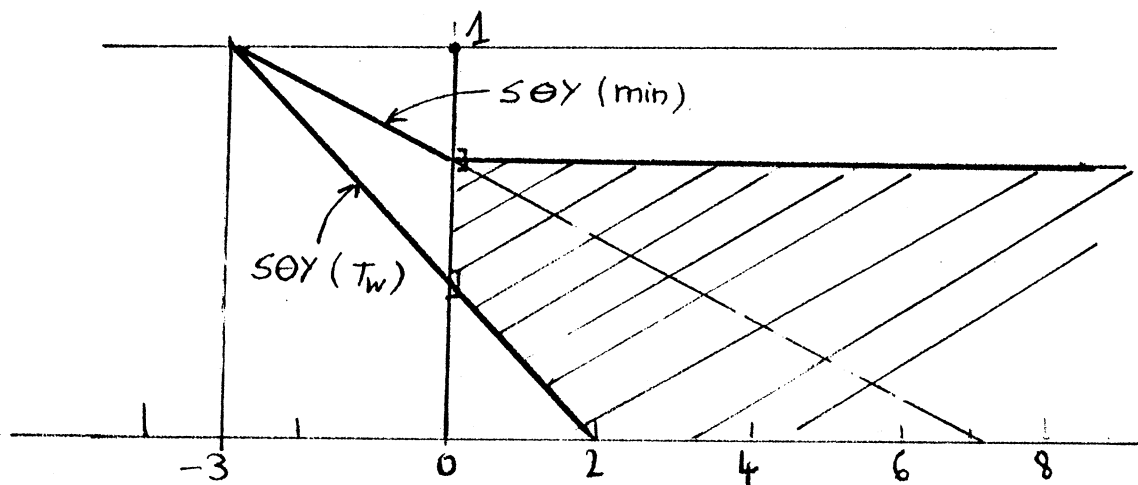
$$S\Theta Y = (-3, 7, 10)_{LL} \quad (\text{au sens de min})$$

$$= (-3, 5, 5)_{LL} \quad (\text{au sens de } T_W)$$

La possibilité asymptotique d'attente $\Pi_a(\lim_{u \rightarrow 0_+} \mu_W(u))$ est donc

. cas non interactif $\Pi_a = 1 - \frac{3}{10} = 0.7$

. cas interactif $\Pi_a = 1 - \frac{3}{5} = 0.4$



De plus avec l'interaction exprimée par la t-norme T_W , on voit que le temps d'attente ne dépassera jamais deux unités de temps. La zone hachurée contient toutes les fonctions d'appartenance $\mu_{W(\perp)}$ obtenues en faisant le calcul arithmétique au sens de \perp .

En particulier, on peut montrer dans certains cas que la limite de W_n , calculé au sens d'une t-norme \perp est à support borné:

$$\exists n_0 \quad n > n_0 \rightarrow W_n = \max(0, \sum_{j=1}^{n_0} (S\Theta Y))$$

C'est le cas par exemple pour $\alpha \perp \beta = \max(0, \alpha + \beta - 1)$ et $L(u) = (\max(0, 1 - x^p))$.

Malgré la possibilité de réduire l'incertitude en introduisant des hypothèses d'interaction entre les variables manipulées, on voit qu'un modèle de file d'attente de type possibiliste est beaucoup plus pauvre qu'un modèle probabiliste. On peut le voir comme un système D/D/1 avec de l'erreur sur le temps de service et le temps inter-arrivées.

7.3.2 - AUTRES MODELES STOCHASTIQUES ENTACHES D'IMPRECISION

Prade /36/ a suggéré d'autres types de modèles de file d'attente qui prennent en compte l'imprécision de nature possibiliste.

. Un système à temps de service flou et loi d'arrivée aléatoire (par exemple loi de Poisson). On peut voir ce modèle comme un ensemble flou de systèmes M/D/1 dont les performances seront entachées d'imprécision de par la mauvaise connaissance du temps de service s . Par exemple le nombre moyen de clients dans le système s'exprime en fonction de l'utilisation $u = \lambda s$ du guichet par (Kleinrock /24/).

$$\bar{N} = u + \frac{u^2}{2(1-u)} = f(u).$$

où λ est le taux d'arrivée.

On peut, à l'aide du principe d'extension, calculer un nombre moyen (flou) de clients dans le système qui vaudra $f(\hat{u})$ où $\hat{u} = \lambda S$, S étant l'intervalle flou contenant les temps de service. Yager /51/ a une approche similaire pour un système M/M/1 dont le taux d'arrivée λ est imprécis.

. Un système à discipline de service incomplètement spécifiée du type : un client qui arrive en $n^{\text{ième}}$ position sera servi "environ" en $n^{\text{ième}}$ position (avec des tolérances exprimées par un intervalle flou de \mathbb{N}). On sera amené à répercuter l'imprécision de la discipline de service sur les performances calculées par une approche classique.

Une direction de recherche qui peut présenter un certain intérêt consisterait à étudier des processus stochastiques où les probabilités de transition sont mal connues, décrites par des intervalles (floues ou non). Ce type de problème a reçu une première analyse dans un article récent de Stanford /44/. Une approche possible de ce problème consiste à supposer que les bornes supérieures et inférieures des valeurs de probabilité sont des probabilités supérieures et inférieures au sens de Dempster (cf. chap.1). C'est l'optique adoptée dans un article de Shafer et Breipohl /41/, à propos d'une étude de fiabilité d'un système informatique.

REFERENCES

- /1/ BELLMAN R.E., ZADEH L.A. (1970)
Decision-making in a fuzzy environment, Management Sciences 17, B141-B154.
- /2/ BUCKLEY J.J. (1983)
Fuzzy programming and the Pareto-optimal set, Fuzzy Sets and Systems, 10, 57-64.
- /3/ CHANAS S., KAMBUROVSKI (1981)
The use of fuzzy variables in PERT, Fuzzy Sets and Systems, 5, 11-20
- /4/ CHARNES A., COOPER W. (1959)
Chance-constrained programming, Management Sciences 6, 73-79
- /5/ CHARNES A., COOPER W. (1961)
Management models and the industrial applications of linear programming, J. Wiley (2 vol.)
- /6/ DUBOIS D. (1977)
Quelques outils méthodologiques pour la conception de réseaux de transport, Thèse Docteur-Ingénieur, ENSAE, Toulouse.
- /7/ DUBOIS D. (1978)
An application of fuzzy sets to bus transportation network modification. Proc. Joint Automatic Control Conference, Philadelphie, Vol.3, 53-60.
- /8/ DUBOIS D. (1980)
Uncertainty in transportation network design : some new-techniques, IV^e European Conference on Operations Research, Cambridge, U.K.
- /9/ DUBOIS D. (1981)
Détermination des conditions de coupe sur une chaîne automatisée d'usinage, Rapport CERT/DERA n°027 (1/7610), Toulouse

- /10/ DUBOIS D. (1981)
A fuzzy set-based method for the optimization of machining operations, Proc. International Conference on Cybernetics and Society, Atlanta.
- /11/ DUBOIS D. (1983)
A fuzzy, heuristic, interaction method for the optimal Network problem, à paraître dans P.P. Wang (Ed.) = Advances in Fuzzy Set Theory and Applications, Plenum Press
- /12/ DUBOIS D., PRADE H. (1978)
Algorithmes de plus courts chemins pour traiter des données floues. RAIRO, Recherche Opérationnelle 12, 213-227.
- /13/ DUBOIS D., PRADE H. (1980)
Fuzzy Sets and Systems : theory and applications. Academic Press. New York
- /14/ DUBOIS D., PRADE H. (1980)
Systems of linear fuzzy constraints. Fuzzy Sets and Systems 3, 37-48
- /15/ DUBOIS D., PRADE H. (1982)
The advantages of fuzzy approach in OR/MS demonstrated on two examples of resource allocation problems in Progress in Cybernetics and Systems Research (R. Trappl, G. Klir, F. Pichler, Eds). Vol.VIII, pp.491-498.
- /16/ DYSON R.G. (1980)
Maximin-Programming, fuzzy linear programming and multicriteria decision making. J. Opnal. Res. Soc. 31, 361-367.
- /17/ ERSCHLER J., ROUBELLAT F., VERNHES J.P. (1976)
Finding some essential characteristics of the feasible solutions for a scheduling problem. Opns. Res. 24, n°4
- /18/ FLACHS J., POLLATSCHETK M.A. (1979)
Further results on fuzzy mathematical programming, Information and Control. 38, 241-257
- /19/ FLOYD R.W. (1962)
Algorithm 97, Shortest path. Comm. ACM, 5, p.345.

- /20/ GONDRAN M., MINOUX M. (1979)
Graphes et algorithmes, Eyrolles, Paris.
- /21/ HAMACHER H., LEBERLING H., ZIMMERMANN H.J. (1978)
Sensitivity analysis in fuzzy linear programming. Fuzzy Sets and Systems 1, 269-282.
- /22/ HANNAN E.L. (1979)
On the efficiency of the product operator in fuzzy programming with multiple objectives. Fuzzy Sets and Systems 2, 259-262.
- /22b/ KABBARA G. (1979)
Optimisation floue. Thèse, Université de Provence, Marseille.
- /23/ KAMBUROVSKI J. (1979)
Critical Path analysis in a project with fuzzy estimates of activity duration times, Instytut Organizacji, i Zarządzania, Politechnica Wroclawska, Pologne
- /24/ KLEINROCK J. (1975)
Queuing Systems, Wiley Interscience.
- /25/ LEBERLING H. (1981)
On finding compromise solution in multicriteria problems using the fuzzy min operator. Fuzzy Sets and Systems 6, 105-118
- /26/ LUHANJULA M.K. (1982)
Compensatory operators in fuzzy linear programming with multiple objectives. Fuzzy Sets and Systems 8, 245-252.
- /27/ LUHANJULA M.K. (1983)
Linear programming under randomness and fuzziness. Fuzzy Sets and Systems 10, 45-55
- /28/ MALCOLM D.G., ROSEBOOM J.H., CLARCK C.E., FAZAR W. (1959)
Application of a technique for research and development program evaluation, Operations Research 7
- /29/ NEGOITA C.V. (1978)
Management applications of systems theory, Birkhäuser Verlag, Basel.
- /30/ OH'EIGEARTAIGH M. (1982)
A fuzzy transportation algorithm. Fuzzy Sets and Systems 8, 235-243.

- /31/ OH'EIGEARTAIGH M. (1982)
Fuzzy linear programming. Non publié
- /32/ ORLOVSKI S. (1977)
On programming with fuzzy constraint sets. Kybernetes 6, 197-201
- /33/ PRADE H. (1977)
Ordonnancement et temps réel. Thèse Docteur-Ingénieur, ENSAE Toulouse.
- /34/ PRADE H. (1977)
Exemple d'approche heuristique, interactive, floue pour un problème d'ordonnancement, Congrès AFCEP Modélisation et Maîtrise des Systèmes, Versailles. Volume 2. Ed. Hommes et Techniques 347-345.
- /35/ PRADE H. (1979)
Using fuzzy set theory in a scheduling problem : a case study, Fuzzy Sets and Systems, 2, 153-165.
- /36/ PRADE H. (1980)
An outline of fuzzy or possibilistic models for queuing systems in Fuzzy Sets and their Applications Policy Analysis and Information System (P.P. Wang, S.K. Chang, Eds) Plenum press. 147-154
- /37/ PRADE H. (1980)
Operations research with fuzzy data, idibem, 155-170.
- /38/ ROY B. (1970)
Algèbre moderne et théorie des graphes, Vol.2, Dunod, Paris.
- /39/ SAKAI Y., TANAKA Y., IDO M. (1981)
Stochastic and fuzzy decision problems in three-workpiece lopping process. Trans. ASME, Jr. of Engineering for industry 103, 192-196.
- /40/ SCULLI D. (1983)
The completion time of PERT networks, J. Opnal Res. Soc. 34, 155-158.

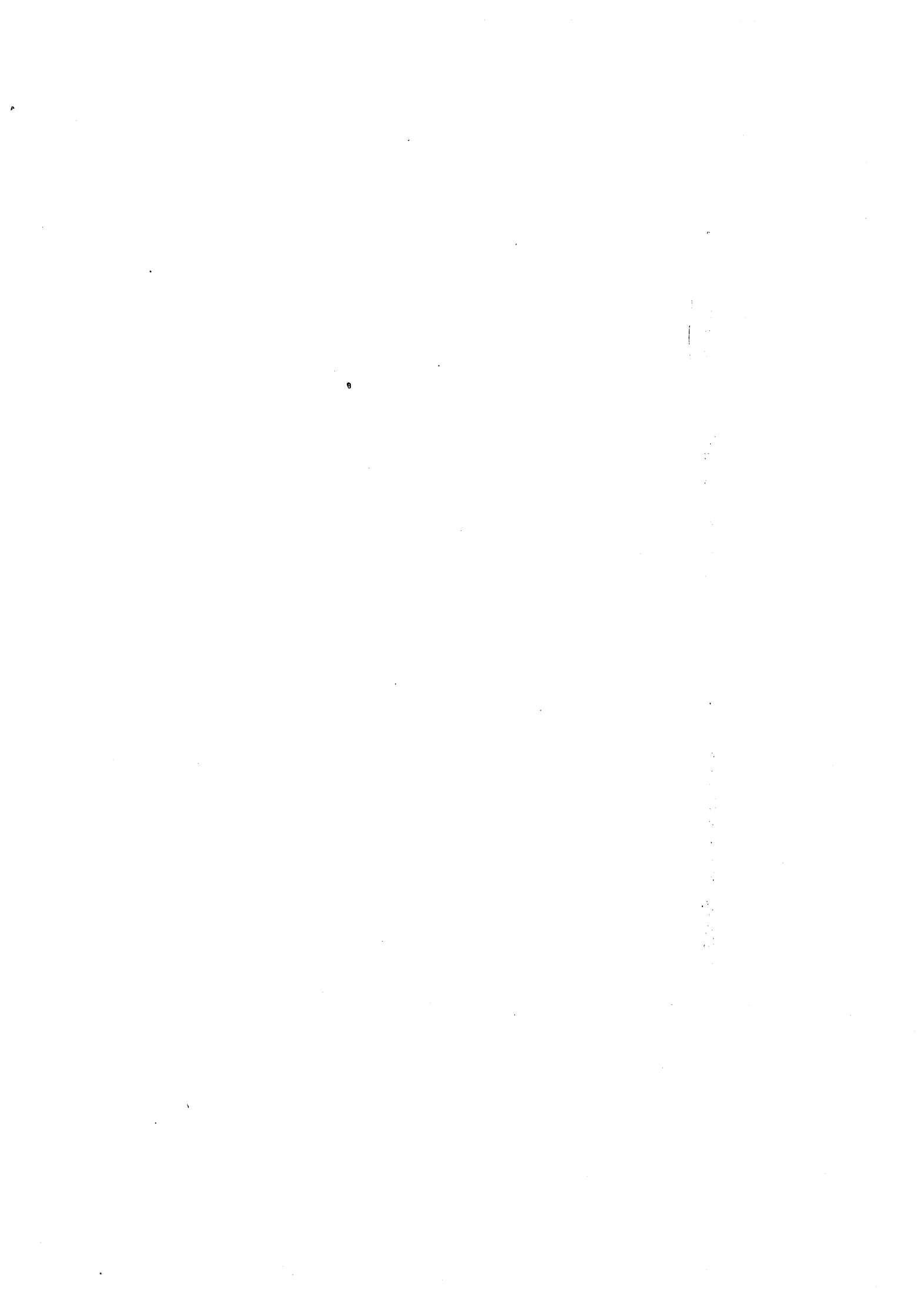
- /41/ SHAFER G., BREIPOHL A.M. (1979)
Reliability described by belief functions. Proc. Annual Reliability and Maintainability Symposium, 23-27
- /42/ SIGAL C.E., PRITSKER A.A.B., SOLBERG J.J. (1980)
The stochastic shortest route problem, Operations Research 28, 1122-1129.
- /43/ SOYSTER A.L. (1973)
Convex programming with set-inclusive constraints. Application to inexact linear programming. Operations Research 21, 1154-1157.
- /44/ STANFORD R.E. (1982)
The set of limiting distributions for a Markov chain with fuzzy transition probabilities, Fuzzy Sets and Systems 7, 71-78
- /45/ STEUER R. (1981)
Algorithms for linear programming problems with interval objective function coefficients. Maths. of O.R. 6, 333-348.
- /46/ TANAKA H., OKUDA T. ASAI K. (1973)
On fuzzy mathematical programming. J. Cybernetics 3, N°4, 37-46.
- /47/ TANAKA H., ASAI K. (1980)
Fuzzy linear programming based on fuzzy functions. Bull. Univ. Osaka Prefecture 29, N°2
- /48/ VENTURA E. (1966)
L'introduction de l'aléatoire dans les réseaux PERT, RFRO, 38, 55-68.
- /49/ VERDEGAY J.L. (1982)
Fuzzy mathematical programming in Fuzzy Information and Decision Processes (M.M. Gupta, E. Sanchez, Eds). North Holland 231-238.
- /50/ WIEDEY G., ZIMMERMANN H.J. (1978)
Media selection and fuzzy linear programming, J. Opnal. Res. Soc. 29, 1071-1084.
- /51/ YAGER R. (1980)
Some observations on probabilistic qualification in approximate reasoning. Information Sciences 22, 217-234.

/52/ ZIMMERMANN (1978)

Fuzzy programming and linear programming with several objective functions, Fuzzy Sets and Systems 1, 45-55

ANNEXE

RÉSULTATS DE THÉORIE DES EQUATIONS FONCTIONNELLES



Cette annexe décrit un certain nombre de théorèmes déjà connus en théorie des équations fonctionnelles, qui s'avèrent très utiles pour définir des mesures d'incertitude ou opérations ensemblistes généralisées. L'ouvrage de base sur le sujet est dû à Aczél /1/. Néanmoins un certain nombre de contributions récentes, développées le plus souvent indépendamment de la théorie de flou, sont particulièrement riches d'enseignement.

1 - EQUATIONS DE L'ASSOCIATIVITÉ ET DE LA BISYMETRIE

Cette équation a été résolue pour la première fois par Abel au début du 19^{ème} siècle. Aczél prouve le résultat suivant /1, p.256/ :

Soit F une fonction réelle de deux variables réelles telle que

$$A1 \quad F(t,u) = F(t,v) \quad \text{ou} \quad F(u,w) = F(v,w)$$

seulement si $u = v$ (Réductibilité)

$$A2 \quad F \text{ prend ses valeurs sur un intervalle (fini ou non)}$$

$$A3 \quad F(F(x,y),z) = F(x,F(y,z)) \quad (\text{Associativité})$$

alors il existe une fonction réelle f d'une variable réelle, continue et strictement monotone telle que

$$F(x,y) = f^{-1}(f(x) + f(y)) \quad (1)$$

En particulier, une conséquence de A1-A3 est la commutativité de F . f est dit "générateur additif" de F . Si f est un générateur additif de F alors $\forall \lambda \neq 0$ λf est aussi un générateur additif de F . Si $e \in \mathbb{R}$ existe tel que $f(e) = 0$ alors on voit que e est élément neutre de F , c'est-à-dire $F(x,e) = F(e,x) = x$. e est forcément unique.

Il y a des fonctions F qui ne se mettent pas sous la forme (1) mais qui sont associatives : $F(x,y) = \min(x,y)$ ou $\max(x,y)$ par exemple.

On rappelle qu'un certain nombre de fonctions réelles strictement monotones ne peuvent pas être associatives. En particulier (cf. Dubois et Prade /4/) dès que $F(x,y) \in [\min(x,y), \max(x,y)]$, $F \neq \min$, $F \neq \max$.

Dans ce cas on remplace la propriété d'associativité par celle de bisymétrie

$$F(F(x,y), F(u,v)) = F(F(x,u), F(y,v)) \quad (2)$$

Il est facile de voir que l'associativité implique la bisymétrie, sous les axiomes A1-A3. Aczél /1,p.281/ montre que la solution générale des équations fonctionnelles définies par les axiomes suivants :

B1 F est continue, définie sur un intervalle fermé borné

B2 F est réductible

B3 F est bisymétrique

B4 F est commutative $F(x,y) = F(y,x)$

B5 F est idempotente $F(x,x) = x$

est :

$$F(x,y) = f^{-1}\left(\frac{f(x)+f(y)}{2}\right) \quad (3)$$

où f est une fonction continue et strictement croissante.

Il est facile de voir que $\forall a, \forall b \neq 0, a + bf$ génère F par (3).

Si on enlève la condition de commutativité alors la solution générale devient

$$F(x,y) = f^{-1}(qf(x)+(1-q)f(y)) \quad q \in [0,1] \quad (4)$$

Dans la suite on particularise l'équation d'associativité à des fonctions de $[0,1]^2$ dans $[0,1]$, mais en omettant l'axiome A1.

2 - NORMES TRIANGULAIRES

Une norme triangulaire T est une application à valeur réelle, définie sur $[0,1] \times [0,1]$ qui satisfait les conditions suivantes :

T1 : $T(0,0) = 0, T(a,1) = T(1,a) = a$ (conditions aux limites)

T2 : $T(a,b) \leq T(c,d)$ si $a \leq c$ et $b \leq d$ (monotonie)

T3 : $T(a,b) = T(b,a)$ (symétrie)

T4 : $T(a,T(b,c)) = T(T(a,b),c)$ (associativité)

Le nom de 'norme triangulaire' (en anglais 'triangular norm') provient du fait que de telles fonctions apparaissent, lors de l'étude des espaces métriques aléatoires dans l'inégalité triangulaire de Menger (/9/)

$$F_{pr}(x+y) \geq T(F_{pq}(x), F_{qr}(y))$$

où F_{pr} , F_{pq} , F_{qr} sont des fonctions de répartition de variables aléatoires représentant respectivement les distances entre les points p et r , p et q , q et r . Il est clair qu'une norme triangulaire n'est pas une norme au sens mathématique habituel.

T confère donc à l'intervalle réel $[0,1]$ une structure de semi-groupe commutatif avec 1 pour élément neutre.

Le caractère général des propriétés T1 à T4 fait que les normes triangulaires sont susceptibles de multiples applications en dehors des espaces métriques aléatoires.

Quatre normes triangulaires sont particulièrement importantes :

- le minimum : $\min(a,b) = \begin{cases} a & \text{si } a \leq b \\ b & \text{sinon} \end{cases}$
- le produit : $\text{prod}(a,b) = a.b$
- l'opération ' T_m ' : $T_m(a,b) = \max(0, a+b-1) = \begin{cases} a+b-1 & \text{si } a+b > 1 \\ 0 & \text{sinon} \end{cases}$
- l'opération ' T_w ' : $T_w(a,b) = \begin{cases} a & \text{si } b = 1 \\ b & \text{si } a = 1 \\ 0 & \text{sinon} \end{cases}$

Entre ces fonctions tiennent les inégalités suivantes :

$$T_w(a,b) \leq \max(0, a+b-1) \leq a.b \leq \min(a,b) \quad (4)$$

De plus, on montre que pour toute norme triangulaire T on a /10/ :

$$T_w(a,b) \leq T(a,b) \leq \min(a,b) \quad (5)$$

Une famille importante de normes triangulaires a pu être caractérisée à l'aide de fonctions d'une seule variable. Plus précisément, Ling /8/ a montré le résultat suivant :

Pour toute norme triangulaire continue qui satisfait de plus la propriété archimédienne

$$T5 \quad T(a,a) < a, \quad a \in]0,1[$$

il existe une fonction f continue et strictement décroissante de $[0,1]$ dans $[0,+\infty)$ telle que

$$T(a,b) = f^*(f(a)+f(b)) \quad (6)$$

où f^* est la pseudo-inverse de f ; la pseudo-inverse d'une fonction strictement décroissante f de $[0,1]$ dans $[0,+\infty)$ est définie par

$$f^*(a) = \begin{cases} 1 & \text{si } a \in [0, f(1)] \\ f^{-1}(a) & \text{si } a \in [f(1), f(0)] \\ 0 & \text{si } a \in [f(0), +\infty) \end{cases} \quad (7)$$

On a alors

$$T(a,b,c) = T(a, T(b,c)) = T(T(a,b), c) = f^*(f(a)+f(b)+f(c)) \quad (8)$$

Notons que 'min' ne satisfait pas T5, d'ailleurs 'min' ne peut pas être mis sous la forme (6). f est la 'fonction génératrice' ou le 'générateur additif' de T et est définie à une constante multiplicative positive près.

Parmi les normes triangulaires continues et archimédiennes, deux sous-familles sont remarquables :

i) les normes triangulaires strictes qui satisfont les conditions T1, T3, T4, la condition plus stricte

$$T'2 \quad \begin{aligned} T(a,b) < T(c,b) & \quad \text{si } 0 < a < c \leq 1 \\ T(a,b) < T(a,d) & \quad \text{si } 0 < b < d \leq 1 \end{aligned}$$

au lieu de T2 et qui sont continues. Schweizer et Sklar /11/ ont établi que toute norme triangulaire stricte pouvait être mise sous la forme

$$T(a,b) = f^{-1}(f(a)+f(b)) \quad (9)$$

où le générateur additif f est tel que

$$f(1) = 0 \text{ et } \lim_{a \rightarrow 0^+} f(a) = +\infty$$

La norme triangulaire stricte typique est le produit. Il s'agit là, en fait, d'un cas particulier satisfaisant les axiomes A1-A2-A3.

En posant $h = e^{-f}$, on obtient une représentation de T sous forme multiplicative, (9) devient

$$T(a,b) = h^{-1}(h(a).h(b)) \quad (10)$$

où h est une fonction continue strictement croissante telle que $h(0) = 0$ et $h(1) = 1$.

ii) Les normes triangulaires dont les fonctions génératrices sont telles que $f(1) = 0$ et $f(0)$ est fini. Alors, pour toute suite $\{a_i\}_{i \in \mathbb{N}}$ de nombres de $]0,1[$, il existe un n_0 tel que

$$\sum_{i=1}^{n_0} f(a_i) > f(0) \text{ et donc } T(a_0, a_1, \dots, a_{n_0}) = f^*(\sum_{i=1}^{n_0} (f(a_i))) = 0.$$

Par suite, de telles normes triangulaires seront appelées nilpotentes.

La norme triangulaire nilpotente typique est l'opération T_m .

3 - CONORMES TRIANGULAIRES

Substituant à T_1 les conditions aux limites

$$S_1 : S(1,1) = 1 ; S(0,a) = S(a,0) = a.$$

et conservant les conditions T_2 , T_3 et T_4 , on définit la classe duale des normes triangulaires, celle des conormes triangulaires. Toute conorme S peut être générée à partir d'une norme T par la formule

$$S(a,b) = 1 - T(1-a, 1-b) \quad (11)$$

et réciproquement. Dans la formule (11) S et T sont dits être des opérations duales.

Ainsi aux quatre normes min, prod, T_m , T_w correspond respectivement les quatre conormes :

$$\text{- maximum : } \max(a,b) = \begin{cases} a & \text{si } a \geq b \\ b & \text{sinon} \end{cases}$$

$$\text{- 'somme probabiliste' : } sp(a,b) = a+b - a.b$$

$$\text{- 'somme bornée' : } sb(a,b) = \min(1,a+b)$$

$$\text{- opération } T_w^* : T_w^*(a,b) = \begin{cases} a & \text{si } b = 0 \\ b & \text{si } a = 0 \\ 1 & \text{sinon.} \end{cases}$$

On a

$$\max(a,b) \leq a+b-ab \leq \min(1,a+b) \leq T_w^*(a,b) \quad (12)$$

et pour toute conorme S,

$$\max(a,b) \leq S(a,b) \leq T_w^*(a,b) \quad (13)$$

Pour toute conorme triangulaire S continue qui satisfait de plus S5

$$S(a,a) > a, a \in]0,1[$$

il existe une fonction φ continue et strictement croissante de $[0,1]$ dans $[0,+\infty)$ telle que

$$S(a,b) = \varphi^*(\varphi(a)+\varphi(b)) \quad (14)$$

où la pseudo-inverse φ^* d'une fonction φ strictement croissante de $[0,1]$ dans $[0,+\infty)$ est définie par :

$$\varphi^*(a) = \begin{cases} 0 & \text{si } a \in [0, \varphi(0)] \\ \varphi^{-1}(a) & \text{si } a \in [\varphi(0), \varphi(1)] \\ 1 & \text{si } a \in [\varphi(1), +\infty) \end{cases} \quad (15)$$

Aux normes triangulaires strictes correspondent par (11) les conormes triangulaires strictes. Si f génère une norme stricte, alors

$\varphi(.) = f(1-.)$ génère la conorme associée $\varphi^{-1}(\varphi(.) + \varphi(.))$; $k = e^{-\varphi}$ est alors le générateur multiplicatif de cette conorme. Aux normes triangulaires nilpotentes correspondent les conormes triangulaires nilpotentes; étant donné une suite $\{a_i\}_{i \in \mathbb{N}}$, il existe alors n_0 tel que $S(a_0, a_1, \dots, a_{n_0}) = 1$. sp et sb sont respectivement les conormes triangulaires, stricte et nilpotente, typiques.

NB1 : Si on considère une opération associative sur $[0, 1]$ d'élément neutre $e \in]0, 1[$ strictement croissante et continue alors son générateur f vérifie

$$f(0) = +\infty \quad f(e) = 0 \quad f(1) = -\infty$$

et est défini à une constante multiplicative près.

On vérifie que e n'est effectivement neutre que sur $]0, 1[$ car

$$v(1, x) = 1, \quad v(0, x) = 0, \quad \forall x \in]0, 1[$$

Mais $v(0, 1)$ n'est pas défini. Ce type d'opération permet de définir des sommes symétriques (voir chap.2 § 3.4.2).

4 - NÉGATIONS FORTES

Trillas /14/, dans le cadre de l'étude des opérateurs de complémentation pour les ensembles flous, a proposé les axiomes suivants pour définir une opération de négation forte.

Une négation forte est une application C de $0, 1$ dans $0, 1$, telle que :

- C1 : $C(0) = 1$
- C2 : $C(C(a)) = a$ (involution)
- C3 : C est strictement décroissante
- C4 : C est continue.

Par exemple, $C(a) = 1 - a$ est une négation forte.

Notons immédiatement que si T est une norme triangulaire et C une négation forte, $S(.,.) = C [T(C(.),C(.))]$ est une conorme triangulaire et réciproquement $T(.,.) = C [S(C(.),C(.))]$. S et T seront alors C-duaux l'un de l'autre (Alsina, Trillas, Valverde /2/).

Trillas /14/ a montré que toute négation forte C pouvait être caractérisée, à l'aide d'une application continue et strictement croissante t de $[0,1]$ dans $[0,+\infty)$ telle que $t(0) = 0$ et $t(1)$ est fini, sous la forme

$$C(a) = t^{-1}(t(1) - t(a)) \quad (16)$$

Notons qu'il existe toujours un seuil $s \in]0,1[$ tel que $C(s) = s$, il vaut $s = t^{-1}(\frac{t(1)}{2})$. De plus si C est une négation forte, alors $1 - C(1-.)$ est aussi une négation forte, généralement distincte de C .

Négations fortes, normes et conormes triangulaires sont reliées entre eux, puisqu'il est manifeste qu'à partir d'un générateur de négation t , on peut obtenir un générateur additif de norme triangulaire nilpotente $f(.) = t(1) - t(.)$ et que t est le générateur additif d'une conorme nilpotente. Inversement, à partir d'un générateur de norme triangulaire nilpotente f , on peut construire un générateur de négation $t(.) = f(0) - f(.)$.

Ainsi, le générateur t d'une négation C peut également générer une norme triangulaire T et une conorme triangulaire S dont il est facile de vérifier qu'elle est la C -duale de T :

$$\begin{aligned} C(a) &= t^{-1}(t(1) - t(a)) \\ T(a,b) &= f^*(f(a)+f(b)) \quad \text{où } f = t(1) - t \\ S(a,b) &= t^*(t(a)+t(b)) \\ &= \min(1, t^{-1}(t(a)+t(b))) \quad \text{si } t \text{ est défini sur } [0,+\infty). \end{aligned} \quad (17)$$

Si C , T et S sont issues du même générateur t au sens de (14), il est aisé de vérifier qu'alors

$$T(a,C(a)) = 0 \quad ; \quad S(a,C(a)) = 1 \quad (18)$$

Réciproquement, pour toute norme (conorme resp.) triangulaire

nilpotente, il existe une négation forte et une conorme (norme resp.) triangulaire nilpotente telles que (15) tienne. Par exemple, $T(a,b) = \max(0, a+b-1)$, $S(a,b) = \min(1, a+b)$ et $C(a) = 1-a$ sont issues de $t(a) = a$.

NB2 : Si on considère une opération associative v définie au NB1, de générateur additif f , alors, il est clair que v est sa propre C-duale pour $C(a) = f^{-1}(f(x))$; cette équation définit bien des négations fortes.

5 - FAMILLES PARAMÉTRÉES D'OPÉRATIONS

A partir d'une famille paramétrée de fonctions pouvant servir de générateurs additifs pour des normes triangulaires, on peut construire une famille paramétrée \mathcal{C} de normes triangulaires et la famille duale (au sens de (8)) \mathcal{J} de conormes. Si les normes de \mathcal{C} sont nilpotentes, alors il existe une famille associée \mathcal{C}' de négations fortes et une famille associée \mathcal{J}' de conormes C-duales (au sens de (13) et (14)) ; de plus, une autre famille \mathcal{C}'' de négations et la famille associée de \mathcal{C}'' de normes C-duales (au sens de (13) et (14)) peuvent être construites directement à partir de \mathcal{J} . On a $\mathcal{C}' = \{1-C(1-), C \in \mathcal{C}\}$, comme on peut le vérifier aisément.

Dans ce qui suit, quelques familles paramétrées remarquables sont passées en revue. Elles ont été introduites par différents auteurs, souvent indépendamment les uns des autres et quelquefois même sans référence à la notion de norme triangulaire.

- L'opération T_p de Schweizer et Sklar /12/

$$T_p(a,b) = [\max(0, a^{-p} + b^{-p} - 1)]^{-1/p}, \quad p \in \mathbb{R} \quad (19)$$

avec la convention $T_0 = \text{prod.}$ Pour p fini, positif ou nul, T_p est une norme triangulaire stricte, pour p fini, négatif c'est une norme triangulaire nilpotente. De plus, $T_{-1} = T_m$, $T_{+\infty} = \min$, $T_{-\infty} = T_w$.

- L'opération de Sugeno /13/

$$S_\lambda(a,b) = \min(1, a+b+\lambda a \cdot b), \quad \lambda \geq -1 \quad (20)$$

Pour $\lambda > -1$, S_λ est une conorme triangulaire nilpotente.

$$S_{-1} = sp ; S_0 = sb ; S_\infty = T_w^*$$

- Le produit de Hamacher /7/

$$H_\gamma(a,b) = \frac{ab}{\gamma + (1-\gamma)(a+b-ab)} , \gamma \geq 0 \quad (21)$$

Pour γ fini, H_γ est une norme triangulaire stricte.

$H_0(a,b) = \frac{a \cdot b^\gamma}{a+b-ab}$; $H_1 = \text{prod}$; $H_{+\infty} = T_w$. Pour $\gamma = 2$, la conorme duale (au sens de (11)) est l'opérateur de Lorentz $\frac{a+b}{1+a \cdot b}$. Le produit de Hamacher est la seule norme triangulaire qui se mette sous la forme d'une fonction rationnelle.

- L'opération de Frank /6/

$$F_s(a,b) = \text{Log}_s \left(1 + \frac{(s^a - 1)(s^b - 1)}{s - 1} \right) , s > 0 \quad (22)$$

$$F_0(a,b) = \min(a,b)$$

$$F_1(a,b) = a \cdot b$$

$$F_\infty(a,b) = T_m(a,b) = \max(0, a+b-1)$$

Pour s fini strictement positif, F_s est une norme triangulaire stricte, de plus, les seules normes triangulaires continues T qui satisfont

$$T(a,b) + S(a,b) = a + b \quad (23)$$

où S est le dual de T sont les opérations F_x , $0 \leq s \leq +\infty$ et leurs sommes ordinales ($s > 0$) définies comme suit :

Soit $\{J_n\}_n$ une famille dénombrable de sous-intervalles fermés propres, adjacents, de $[0,1]$. A chaque $J_n = [p_n, q_n]$ est associée une norme triangulaire archimédienne T_n (i.e. qui satisfait T5). La norme triangulaire T définie par :

$$T(a,b) = \begin{cases} p_n + (a - p_n) T_n \left(\frac{a - p_n}{q_n - p_n}, \frac{b - p_n}{q_n - p_n} \right) & \text{si } (a,b) \in [p_n, q_n]^2 \\ \min(a,b) & \text{sinon.} \end{cases} \quad (24)$$

T est appelée somme ordinale des T_n . T est une norme triangulaire. Pour plus

de détails, voir référence /6/. La somme ordinaire de conormes satisfaisant S5 se définit de façon similaire changeant T_n en S_n et min en max dans (24).

- L'opération de Yager /15/

$$Y_q(a,b) = 1 - \min(1, [(1-a)^q + (1-b)^q]^{1/q}) \quad q \geq 0 \quad (25)$$

Pour $q > 0$, Y_q est une norme triangulaire nilpotente. $Y_0 = T_w$; $Y_1 = T_m$; $Y_{+\infty} = \min$.

- Un exemple d'opération non-archimédienne (Dubois, Prade, /5/)

$$\sigma_\alpha(a,b) = \frac{ab}{\max(a,b,\alpha)} \quad \alpha \in [0,1] \quad (26)$$

On voit que $\sigma_0 = \min$, $\sigma_1 = \text{produit}$.

σ_α ne peut pas être mis sous la forme (6), c'est une norme triangulaire qui n'a pas de générateur additif. σ_α est une somme ordinaire basée sur le produit.

Reprenant les notations introduites plus haut, on a $n = 1$, $J_1 = [0,\alpha]$, $T_1 = \text{prod}$. Alors, on a bien

$$\text{si } (a,b) \in [0,\alpha]^2, \sigma_\alpha(a,b) = \frac{ab}{\alpha} = 0 + (\alpha-0) \cdot \left(\frac{a-0}{\alpha-0}\right) \cdot \left(\frac{b-0}{\alpha-0}\right)$$

$$\text{sinon} \quad \sigma_\alpha(a,b) = \min(a,b)$$

Le dual de σ_α est une somme ordinaire basée sur la somme probabiliste.

- L'opération de Dombi /3/,

$$k_\lambda(a,b) = \frac{1}{1 + \sqrt[\lambda]{\left(\frac{1}{a} - 1\right)^\lambda + \left(\frac{1}{b} - 1\right)^\lambda}}$$

une t-norme stricte générée par $f(x) = \left(\frac{1}{x} - 1\right)^\lambda$, pour $\lambda \geq 1$. Pour $\lambda = 1$ on retrouve le produit de Hamacher $H_0(a,b)$, déjà rencontré plus haut. De plus $k_{+\infty} = \min$.

Le tableau suivant (Dubois, Prade /5/) récapitule les différentes familles de normes triangulaires dont il a été question, les conormes associées, ainsi que les générateurs additifs et les négations correspondantes s'il y a lieu. On pourra remarquer que la C-duale (duale resp.) de la norme de Schweizer et Sklar est la duale (C-duale resp.) de la norme proposée par Yager. Dans cet exemple, les deux négations associées s'échangent dans la transformation $C(.) \rightarrow 1-C(1-.)$.

6 - NORMES TRIANGULAIRES DISTRIBUTIVES

Une norme triangulaire T est dite distributive si et seulement si

$$T(a+b-T(a,b),c) = T(a,c) + T(b,c) - T(a,b,c) \quad (24)$$

La définition est la même pour les conormes.

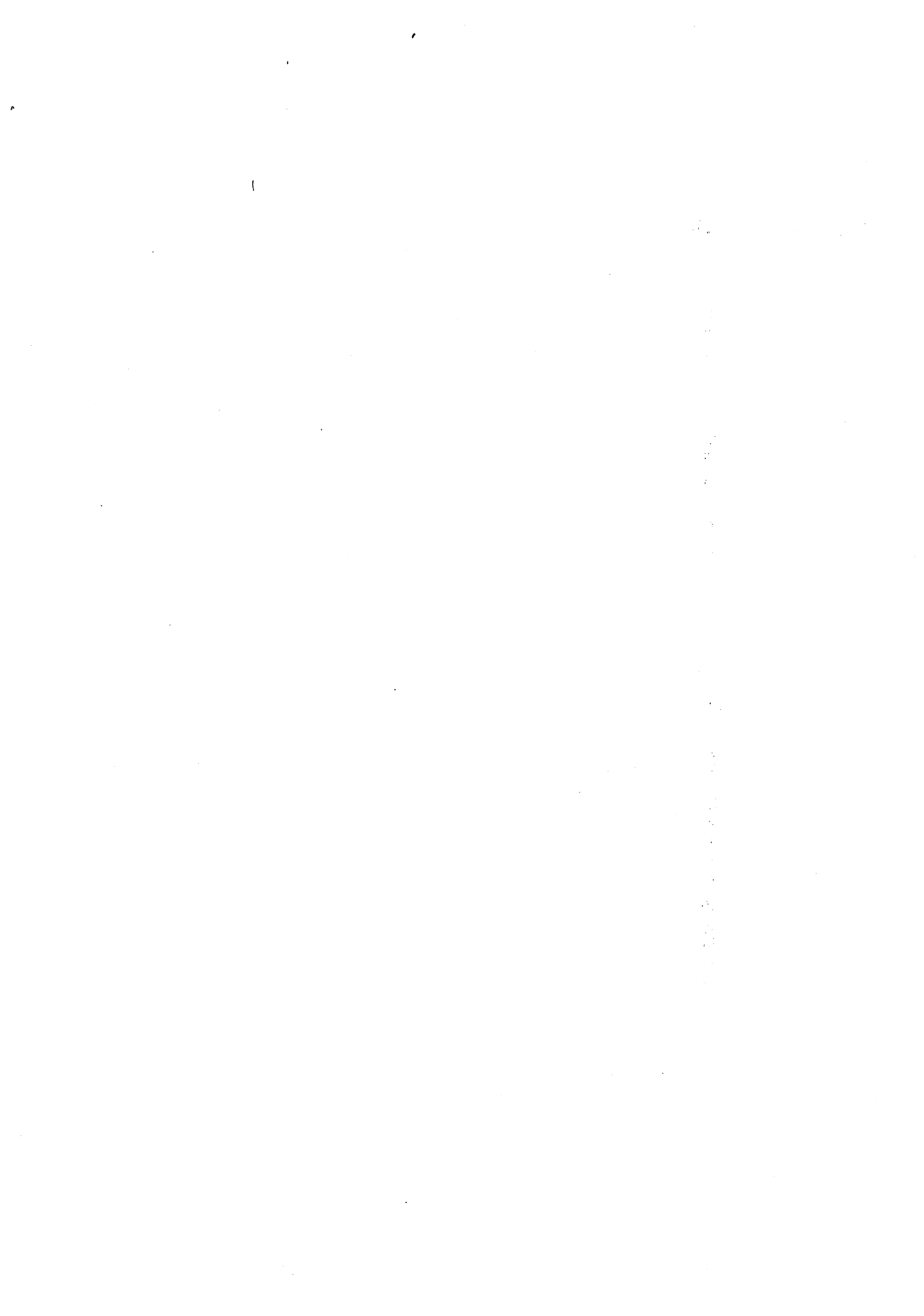
Frank /6/ a montré que les normes triangulaires continues distributives étaient, min, prod et les sommes ordinales basées sur prod (par exemple σ_α). Le résultat dual tient pour les conormes.

Références et notations	Norme	Conorme	Générateur additif de la norme	Négation	Norme C-duale de la conorme	Conorme C-duale de la norme
Hamacher (1975) H_Y	$\frac{ab}{\gamma + (1-\gamma)(a+b)ab}$ $\gamma \geq 0$	$\frac{a+b-(2-\gamma)ab}{1-(1-\gamma)ab}$	$f(a) = \text{Ln} \frac{\gamma - (\gamma-1)a}{a}$	norme (stricte)		
Sugeno (1974) S_λ	$\max(0, (\lambda+1) \cdot (a+b-1) - \lambda ab)$ $\lambda > -1$	$\min(1, a+b+\lambda ab)$ $\lambda > -1$	$\varphi(a) = \text{Ln}(1+\lambda a)$	$\frac{1-a}{1+\lambda a}$ (associée à la conorme)	$\max(0, \frac{a+b+\lambda ab-1}{\lambda+1})$	
Schweizer & Sklar (1963) T_p	$\max(0, a^{-p} + b^{-p} - 1)^{-1/p}$ $p \in \mathbb{R}$	$1 - \max(0, (1-a)^{-p} + (1-b)^{-p} - 1)^{-1/p}$	$f(a) = 1 - a^{-p}, p < 0$ $= a^{-p-1}, p > 0$	$\sqrt[p]{1-a^{-p}}$ (associée à la norme)		$\min(1, a^{-p} + b^{-p})^{-1/p}$
Yager (1980a) Y_q	$1 - \min(1, (1-a)^q + (1-b)^q)^{1/q}$ $q > 0$	$\min(1, a^q + b^q)^{1/q}$	$f(a) = (1-a)^q$	$1 - \sqrt[q]{1-(1-a)^q}$ (associée à la norme)		$1 - \max(0, (1-a)^q + (1-b)^q)^{1/q}$
Frank (1979) F_s	$\text{Log}_s \left[1 + \frac{(s^a-1)(s^b-1)}{s-1} \right]$ $s > 0$	$1 - \text{Log}_s \left[1 + \frac{(s^{1-a}-1)(s^{1-b}-1)}{s-1} \right]$	$f(a) = \text{Log}_s \frac{s-1}{s^{a-1}}$	norme (stricte)		
Dubois, Prade (1979) σ_α	$\frac{ab}{\max(a, b, \alpha)}$ $\alpha \in [0, 1]$	$\frac{a+b-\alpha - \min(a, b, 1-\alpha)}{\max(1-a, 1-b, \alpha)}$				
Dombi (1982)	$\frac{1}{1 + \sqrt[\lambda]{(\frac{1}{a}-1)^\lambda + (\frac{1}{b}-1)^\lambda}}$	$1 - \frac{1}{1 + \sqrt[\lambda]{(\frac{a}{1-a})^\lambda + (\frac{b}{1-b})^\lambda}}$	$f(a) = (\frac{1}{x} - 1)^\lambda$	norme (stricte)		

REFERENCES

-
- /1/ ACZEL J. (1966)
Lectures on functional equations and their applications.
Academic Press N.Y.
- /2/ ALSINA C., TRILLAS E., VALVERDE L. (1980)
Op. cit. cf. /1/ du chapitre 2
- /3/ DOMBI J. (1982)
Op. cit. cf. /7/ du chapitre 2.
- /4/ DUBOIS D., PRADE H. (1979)
Op. cit. cf. /10/ du chapitre 2
- /5/ DUBOIS D., PRADE H. (1982)
Op. cit. cf /10/ du chapitre 1
- /6/ FRANK M.J. (1979)
On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$
Aequationes Math. 19, 194-226.
- /7/ HAMACHER H. (1975)
Op. cit. cf. /27/ du chap. 2
- /8/ LING C.H. (1965)
Representation of associative functions. Publ. Math. Debrecen
12, 189-212
- /9/ MENGER K. (1942)
Statistical metrics. Proc. Nat. Acad. Sci. USA 28, 535-537
- /10/ SCHWEIZER B., SKLAR A. (1960)
Statistical metric spaces. Pacific J. Math. 10, 313-334
- /11/ SCHWEIZER B., SKLAR A. (1961)
Associative functions and statistical triangle inequities
Publ. Math. Debrecen 8, 169-186
- /12/ SCHWEIZER B., SKLAR A. (1963)
Associative functions and abstract semi-groups Publ. Math.
Debrecen 10, 69-81

- /13/ SUGENO M. (1974)
Op. cit. cf /30/ du chapitre 1
- /14/ TRILLAS E. (1979)
Op. cit. cf /74/ du chapitre 2
- /15/ YAGER R. (1980)
Op. cit. cf /79/ du chapitre 2



AUTORISATION DE SOUTENANCE

VU les dispositions de l'article 5 de l'arrêté du 16 Avril 1974,

VU les rapports de M. *Sakacvitch*

M. *NOVA*

M. *Bouscassan*

M. *Dubier* *DUBIER* est autorisé
à présenter une thèse en soutenance pour l'obtention du grade de
DOCTEUR D'ETAT ES SCIENCES.

Fait à GRENOBLE, le 20 OCT. 1983

Le Président de l'U.S.N.G.

E. Tanche

Le Président

E. TANCHE

Le Président de l'I.N.P.G.

D. BLOCH

Président
de l'Institut National Polytechnique
de Grenoble

P.O. le Vice-Président,

[Signature]