Chapitre 3

Opérateurs logiques

Afin de généraliser la logique booléenne où on ne compte que deux valeurs de vérité, plusieurs systèmes de logiques trivaluées ont été proposés à partir de 1920, considèrant le vrai (1 ou T), le faux (0, ⊥ ou F), et une troisième valeur généralement interprétée comme «l'indéterminé» (1/2 ou I) ou le «possible», par Kleene, Lukasiewicz et Przymusinski. Puis apparurent des systèmes avec un nombre fini de valeurs {0, 1/n, 2/n, ... , (n-1)/n, 1} et enfin des systèmes à valeurs continues dans [0, 1]. Nous étudions dans ce chapitre, différentes généralisations des opérateurs logiques booléens, ainsi que le problème de la déduction approchée et de l'agrégation avec ces formulations.

3.1. Les connecteurs logiques

LES PREMIERS SYSTÈMES DE LOGIQUE TRIVALUÉE

Pour {F, I, T}, désignant les valeurs logiques «faux», «indéterminé» et «vrai», également notées {0, 1/2, 1} plusieurs généralisations de la logique binaire ont été proposées. Dans tous les cas la négation échange T et F pour laisser fixe I.

La logique L3 de Lukasiewicz (1920)

$p \wedge q$

q \ p	0	1/2	1
1	0	1/2	1
1/2	0	1/2	1/2
0	0	0	0

$p \vee q$

q \ p	0	1/2	1
1	1	1	1
1/2	1/2	1/2	1
0	0	1/2	1

$p \rightarrow q$

q \ p	0	1/2	1
1	1	1	1
1/2	1	1	1/2
0	1	1/2	0

Cette logique, comme les suivantes, vérifie les lois de Morgan et définit l'équivalence par : $(p \leftrightarrow q) = (p \rightarrow q) \wedge (p \rightarrow q)$

Elle ne vérifie pas $(p \rightarrow q) = \neg p \vee q$. L'implication peut être rendue par la formule :

$(p \to q) = [\text{si } p \leq q \text{ alors } 1 \text{ sinon } 1 - p + q]$, qui permet la généralisation à un nombre fini de valeurs de vérité ou à toutes les valeurs de [0, 1].
La logique B3 de Bochvar (1939)

q	$p \wedge q$		
1	0	1/2	1
1/2	1/2	1/2	1/2
0	0	1/2	0
	0	1/2	1

q	$p \vee q$		
1	1	1/2	1
1/2	1/2	1/2	1/2
0	0	1/2	1
	0	1/2	1

q	$p \to q$		
1	1	1/2	1
1/2	1/2	1/2	1/2
0	1	1/2	0
	0	1/2	1

On peut définir :
$p \wedge q = [\text{si au moins un des deux vaut } 1/2 \text{ alors } 1/2, \text{ si } p = q = 1 \text{ alors } 1 \text{ sinon } 0]$.
En ce cas $(1/2) \wedge 0 = 1/2$ ce qui semble peu intuitif.
Quant à la disjonction :
$p \vee q = [\text{si au moins un est } 1/2 \text{ alors } 1/2, \text{ si } p = q = 0 \text{ alors } 0 \text{ sinon } 1]$.
L'implication vérifie $(p \to q) = \neg p \vee q$, elle peut être définie par :
$(p \to q) = [\text{si au moins un est } 1/2 \text{ alors } 1/2, \text{ si } p = 1 \text{ et } q = 0 \text{ alors } 0, \text{ sinon } 1]$.

La logique S3 de Sobocinski (1952). L'implication vérifie $(p \to q) = \neg p \vee q$.

q	$p \wedge q$		
1	0	1	1
1/2	0	1/2	1
0	0	0	0
	0	1/2	1

q	$p \vee q$		
1	1	1	1
1/2	0	1/2	1
0	0	0	1
	0	1/2	1

q	$p \to q$		
1	1	1	1
1/2	1	1/2	0
0	1	0	0
	0	1/2	1

Pour les logiques K3 de Kleene (1938) et H3 de Heyting (1956), $p \wedge q$ et $p \vee q$ sont définis comme dans L3.

q	K3	$p \to q$		
1	1	1	1	
1/2	1	1/2	1/2	
0	1	1/2	0	
	0	1/2	1	p

q	H3	$p \to q$		
1	1	1	1	
1/2	1	1	1/2	
0	1	0	0	
	0	1/2	1	p

q	$p \to q$		
1	1	1	1
1/2	1	0	0
0	1	0	0
	0	1/2	1

L'implication de Kleene vérifie $(p \to q) = \neg p \vee q$, celle de Heyting, n'ayant pas de symétrie par rapport à la seconde diagonale, ne le vérifie pas non plus que la contraposition $(\neg q \to \neg p) = (p \to q)$.
La définition de Gödel donnée par $(p \to q) = [\text{si } p \leq q \text{ alors } 1 \text{ sinon } q]$, généralise cette dernière implication.
A noter le dernier tableau, qui constitue une implication obtenue à partir de :
$p \wedge q$ = si l'un est faux alors 0 sinon 1, qui n'a que des résultats binaires.
Gödel a prouvé qu'aucune logique multivaluée finie ne pouvait engendrer une théorie (un calcul propositionnel) identique à celle du calcul propositionnel intuitioniste (annexe 2).
Par contre, ce dernier augmenté de l'axiome $(A \to B) \vee (B \to A)$ engendre la même théorie que la logique de Gödel ayant un nombre dénombrable de valeurs de vérité. En 1936, Jaskowski montre une logique multivaluée infinie engendrant la même

théorie que le calcul propositionnel intuitioniste, enfin Tarski a montré que H3 était équivalent à ce dernier augmenté de l'axiome : $(\neg A \rightarrow B) \rightarrow (((B \rightarrow A) \rightarrow B) \rightarrow B)$.
Les logiques trivaluées ou avec un nombre supérieur, mais fini de valeurs, ont été appliquées à des problèmes concrets, y compris grâce à des seuils, pour des circuits électroniques [Kameyana, Higushi 82].
Cependant l'intérêt pour ces logiques a chuté depuis les années 70 et 80. En logique «floue», on choisit habituellement : $p \wedge q = \min(p, q)$ et $p \vee q = \max(p, q)$ qui sont les généralisations de L3.

NÉGATION

Dans [0, 1] la négation utilisée est toujours le complément à 1 grâce à la simple formule $N(a) = 1 - a$, cependant toute application continue involutive et strictement décroissante telle que $N(0) = 1$ (négation forte) peut convenir comme définition.

Exercice 3.1
Montrer que les applications $N(x) = (1 - x^2)^{1/2}$ ou $N(x) = (1 - \sqrt{x})^2$.sont des négations. Le complément de Sugeno $N_\lambda(x) = (1 - x) / (1 + \lambda x)$ en est-il une ?

DÉFINITIONS DES T-NORMES ET T-CONORMES

Les t-normes [Menger 42] étendent la conjonction ordinaire. Les t-normes T sont les applications commutatives, associatives, croissantes de $[0, 1]^2$ dans [0, 1] (telles que $a \leq c$ et $b \leq d$ entraînent $T(a, b) \leq T(c, d)$) vérifiant $T(0, 0) = 0$ et où 1 est neutre, pour tout a $T(a, 1) = a$ (la stricte monotonie n'est pas exigée par la définition).
Une t-norme est dite diagonale si et seulement si $T(x, x) \leq x$ pour $x < 1$.
Les plus connues sont :

$T_0(a, b) =$ si $\max(a, b) = 1$ alors $\min(a, b)$ sinon 0	(produit drastique de Weber)
$T_1(a, b) = \max(0, a + b - 1)$	(Lukasiewicz)
$T_{1,5}(a, b) = ab / (2 - a - b + ab)$	(Einstein)
$T_2(a, b) = ab$	(algébrique ou probabiliste)
$T_{2,5}(a, b) = ab / (a + b - ab)$	(Hamacher)
$T_3(a, b) = \min(a, b)$	(Zadeh)

Exercice 3.2
Vérifier $T_0 \leq T_1 \leq T_{1,5} \leq T_2 \leq T_{2,5} \leq T_3$ et que toute autre T-norme est comprise entre T_0 et T_3. Montrer que T_1 possède des diviseurs de 0 ainsi $T(1/2, 1/2) = 0$.

Exercice 3.3
On appelle générateur additif f une fonction numérique strictement décroissante telle que $f(0) = +\infty$ et vérifiant $f(1) = 0$.
Montrer que $T(a, b) = f^{-1}(f(a) + f(b))$ ramené à [0, 1] est une t-norme et l'exprimer pour les quatre fonctions $f(x) = 1 - x$, $-\ln(x)$, $1 / x - 1$, $\cot(\pi x/2)$.
Quelles sont ces t-normes dans les trois premiers cas ? (réponses Lukasiewicz, probabiliste, Hamacher).
Montrer qu'en général, T est continue et strictement monotone.

Soit $f(1) = 0$ et $f(x) = 2 - x$ pour $0 \leq x < 1$, prouver que T est la norme de Weber, en donnant cette fois pour pseudo-inverse $f^{-1}(y) = \sup\{t \in [0, 1] / f(t) > y\}$.

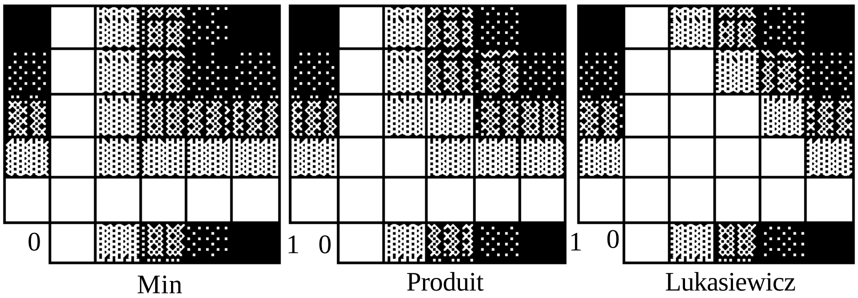

Figure 3.1 Illustration de trois conjonctions avec 5 valeurs, de 0 à 1 du blanc au noir.

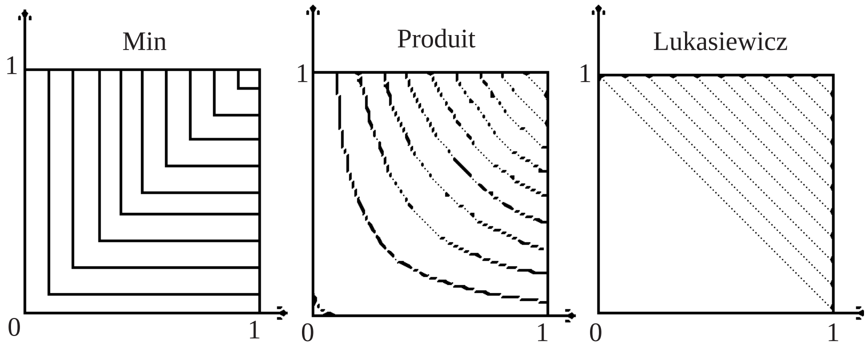

Figure 3.2 Plus finement dans $[0, 1]^2$, on peut tracer (ici 11) lignes de niveaux.

Exercice 3.4
Si f est une bijection continue croissante de [0, 1], montrer que $N(x) = f^{-1}(1 - f(x))$ est une négation et que l'on obtient une t-norme en définissant :

$$T(x, y) = [\text{si } f(x) + f(y) > 1 \text{ alors } f^{-1}(f(x) + f(y)) \text{ sinon } 0].$$

Les définitions de Sugeno $n(x) = (1 - x) / (1 + px)$ et de Yager $n(x) = (1 - x^p)^{1/p}$ peuvent-elles être des négations ?

Exercice 3.5
Si f est une bijection continue croissante de [0, 1], montrer que :
$T(x, y) = f^{-1}(f(x).f(y))$ est une t-norme.
Exprimer T pour $f(x) = (a^x - 1) / (x - 1)$, pour $4\text{atan}(x) / \pi$ et pour $x^2 / (2x^2 - 2x + 1)$..

Exercice 3.6
Une t-norme T est archimédienne si pour tous x, y dans]0, 1[, il existe une puissance n de x vis à vis de T, pour laquelle $x^{(n)} = T(x, ..., x) < y$. Montrer qu'alors pour tout x différent de 1 la limite de $x^{(n)}$ est 0 quand n tend vers l'infini. Montrer la réciproque.
T est régulière si $\forall x, y, z \;\; T(x, y) = T(x, z)$ et $x \neq 0 \Rightarrow y = z$.
Montrer que T est régulière si et seulement si elle est strictement monotone.

T est «diagonale» si $\forall x < 1$ $T(x, x) < x$
Montrer que «min» n'est ni archimédienne, ni régulière, ni diagonale, mais continue.
Montrer que la norme de Weber est archimédienne et diagonale mais ni continue ni régulière.
Montrer que si T est continue, la régularité entraîne que T est diagonale et archimédienne.
$T(x, y) =$ si $x, y \leq 1/2$ alors 0, sinon si $1/2 < x, y$ alors $2xy - x - y + 1$ sinon $\min(x, y)$ est la norme de Schweizer et Sklar. Montrer qu'elle est diagonale mais ni continue, ni archimédienne, ni régulière.

LES T-CONORMES

Ce sont les applications S commutatives associatives croissantes depuis $[0, 1]^2$ dans $[0, 1]$ (telles que $a \leq c$ et $b \leq d$ entraînent $S(a, b) \leq S(c, d)$), vérifiant $S(1, 1) = 1$ et telles que pour tout a, on ait $S(a, 0) = a$ (0 pour neutre). En notant $N(a) = 1 - a$ la négation, les deux notions sont évidemment duales ce qui signifie :
$S(a, b) = N(T(N(a), N(b)))$. On a donc les t-conormes associées :

$S_0(a, b) =$ si $\min(a, b) = 0$ alors $\max(a, b)$ sinon 1	(Weber)
$S_1(a, b) = \min(1, a + b)$	(Lukasiewicz)
$S_{1,5}(a, b) = (a + b) / (1 + ab)$	(Einstein)
$S_2(a, b) = a + b - ab$	(Probabiliste)
$S_{2,5}(a, b) = (a + b - 2ab) / (1 - ab)$	(Hamacher)
$S_3(a, b) = \max(a, b)$	(Zadeh)

On peut alors vérifier $S_3 \leq S_{2.5} \leq S_2 \leq S_{1.5} \leq S_1 \leq S_0$ et que toute autre t-conorme est comprise entre S_3 et S_0. Regardons par exemple simplement le cas où $a = 0$ et $0 < b$, alors $b \leq S(0, b) \leq b$ donc $S_3 \leq S \leq S_0$, par ailleurs si $0 < a \leq b$, $b = S(0, b) \leq S(a, b) \leq 1$ donc c'est encore vrai.
Voir des études systématiques des t-normes et t-conormes $[0, 1]^2 \to [0, 1]$ dans [Yager 80], [Dubois Prade 85], [Bonissone 87], [Mizumoto Masaharu 89].

ROBUSTESSE [Nguyen H.T. 96]

Pour une fonction f de $[0, 1]^n$ dans $[0, 1]$, on peut définir le risque en ∂ de f par :
$\Delta_f(\partial) = \sup \{\| f(x) - f(y) \| / |x - y| < \partial\}$, et on montre que parmi les t-normes (resp. t-conormes), la moins sensible est min (resp. max). Par contre, en se restreignant aux t-normes différentiables, c'est la t-norme probabiliste qui est la moins sensible.

Exercice 3.7
On dit qu'une t-conorme est diagonale si $x < S(x, x)$.
On dit que la fonction f est un générateur pour la t-conorme S, si f est continue et strictement croissante de $[0, 1]$ vers $[0, +\infty]$ et que $S(x, y) = f^{-1}(\min(f(1), f(x) + f(y)))$.
Montrer que si f(1) est fini, alors S est diagonale et T nilpotente c'est à dire qu'il existe une puissance (un nombre d'occurence de x) telle que $T(x, x, ...,x) = 0$

Exercice 3.8
On définit l'opération de Yager par $S_n(a, b) = \min(1, (a^n + b^n)^{1/n})$ pour $n > 0$, montrer que la limite de $S_n(a, b)$ lorsque n tend vers l'infini est max(a, b).
On prendra $\ln(S_n(a, b))$ avec $a < b$ et le min $\neq 1$.

Exercice 3.9
Que donne la famille de t-normes de Hamacher $T(x, y) = xy / [p + (1-p)(x + y - xy)]$ avec $p \geq 0$ si $p = 0$, si $p = 1$?, si p tend vers l'infini ? (réponses : Hamacher, produit, Weber). Exprimer la t-conorme associée S.

Exercice 3.10
Montrer que l'on a des t-normes pour les familles :
a) Dombi : $T(x, y) = 1/[1 + [(1/x - 1)^p + (1/y - 1)^p]^{1/p}]$ cas de $p = 1$?
b) Schweitzer : $T(x, y) = [\max(0, x^p + y^p - 1)]^{1/p}$. Montrer que les limites en $-\infty$, 0 et $+\infty$ donnent respectivement les normes min, probabiliste et drastique. Pour $p = 1$?
c) Franck : si $p > 0$ et $p \neq 1$ $T(x, y) = \log_p[1+ (p^x - 1)(p^y - 1) / (p - 1)]$
Quelles sont les limites de T lorsque p tend vers 0, 1, et $+\infty$?
(Réponses : Zadeh, produit, Lukasiewicz)
Exprimer S et vérifier que pour tous x, y : $T(x, y) + S(x, y) = x + y$
d) Weber : si $p \geq -1$ $T(x, y) = \max[0, (1 + p)(x + y - 1) - pxy]$.
Regarder les cas de $p = -1$ et $p = 0$.
e) Dubois : si $0 < p \leq 1$, $T(x, y) = xy / \max(x, y, p)$

REMARQUE, LIEN ENTRE T-NORME ET NORME

Le produit d'espaces normés est défini habituellement par $n(x, y) = \sup(n_1(x), n_2(y))$, cependant il est possible de généraliser par des définitions topologiquement équivalentes en dimension finie (théorème de Riesz).
Soient (E_1, n_1) et (E_2, n_2) deux espaces normés et S une application de R^{+2} dans R^+, en définissant n sur le produit E_1*E_2 par $n(x, y) = S(n_1(x), n_2(y))$ on peut montrer certains liens entre les propriétés de n et celle de S.
a) S est sous-additive pour l'addition des couples $\Leftrightarrow$ n vérifie l'inégalité triangulaire
b) Si S est croissante $[a < c$ et $b < d \Rightarrow S(a, b) \leq S(c, d)]$ alors :
$S(1, 1) = 1 \Leftrightarrow$ Le produit des boules unités de E_1 et E_2 est la boule unité de E_1*E_2
c) La norme n se projette suivant n_1 sur E_1, n_2 sur $E_2 \Leftrightarrow [\forall a\ \ S(a, 0) = S(0, a) = a]$
d) $[n(x, y) = 0 \Rightarrow x = y = 0] \Leftrightarrow [\forall a, b\ \ S(a, b) = 0 \Leftrightarrow a = 0 \text{ et } b = 0]$
S se comporte alors comme une disjonction.
Cependant le fait que n soit une norme n'est pas équivalent au fait que $S_{/[0, 1]^2}$ soit une conorme triangulaire. [Menger 42, Schweitzer 83]
En effet, si nous prenons par exemple dans R^2 la norme définie par :
$n(x, y) = |x| + |x + y|$, nous avons alors $S(1, 1) = 2$ et S non associative. En prenant $n(x, y) = |x| + 2|y|$, n est également une norme telle que S n'est pas commutative et $S(1, 1) = 3$. Enfin, réciproquement si S est la t-conorme algébrique définie par contre par $S(a, b) = a + b - ab$, alors n ne vérifie pas la propriété $[\forall\lambda, X\ \ n(\lambda X) = |\lambda|\, X]$, mais vérifie l'inégalité triangulaire.

Cependant, si nous considérons les trois normes les plus utilisées pour le produit d'espaces normés, nous retrouvons comme applications S associées :
- Conorme de Zadeh $S(a, b) = \max(a, b)$ pour la norme «sup»

- Conorme de Lukasiewicz S(a, b) = min (1, a + b); elle correspond à la norme de Hamming où N_1 est définie par $N_1(x, y) = n_1(x) + n_2(y)$
- Conorme de Yager $S(a, b) = (a^2 + b^2)^{1/2}$ correspond à la norme euclidienne.

La question du choix d'une norme triangulaire pour la conjonction des prémisses d'une règle floue, correspond donc partiellement à celle du choix de la norme dans R^n à propos de la définition des voisinages flous qui va suivre.

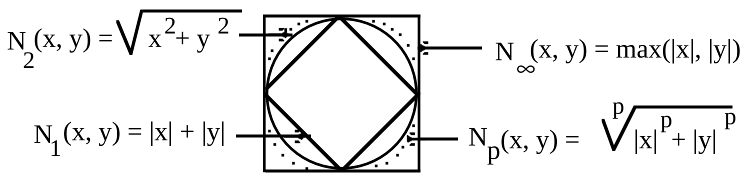

Figure 3.3 En dimension 2, la boule-unité pour ces trois normes plus la norme «p», on montre facilement que celle-ci tend vers la norme «sup» quand p tend vers l'infini.

Plaçons-nous dans $[-1, 1]^n$, et notons B(a, r) la «boule floue» ou «hypercône» centrée en $a \in [-1, 1]^n$ et de rayon $r \in R^+$ le sous-ensemble flou définie par la fonction d'appartenance : $\mu_{B(a, r)}(x) = \max(0, 1 - N(a - x) / r)$
Si T est une t-norme et N la norme associée (un des trois cas précédent), on aura donc

$$\mu_{B(a, r)}(x) = T(\mu_{B(a1, r)}(x_1), \dots, \mu_{B(an, r)}(x_n))$$

où $a = (a_1, \dots, a_n)$ et $x = (x_1, \dots, x_n)$.
Voyons à présent, en dimension deux l'image d'une boule floue : soit ZE (comme zéro), le prédicat défini sur [-1, 1] par la fonction d'appartenance $\mu(x) = \max(0, 1 - |x|)$, il est figuré comme coupes suivant x = 0 ou bien y = 0 de chacun des quatre figures suivantes. On définit, c'est la cas des systèmes de règles floues, la valeur de «x est ZE» et «y est ZE» par $u = T(\mu(x), \mu(y))$, ce qui donne plusieurs définitions possibles pour ZE^2 suivant les différentes t-normes employées.

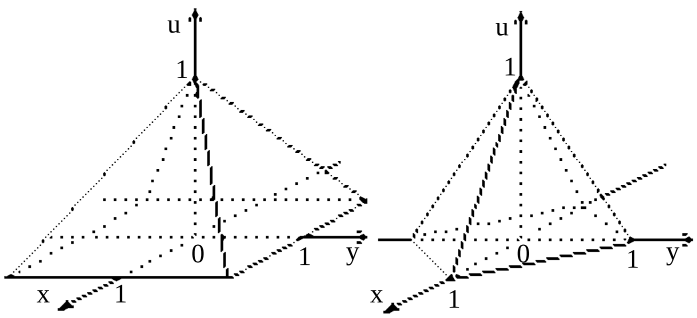

Figure 3.4 Boule-unité floue $T_{Zad}(x, y) = \min(x, y)$, $T_{Luk}(x, y) = \max(0, x + y - 1)$.

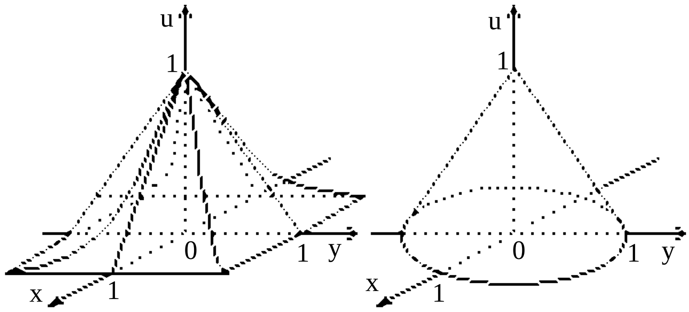

Figure 3.5 $T_{Yag}(x, y) = 1 - [(1 - x)^2 + (1 - y)^2]^{1/2}$, $T_{Proba}(x, y) = xy$.

3.2. Le problème de l'agrégation

Un opérateur d'agrégation en général, est une fonction H de R^n dans R croissante au sens large pour chacun de ses arguments, commutative et idempotente ce qui doit signifier pour tout nombre a, H(a, a, a, ..., a) = a. Cette définition valable pour des valeurs réelles exactes, doit être étendue à des ensembles flous, car dans la pratique le problème sera la plupart du temps de prendre une décision (c'est à dire déterminer une valeur exacte) rendant compte de plusieurs valeurs précises ou non, certaines ou non, bref, de réaliser un «consensus» entre valeurs floues.
Le problème est, soit de donner une valeur de vérité à un fait lorsqu'on en a démontré plusieurs, soit, plus généralement, de donner une assignation floue à une variable lorsqu'on en connaît plusieurs. Ce problème n'est résolu qu'empiriquement dans les applications, par divers opérateurs, mais aucune option générale ne peut convenir. Néanmoins, on trouve dans toutes les approches destinées à résoudre ce problème, l'idée de renforcement, c'est à dire une accentuation de la valeur la mieux représentée, comme on l'a vu pour l'opérateur de Dempster (chapitre 2).
Dans ce qui suit, si $x_1, \dots, x_n$ sont des valeurs pondérées par des poids $w_1, \dots, w_n$ dont la somme est $\sum w_i = 1$, on peut définir :

LES MOYENNES GÉNÉRALISÉES

$$H(x_1, \dots, x_n) = f^{-1}(\sum w_i f(x_i))$$

Rappelons que la moyenne arithmétique r de n valeurs $x_1, \dots, x_n$ est obtenue en prenant la fonction f = I, la moyenne géométrique g si f = ln (logarithme néperien), la moyenne harmonique h si f = «inverse», quadratique q si f = «carré», et que pour deux réels positifs a et b on a toujours les inégalités :

Avec $\frac{2}{h} = \frac{1}{a} + \frac{1}{b}$, on a : $h \leq g = \sqrt{ab} \leq r = \frac{a+b}{2} \leq q = \sqrt{\frac{a^2 + b^2}{2}}$

Prendre l'opération de moyenne géométrique si on veut maximiser, favorise les couples moyens au détriment des couples de plus forte étendue. Par exemple pour des candidats ayant la même moyenne arithmétique 0.5 sur deux notes, celui qui possède les notes 0.5 et 0.5 aura la moyenne géométrique 0.5 alors que celui qui a obtenu 0.4 et 0.6 aura 0.49, de plus le couple de notes 0.3 et 0.7 aura 0.46, et le couple 0.1 et 0.9 obtiendra 0.3. On voit donc que tous ces couples de notes de même moyenne arithmétique, ont une moyenne géométrique d'autant plus basse que leur écart est important.

MODÉLISATION ET AGRÉGATION DE PRÉFÉRENCES FLOUES [PERNY 92]

Dans cette thèse sont étudiées différentes méthodes d'agrégation pour exprimer une décision parmi des classes $\{C_1, \dots, C_m\}$ suivant des critères notés $v_1, \dots, v_n$ pondérés par des poids, on y trouve aussi l'idée d'accentuer les notes v_i par une fonction sigmoïde sur [0, 1]. A partir d'une relation binaire S sont définies des relations de :
Préférence $>_S(a, b)$ = si $S(a, b) > S(b, a)$ alors $\min(S(a, b), 1 - S(b, a))$ sinon 0
Indifférence $I_S(a, b) = \min(S(a, b), S(b, a))$
Incomparabilité $R_S(a, b) = \min(1 - S(a, b), 1 - S(b, a))$
Non-préférence $\sim_S(a, b) = \max(I_S(a, b), R_S(a, b))$

L'OPÉRATEUR Γ DE ZIMMERMANN-ZYSNO

$Z_\gamma(x, y) = (xy)^{1-\gamma}(x + y - xy)^\gamma$, réalise, pour $0 \le \gamma \le 1$, un intermédiaire entre $Z_0 = T_{prob}$ et $Z_1 = S_{prob}$, et il se généralise en une opération malheureusement non associative : $H(x_1, \dots, x_n) = (\prod x_i^{w_i})^{1-\gamma}(1 - \prod(1- x_i^{w_i}))^\gamma$ pour plusieurs variables. Néanmoins pour 1, on a, si on veut minimiser, une analogie avec la moyenne géométrique.

EXEMPLE DU PROBLÈME DE L'AGRÉGATION DANS LE DOMAINE DE L'ORDONANCEMENT DES TÂCHES [Felix 91]

Supposons par exemple que la qualité d'une tâche est l'agrégation de trois critères MDD (date exigée), PT (temps de travail) et RO (occupation des ressources) avec comme ordres de priorités 1, 0.5, 0.7. Chacun de ces trois critères est le $Z\gamma$ avec γ = 0.7 (opérateur de Zimmermann) de deux ou trois buts.
Toute action a est évaluée par $\sum_{(z \text{ fils de } a)} \text{priorité}(z) (\mu_{Uz}(a) - \mu_{Bz}(a))$ et on décide de l'action ayant l'évaluation maximale. Chaque but étant associé à deux ensembles flous, son support Uz, et son ensemble d'empêchement Bz, s'il n'y a que quelques actions à choisir, chaque ensemble est discret.
Les différents buts sont en outre reliés par 14 relations comme par exemple :
$\mu_{\text{indifférent}}(z_1, z_2) = \min [\mu_{\neg\text{inclu}} (U_1, U_2), \mu_{\neg\text{inclu}}(U_2, B_2),$
$\mu_{\neg\text{inclu}} (U_1, B_2), \mu_{\neg\text{inclu}} (U_1, B_2))]$
$\mu_{\text{presque coopérant}} (z_1, z_2) = \min(\mu_{\neg\text{inclu}} (U_1, B_2), \mu_{\neg\text{inclu}} (U_2, B_1)$
$\mu_{\text{entrave}} (z_1, z_2) = \min(\mu_{\neg\text{inclu}} (U_1, U_2), \mu_{\text{inclu}}(U_1, B_2),$ $\mu_{\subseteq}$ étant défini ici par card $(A \cap B)$ / card (A)
Dans ce système, seules les relations supérieures à 0.7 sont retenues.
Le problème de l'ordonnancement des tâches n'est pas facile à étendre pour des durées incertaines et des échéances imprécises. Ordinairement, si la tâche i possède la durée C_i avec la période T_i, le critère de charge à respecter est $\sum C_i / T_i \le 1$.

LES OWA (ordered weighted averaging)

Si W (w_1, ... , w_n) est un vecteur de poids de somme 1 et si y_i désigne le i-ième plus grand parmi les x_i rangés dans l'ordre décroissant, [Yager 88] définit OWA (x_1, ... , x_n) = $\sum w_i y_i$ Si chaque poids vaut 1/n on retrouve bien sûr la moyenne arithmétique. On a, si W = (1, 0 ... , 0), OWA = max, et si W = (0, ... , 0, 1) alors OWA = min. D'autre part, pour toute t-norme T et sa t-conorme associée S : T ≤ OWA ≤ S
Les OWA cherchent à considérer les meilleures valeurs parmi plusieurs. Ils formalisent plus précisément la notion de «la plupart des critères sont vérifiés», ces opérateurs peuvent donc servir à réaliser un consensus (suivant les poids donnés) plutôt des valeurs voisines, en éliminant les valeurs extrèmes.
Cependant, l'inconvénient, est qu'il faut donner un vecteur de poids, ceci peut être réalisé presque automatiquement par la méthode de [O'Hagan 90] où seul un coefficient «d'optimisme» α est demandé (plus il est proche de 1, plus les grandes valeurs seront privilégiées). La méthode consiste à maximiser $-\sum w_i \log(w_i)$ sous les contraintes $0 \leq w_i \leq 1$, $\sum w_i = 1$, $\sum w_i(n - i) / (n - 1) = \alpha$.

LES INTÉGRALES FLOUES DE CHOQUET OU SUGENO

Si c est une mesure sur l'ensemble des critères ou attributs {A_1, ... , A_n} et si f est la fonction discrète définie par les degrés de certitudes (x_1, ... , x_n) dans l'ordre croissant sur les A_i, alors :

$$H(x_1, ..., x_n) = \int_{Sug} f_o c = \max_{1 \leq i \leq n} \min(f(x_i), c\{x_i, x_{i+1}, ..., x_n\})$$

On a donc une généralisation au cas continu :

$$\int_{Sug} f_o c = \sup_{\alpha \in [0, 1]} \min(\alpha, c\{x / f(x) > \alpha\})$$

De plus on peut accorder des poids w_i aux critères A_i suivant lesquels on veut décider.

INTÉGRALE FLOUE DE SUGENO [Sugeno 74]

Si f est une application de X un espace mesuré avec la mesure c, dans [0, 1] où c est une «mesure floue» de confiance, on définit l'intégrale de Sugeno de f comme :

$$\int_{Sug} f_o c = \sup_{\alpha \in [0, 1]} \min(\alpha, c\{x / f(x) > \alpha\})$$

Si X est discret et si f est croissante, l'intégrale de Sugeno est alors calculée par :
$\max_{1 < i < n} \min(f(x_i), c(A_i))$ avec $A_i = \{x_i, x_{i+1}, ..., x_n\}$. Au cas où f est l'identité et c uniforme (c(I) est la longueur d'un intervalle I), alors l'intégrale donne le terme médian (le k-ième terme s'il y a 2k-1 termes discrets).
Si f n'est pas croissante, la définition est la même à condition de ranger préalablement les valeurs prises par f dans l'ordre croissant.
L'intégrale de Choquet [Choquet 53] consiste en une autre définition, qui, elle, coincide avec l'intégrale de Lebesgue dans le cas où c est additive, c'est à dire :

$$c(A \cup B) = c(A) + c(B).$$

$\int_{Cho} f_o c = \int_{[0, 1]} c\{x / f(x) > \alpha\} d\alpha = a_1 + (a_2 - a_1)c(A_2) + (a_3 - {}_2)c(A_3) + ...$ dans le cas discret.

REMARQUE Une mesure est dite décomposable s'il existe une t-conorme S telle que : $c(A \cup B) = S[c(A), c(B)]$, en particulier $c(A \cup B) = c(A) + c(B) + \lambda c(A)c(B)$ avec le coefficient $\lambda > -1$ qui sont les λ-mesures de Sugeno.

PROPRIÉTÉS

Les deux opérateurs sont croissants et continus mais non linéaires.
Si c est une mesure de possibilité (nécessité) et f la fonction d'appartenance à un ensemble flou, alors l'intégrale de Sugeno est la possibilité (nécessité) de cet ensemble flou.
Si f est constante égale à k, alors les deux intégrales Choquet et Sugeno ont bien la valeur k.
Si $c \leq c'$ alors $\int_{Sug} f_o c \leq \int_{Sug} f_o c'$ et $\int_{Cho} f_o c \leq \int_{Cho} f_o c'$
$\mathrm{Inf}_X (f) \leq \int f_o c \leq \sup_X (f)$ pour les deux définitions car on retrouve inf et sup pour deux mesures particulières que sont $c_0(\Omega) = 1$ et 0 ailleurs, et $c_1(\varnothing) = 0$ et 1 ailleurs.
Continuité pour les deux définitions : si $\lim (f_n) = f$ sur R, alors $\lim \int (f_n)_o c = \int f_o c$

Exercice 3.11
Calculer les deux intégrales pour $f(x) = x$ puis x^2 avec $c([a, b]) = (b-a)^2$ sur [0, 1]

APPLICATION À L'ANALYSE MULTICRITÈRE [Grabisch 93]

On évalue chaque objet, par exemple une voiture, par son degré de performance à l'égard d'un certain nombre de critères (prix, confort, consommation....) dont les attributs sont définis par des ensembles flous. Un utilisateur donnant ses poids à chacun de ces critères, on cherchera l'objet ayant la plus forte intégrale avec la mesure floue définie sur l'ensemble des critères.

APPLICATION À UNE MESURE DU FLOU [Benvenuti 94]

On peut définir A < B (A moins flou que B) par $\mu_A \leq \mu_B \leq 0.5$ ou $0.5 \leq \mu_B \leq \mu_A$, grâce à cette définition on peut montrer que c'est un ordre partiel et que les ensembles exacts sont minimaux et l'ensemble (noté E/2) défini par $\mu = 0.5$ est l'élément maximal unique. D'autre part le sup de deux ensembles existe toujours, mais pas l'inf.
En prenant une fonction f croissante sur [0, 0.5] décroissante sur [0.5, 1] et symétrique par rapport à 0.5, vérifiant $f(0.5) = 1$ et $f(0) = f(1) = 0$, un degré de flou peut aussi être défini par :

$$d(F) = \int_{Sug} f_o\, \mu_{F o} c = \sup_{\alpha \in [0, 1]} \min(\alpha, c\{x \,/\, f(\mu_F(x)) > \alpha\})$$

Exercice 3.12
Montrer que $d(F) = 0 \Leftrightarrow$ F ensemble exact, $d(E/2) = 1$, $d(\neg F) = d(F)$, et que d est croissante vis à vis de cette relation «moins flou que» : $A < B \Rightarrow d(A) < d(B)$

AGRÉGATION D'ENSEMBLES FLOUS

On peut utiliser une des méthodes ci-dessus appliquées aux valeurs des fonctions d'appartenance pour chaque x. Cependant si $h = \sup \min (\mu_A, \mu_B)$ est le degré de chevauchement de deux ensembles flous A et B, [Dubois, Prade 92, 94] proposent un opérateur d'agrégation donnant une distribution normalisée réalisant d'autant plus la conjonction que le conflit est faible (h proche de 1), et d'autant plus la disjonction que le conflit est fort (h proche de 0).

$$\mu_A * \mu_B = \max\ [\min(\mu_A, \mu_B) / h), \min((1 - h), \max(\mu_A, \mu_B))]$$

L'inconvénient est que cette formule n'est pas associative et surtout que l'ensemble flou produit peut ne pas être convexe, le pic à 1 peut fort bien être encadré de deux minima puis de deux plateaux plus élevés. Une extension de la formule à plusieurs distributions est néanmoins donnée de la façon suivante. Si $\mu_1, \mu_2, \ldots, \mu_r$ sont des ensembles flous normalisés sur un univers U, on pose m le nombre maximal de ces sources présentant une concordance totale, c'est à dire :
$m = \max \{card(I) / I \subset \{1, ..., r\}$ et $\sup(\inf\{\mu_i / i \in I\}) = 1\}$, et n le nombre maximal de sources présentant une certaine concordance (un mutuel chevauchement), soit $n = \max \{card(I) / I \subset \{1, ..., r\}$ et $\sup(\inf\{\mu_i / i \in I\}) > 0\}$.
On a donc $m \leq n$ et on pose :
$\pi_k(x) = \max \{\min \{\mu_i(x) / i \in I\} / card(I) = k$ et $I \subset \{1, ..., r\}\}$ la meilleure possibilité obtenue avec k sources sans les préciser au sein des $\mu_1, \mu_2, \ldots, \mu_r$.
Le plus grand degré de concordance établi sur n ensembles flous parmi $\mu_1, \mu_2, \ldots, \mu_r$, ayant une certaine concordance est :

$$h = \max \{\sup(\{\min \{\mu_i(x) / i \in I\} / x \in U\}) / card(I) = n \text{ et } I \subset \{1, ..., r\}\}.$$

La fonction agrégée des $\mu_1, \mu_2, \ldots, \mu_r$ sur l'univers U est alors :

$$Ag(x) = \max [(\pi_n(x) / h, \min((1 - h), \pi_m(x))].$$

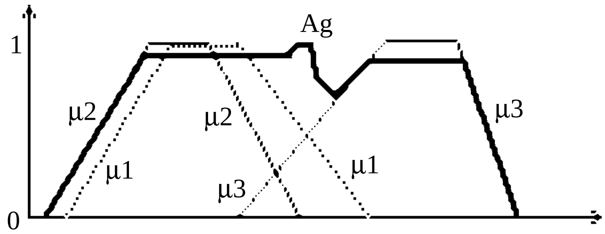

Figure 3.6 Agrégation Ag de deux ensembles flous suivant Dubois-Prade

Cette méthode est utilisée pour la fusion de capteurs donnant une évaluation de distance non fiable [Oussalah, Maaref, Barret 96].

Dans le même ordre d'idée, [Kelman, Yager 95], en considérant un opérateur d'agrégation H sur R^n, formulent l'agrégation des ensembles flous $A_1, \ldots, A_n$ par $H(A_1, \ldots, A_n)$ d'après une généralisation du principe d'extension de Zadeh, comme un ensemble flou B de fonction :

$$\mu_B(H(x_1, \ldots, x_n)) = \min(\mu_{A1}(x_1), \ldots, \mu_{An}(x_n), R(x_1, x_2), R(x_1, x_3), \ldots, R(x_{n-1}, x_n))$$

au moyen de R une «mesure de compatibilité» qui est une relation floue réflexive et symétrique et vérifiant [d(x, y) < d(x, z) alors R(x, y) > R(x, z)] si d est une distance sur l'univers considéré.
Au cas où R est la relation toujours égale à 1, on retrouve le principe d'extension.

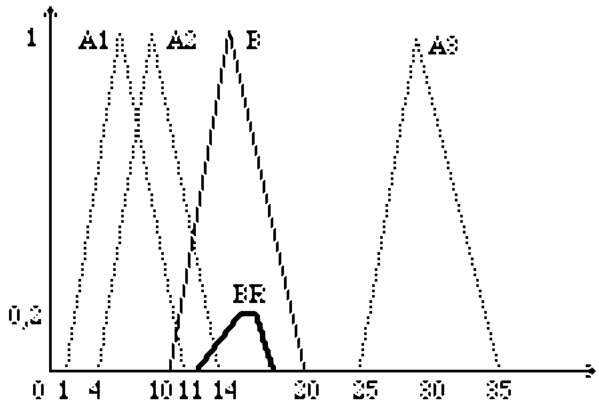

Figure 3.7 En prenant l'exemple de $R(x, y) = \max(0, 1 - 0.05|x - y|)$ dans l'intervalle [0, 40], on obtient un ensemble flou BR de plus grande valeur 0.2 suivant cette compatibilité, et une agrégation normalisée sans la considérer.

Exercice 3.13

On définit une somme symétrique comme une application de $[0, 1]^p$ vers [0, 1] continue et croissante pour chacun de ses arguments, commutative (ne dépendant pas de l'ordre de ses arguments) vérifiant :
$S(0, ..., 0) = 0$ et $S(1, ..., 1) = 1$. S est dite auto-duale si elle vérifie de plus :
$S(1 - x, 1 - y, 1 - z, ...) = 1 - S(x, y, z, ...)$. Trouver une telle fonction.

Signalons pour finir que pour les couples de [0, 1], cette fois, l'opérateur de Dempster réalise une sorte de «moyenne» qui en fait renforce les valeurs données, soit vers le vrai, soit vers le faux. Dempster est associatif mais non idempotent (voir chapitre 2).

3.3. Différentes formes de l'implication

En désignant par le symbole $\rightarrow$ l'implication booléenne classique, nous passons en revue les différentes généralisations, toutes notées de la même façon.
On peut définir une implication associée à une t-norme T par :
$(p \rightarrow q) = \sup \{r / T(p, r) \leq q\}$ ainsi Gödel-Sanchez, Goguen et Lukasiewicz sont-ils associés aux normes du même nom. Ces opérateurs vérifient pour la plupart la relation $p \leq q \Rightarrow (p \rightarrow q) = 1$ et aussi la définition classique $(p \rightarrow q) = \neg(T(p, \neg q))$

DÉFINITION DE REICHENBACH

$(p \rightarrow q) = 1 - p + p.q$ résulte de $(p \rightarrow q) = \neg p \vee q$ mais avec le produit pour la conjonction.

DÉFINITION DE WILLMOTT

$(p \rightarrow q) = \max(1 - p, \min(p, q))$
c'est une implication tempérée par $(p \vee \neg p) \geq 0.5$

DÉFINITION DE MAMDANI

$(p \rightarrow q) = \min(p, q)$ caractérisée par une symétrie, cette définition pose problème car, ainsi que Larsen $(p \rightarrow q) = p.q$, ce n'est pas une généralisation de l'implication booléenne.

DÉFINITION DE RESCHER ET DE PRZYMUSINSKI

$$(p \rightarrow q) = \text{si } p \leq q \text{ alors } 1 \text{ sinon } 0$$

DÉFINITION DE KLEENE ET DIENES

On généralise simplement la logique binaire classique où $p \rightarrow q$ est légitimé en disant que p ne peut être vrai sans que q le soit, en d'autres termes, il faut que $p \wedge \neg q$ soit faux, on a donc :
$(p \rightarrow q) = \neg(p \wedge \neg q) = \neg p \vee q = \max(1 - p , q)$ résulte de $(p \rightarrow q) = \neg p \vee q$ en logique min-max.

DÉFINITION DE GÖDEL (reprise par Brouwer)

$$(p \rightarrow q) = \text{si } p \leq q \text{ alors } 1 \text{ sinon } q$$

Elle est associée aux t-normes min ou de Weber pour :
$(p \rightarrow q) = \sup \{r / T(p, r) \leq q\}$.
Cette définition est utilisée dans le système-expert Diabeto [Buisson, Farreny, Prade 86], non pour donner une valeur de vérité à la proposition $(p \rightarrow q)$, mais pour donner une distribution de possibilité de q conditionné par p.

DÉFINITION PROPOSÉE PAR GOGUEN (probabiliste)

$(p \rightarrow q) = [\text{si } p = 0 \text{ alors } 1 \text{ sinon } \min(1 , q / p)] = [\text{si } p \leq q \text{ alors } 1 \text{ sinon } q / p]$.
Il est à noter que cette formule coïncide avec celle de Lukasiewicz, mais uniquement dans le cas de la logique à trois valeurs.
L'interêt de cette définition réside dans le fait que la résolution de l'équation en q :
$(p \rightarrow q) = r$ se fait facilement par : $q = \text{si } r < 1 \text{ alors } p.r \text{ sinon } [p, 1]$
Cette question est en effet importante en ce qui concerne le mécanisme d'inférence d'une règle $p \rightarrow q$ pondérée par un coefficient dans un système-expert.

DÉFINITION DE LUKASIEWICZ

Reprise par Zadeh et utilisée dans le système-expert Protis, est :

$$(p \rightarrow q) = \min(1 , 1 - p + q) = 1 - \max(0, p - q)$$

Cette formule a l'avantage de respecter la contraposition $(\neg q \rightarrow \neg p) = (p \rightarrow q)$
Dans un système multivalué comportant m valeurs $t_0, t_1, \ldots t_{m-1}$ elle se définit par :

$$t_i \rightarrow t_j = [\text{si } (i \leq j) \text{ alors } t_{(m-1)} \text{ sinon } t_{(m-i+j-1)}]$$

AUTRES DÉFINITIONS

On trouve encore $(p \rightarrow q) = (p \wedge q) \vee (1 - p)$ définition de Zadeh analogue à celle de Wilmott, et :
$(p \rightarrow q) = [(p \wedge q) \quad (1 - p)] \wedge [q \vee (1 - p)]$ [Turksen 88]
$(p \rightarrow q) = p.q$ (Opérateur de Larsen)

$(p \rightarrow q) = p(1 - |p - q|)$ [Vincent, Dujet 94] qui généralise Mamdani.
$(p \rightarrow q) = q^p$ généralise aussi l'implication booléenne avec la convention $0^0 = 1$.

REMARQUE Les définitions de Mamdani et Larsen ne sont pas à proprement parler des implications car elles ne généralisent pas l'implication booléenne.
Pour p = 0 et q = 1 on doit en effet avoir $(p \rightarrow q) = 1$ et non pas 0.

EXEMPLES D'IMPLICATIONS AVEC 3, 4, OU 5 VALEURS DE VÉRITÉ

	P	Q	KLEENE	LUKASIEWICZ	GODEL
3 valeurs	1	Q	Q	Q	Q
	1/2	1	1	1	1
	1/2	1/2	1/2	1	1
	1/2	0	1/2	1/2	0
	0	Q	1	1	1
4 valeurs	1	Q	Q	Q	Q
	2/3	1	1	1	1
	2/3	2/3	2/3	1	1
	2/3	1/3	1/3	2/3	1/3
	2/3	0	1/3	1/3	0
	1/3	1	1	1	1
	1/3	2/3	2/3	1	1
	1/3	1/3	2/3	1	1
	1/3	0	2/3	2/3	0
	0	Q	1	1	1
5 valeurs	1	Q	Q	Q	Q
	3/4	1	1	1	1
	3/4	3/4	3/4	1	1
	3/4	1/2	1/2	3/4	1/2
	3/4	1/4	1/4	1/2	1/4
	3/4	0	1/4	1/4	0
	1/2	1	1	1	1
	1/2	3/4	3/4	1	1
	1/2	1/2	1/2	1	1
	1/2	1/4	1/2	3/4	1/4
	1/2	0	1/2	1/2	0
	1/4	1	1	1	1
	1/4	3/4	3/4	1	1
	1/4	1/2	3/4	1	1
	1/4	1/4	3/4	1	1
	1/4	0	3/4	1	1
	0	Q	1	1	1

Exercice 3.14
Comme on le voit, le résultat peut être fort différent pour ces trois implications. Cependant on peut constater que l'implication de Lukasiewicz est toujours supérieure aux deux autres, montrer en étudiant les formules que cela est toujours vrai.

CONTRAPOSITION

On vérifie facilement que pour les fonctions booléennes qui définissent l'implication au sens de Kleene ou de Lukasiewiecz, l'égalité :

$(p \to q) = (\neg q \to \neg p)$ est alors assurée.

Par contre la définition de Gödel n'assure plus la contraposition car :

$(\neg q \to \neg p)$ = si $p \leq q$ alors 1 sinon $1 - p$

$(p \to q)$ = si $p \leq q$ alors 1 sinon $1 - q$

RAISONNEMENT POSSIBILISTE

Si on se tient seulement à la théorie des possibilités et si P est connu avec un intervalle [x, y] et une règle $P \to Q$ avec un degré de suffisance [s, ...], par modus ponens $nc(Q) \geq \min(x, s)$, par modus tollens $nc(\neg Q) = 1 - ps(Q)$ donc $ps(P) \leq \max(1 - s, ps(Q))$ et c'est tout ce qui peut être dit.

Si un degré de nécessité [n,...] est donné pour la règle inverse $Q \to P$ c'est à dire $\neg P \to \neg Q$, alors $nc(\neg Q) = 1 - ps(Q) \geq \min(nc(\neg P), n)$ soit $ps(Q) \leq S(y, 1-n)$

L'approche Bayésienne, préferée par certains systèmes, consiste à attribuer des coefficients [s, S] et [r, R] pour les règles respectives $P \to Q$ et $\neg P \to Q$ alors par Bayes : $nc(Q) = S(T(s, y), T(r, 1 - y)) \leq ps(Q) = S(T(S, y), T(R, 1 - x))$

En posant sous forme de possibilité conditionnelle $ps(p \text{ et } q) = \min(ps(p), ps(q / p))$ et $nc(q / p) = 1 - ps(\neg q / p)$ on obtient une écriture matricielle (mais avec les opérations min-max) :

$$\begin{bmatrix} ps(Q) \\ ps(\neg Q) \end{bmatrix} = \begin{bmatrix} ps(Q \mid P) & ps(Q \mid \neg P) \\ ps(\neg Q \mid P) & ps(\neg Q \mid \neg P) \end{bmatrix} \begin{bmatrix} ps(P) \\ ps(\neg P) \end{bmatrix}$$

C'est la relation de Bayes transposée aux coefficients de possibilités pour Q et pour $\neg Q$ qui s'écrit matriciellement [Farreny, Prade, Wyss 86].

LE PROBLÈME DE LA RÉSOLUTION $(p \to q) = r$

Aucune définition de $p \to q$ ne donne de réponse simple à la solution q de l'équation $(p \to q) = r$, seule la formule de Goguen peut donner une borne inférieure $q = p*r$ d'un calcul pratique, c'est ce qui a été utilisé dans certains systèmes-experts (coefficients de vraisemblance).

Une autre borne inférieure utilisée est $q = \min(p, r)$, et c'est avec la définition de Gödel que cette formule s'accorde le mieux.

LE POINT DE VUE APPORTÉ PAR LA THÉORIE DES RELATIONS FLOUES

Dans [Kaufmann 77] le calcul propositionnel est également introduit à partir de :

$\neg p = 1 - p$ $\quad$ $p \wedge q = \min(p, q)$ $\quad$ $p \vee q = \max(p, q)$

La recherche des «implications» a conduit à étudier les relations floues entre deux ensembles X, Y : si les propositions sont des sous-ensembles flous, donner une définition de $A \to B$ où A, B sont sous-ensembles flous respectivement de X, Y; équivaut à se donner une relation, dite de causalité, R permettant de trouver B connaissant A, c'est-à-dire : R(A) = B (voir aussi [Turksen 89])

C'est alors la définition de Gödel qui joue le plus grand rôle, en effet si on définit $\rightarrow$ l'opérateur de Sanchez-Gödel par : $(x \rightarrow y) = [\text{si } x \leq y \text{ alors } 1 \text{ sinon } y]$
la composition de deux relations floues S entre X, Y puis R entre Y, Z par :

$$\mu_{R \circ S}(x, z) = \max\{ \min(\mu_S(x, y), \mu_R(y, z)) / y \in Y\}$$

et enfin l'α-composition par $\mu_{S \rightarrow R}(x, z) = \min\{ \mu_S(x, y) \rightarrow \mu_R(y, z) / y \in Y\}$
on démontre les théorèmes suivants :
1° Si Q et T sont deux relations floues entre X et respectivement Y et Z, et s'il existe une solution R à l'équation $R \circ Q = T$ alors il existe une solution maximale de Y dans Z qui est $Q^{-1} \rightarrow T$
2° Si A, B sous-ensembles flous respectivement de X,Y et s'il existe une relation floue R telle que R(A) = B alors il existe une telle relation maximale qui est $(A \rightarrow B)$ où $\mu_{A \rightarrow B}(x,y) = (\mu_A(x) \rightarrow \mu_B(y))$
3° Si R et T sont deux relations floues entre respectivement Y et X , et Z , et s'il existe une solution Q à l'équation $R \circ Q = T$ alors il existe une solution maximale de X dans Y qui est $(R \rightarrow T^{-1})^{-1}$
C'est le deuxième résultat qui nous intéresse. Ainsi par exemple : si $X = \{x_1 \ldots x_5\}$ et $Y = \{y_1, y_2, y_3\}$, que la «cause» soit le sous-ensemble flou :
$A = \{ x_1|0.8 , x_2|1 , x_3|0 , x_4|0.3 , x_5|0.5 \}$, et $B = \{ y_1|0.9 , y_2|0.2 , y_3|0.7 \}$ l'effet, alors $R = (A \rightarrow B)$ est donné par le tableau :

α	y1	y2	y3
x1	1	0.2	0.7
x2	0.9	0.2	0.7
x3	1	1	1
x4	1	0.2	1
x5	1	0.2	1

On vérifie d'ailleurs que la matrice-ligne A multipliée (min-max) par la matrice ci-dessus, donne la matrice-ligne B.

UNE GÉNÉRALISATION DE L'ÉQUIVALENCE

En attachant à chaque règle $P \Rightarrow Q$ (éventuellement $A(x) \Rightarrow B(y)$) un couple (s, n) représentant cette fois les plausibilités de la règle et de la règle inverse on cherche la plausibilité pour un couple de valeurs précises (x, y) d'avoir B(y) conditionné par A(x), en d'autres termes pour une règle de la forme : si «X est A» alors$_{(s,n)}$ «Y est B».

On peut proposer [Gacôgne 90] : $\begin{bmatrix} ps(Q) \\ ps(\neg Q) \end{bmatrix} = \begin{bmatrix} 1 & 1-n \\ 1-s & 1 \end{bmatrix} \begin{bmatrix} ps(P) \\ ps(\neg P) \end{bmatrix}$

Ce qui conduit à :
$ps(Q) = \max(ps(P), 1 - \max(n, nc(P)))$ et $nc(Q) = \min(nc(P), \max(s, 1 - ps(P)))$.
Les différents systèmes existants utilisent des généralisations de l'implication de Kleene ou de Gödel en distinguant les règles suivant des conditions certaines, ou floues et une conclusion incertaine floue ou non floue [Buisson Farreny Prade 86]. Si on choisit comme dans [Martin-Clouaire, Prade 86], une définition qui tente une généralisation comme la plus simple fonction π donnant les quatre valeurs du tableau suivant dans lequel on compare avec l'implication $P \rightarrow Q$ classique.

p	q	$p \to q$	$p \leftarrow q$	$p \leftrightarrow q$	$p \leftrightarrow_{s,n} q$
1	1	1	1	1	1
1	0	0	1	0	1 - s
0	1	1	0	0	1 - n
0	0	1	1	1	1

On voit donc seulement avec ces quatre cas de la logique binaire classique, que plus s est grand et n petit, plus on se rapproche de l'implication et plus s et n sont grands, plus on se rapproche de l'équivalence.
Précisément, si s = 1 et n = 0 on a l'implication, si s = 0 et n = 1, on a l'implication inverse, et pour s = n = 1, on retrouve l'équivalence. Il s'agit donc d'une généralisation de l'équivalence pouvant trouver son utilité dans la mesure où le raisonnement dit de «sens commun» s'intéresse toujours lors d'une relation de cause à effet, à la relation inverse.

L'APPROCHE DES VOISINAGES [AKDAG 92]

Dans une logique multivaluée finie, on prend généralement 7 valeurs linguistiques : L_7 = {faux, peuvrai, assezvrai, moyen, plutôtvrai, trèsvrai, vrai}.
$L_n = \{t_0, ..., t_{n-1}\}$ est muni d'un ordre total $t_i < t_j$ si $i < j$, du min et max et de la négation $\neg t_i = t_{n-1-i}$.
L'implication choisie est celle de Luckasiewicz $(t_i \to t_j) = t_{[n-1-\max(0,\ i\ -\ j)]}$ = si $i \le j$ alors t_{n-1} sinon $t_{n-1-i+j}$
Grâce à une définition de voisinage $x\ V_i\ y \Leftrightarrow (x \to y) \ge i$ et $(y \to x) \ge i$, le t_k-voisinage de t_i est alors $\{t_j \ /\ \max(0, k\text{-}n\text{+}1\text{+}i \le j \le n\text{-}1\text{+}i\text{-}k\}$

3.4. Le modus ponens généralisé

Le modus ponens généralisé consiste en la donnée d'une règle (X est A) $\to$ (Y est B) où A et B sont des relations floues unaires, et d'un fait X est A', on en déduit Y est B', naturellement suivant certaines modalités où A' étant «proche» de A, alors B' est construit comme «voisin» de B.
Pour une t-norme T et un opérateur π d'implication défini par :

$$\pi(x, y) = (\mu_A(x) \to \mu_B(y))$$

la définition donnée par [Zadeh 73] en est $\mu_{B'}(y) = \sup_{x \in U} (T(\mu_{A'}(x), \pi(x,y))$.

$$\boxed{\mu_{B'}(y) = \sup_{x \in \mathrm{supp}(A')} (T(\mu_{A'}(x), \mu_A(x) \to \mu_B(y))}$$

On peut voir l'analogie avec la formule probabiliste $p(B) = \Sigma p(A_i).p(B/A_i)$, mais aussi une définition résultant du principe d'extension de Zadeh.

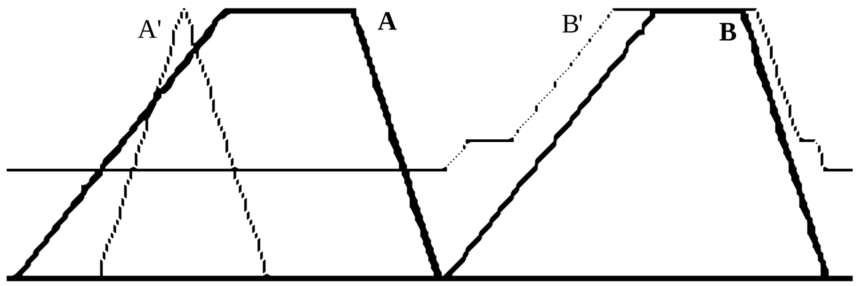

Figure 3.8 Exemple avec la t-norme de Zadeh et l'implication de Luckasiewicz

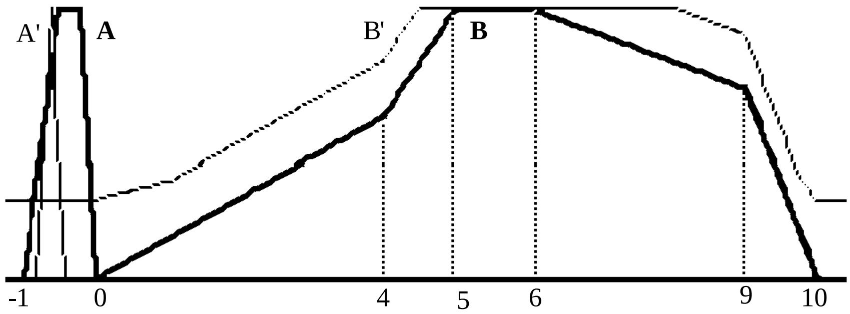

Figure 3.9 Exemple pour la t-norme Hamacher et l'implication de Luckasiewicz

Exercice 3.15
Pour A' = A puis A' coupant A triangulaire et B triangulaire, représenter B' avec π de Mamdani, Lukasiewicz, Kleene et Gödel.

Exercice 3.16
Comparaison des implications lors d'inférence floue (X est A) $\rightarrow$ (Y est B)
Montrer que si $\mu_{A'} \leq \mu_A$ alors on a $\mu_{B'} = \mu_B$ avec Kleene, Gödel, Lukasiewiecz, Reichenbach, Willmott, Mamdani, tandis que l'on a $\mu_{B'} \leq \mu_B$ pour Rescher.
Si maintenant on a $\mu_{A'} \geq \mu_A$ alors $\mu_{B'} \geq \mu_B$ pour Kleene, Gödel, Goguen, Lukasiewicz, Reichenbach, Willmot, alors que $\mu_{B'} = \mu_B$ pour Mamdani.

Exercice 3.17
Au cas d'une observation précise, c'est à dire qu'il existe x_0 tel que $\mu_{A'}(x_0) = 1$ et $\mu_{A'}(x) = 0$ pour tout autre x, montrer alors qu'il y a incertitude avec Kleene, Lukasiewicz, Reichenbach, Willmott, mais pas avec Mamdani, et pas avec Gödel, Goguen et Rescher dans le cas où $\mu_A(x_0) \neq 0$ pour ces trois derniers.

Exercice 3.18
Montrer qu'en général, si μ = ps (A' est A) alors, en dehors du support de B, B' vaut 1 - μ pour Kleene et Reichenbach.

Exercice 3.19
Cas d'une conclusion précise : B est défini par $\mu(y_0) = 1$ et si $y \neq y_0$ alors $\mu(y) = 0$, en ce cas $\pi(x, y_0) = \mu_A(x)$ pour Mamdani, et $\pi(x, y_0) = 1$ sauf pour Willmott et Mamdani.
Une étude systématique a été réalisée dans [Després 88].

ALGORITHME APPROCHÉ DE CALCUL DU MODUS PONENS GÉNÉRALISÉ, DANS LE CAS DES Λ-TRAPÈZES

Etant donnés les λ-trapèzes $A = (a_A, b_A, \alpha_A, \beta_A, \lambda_A)$, $B = (a_B, b_B, \alpha_B, \beta_B, \lambda_B)$, intervenant dans la règle $A \rightarrow B$, et la valeur floue $A' = (a_{A'}, b_{A'}, \alpha_{A'}, \beta_{A'}, \lambda_{A'})$ approchant A, l'algorithme renvoye un λ-trapèze $B' = (a_{B'}, b_{B'}, \alpha_{B'}, \beta_{B'}, \lambda_{B'})$.
On calcule les niveaux $\xi = \sup \{\mu_{A'} / \mu_A = 0\}$, $\zeta = \inf \{\mu_A / \mu_{A'} = 1\}$. Un autre indicateur a été utilisé, c'est $\eta = \inf\{\mu_A / \mu_{A'} \geq ps\ (X \text{ est } A)\}$ sorte de degré de proximité.
Dans le cas où on prend la conjonction min et l'implication de Gödel, on examine dans l'ordre les cas :

Conditions :

Si $A' \subset A$ alors $B' = (a_B, b_B, \alpha_B, \beta_B, \max(\lambda_B, \lambda_{A'}))$
Si $A \cap A' = \varnothing$ alors $B' = (a_B, b_B, \alpha_B, \beta_B, 1)$
Si $\text{supp}(A') \subset \text{supp}(A)$,
$$B' = (a_B - (1-\zeta)\alpha_B, b_B + (1-\zeta)\beta_B, \zeta\alpha_B, \zeta\beta_B, \max(\lambda_B, \lambda_{A'}))$$
Si $\text{noy}(A') \subset \text{noy}(A)$ alors $B' = (a_B, b_B, \alpha_B, \beta_B, \max(\xi, \lambda_{A'}))$
Sinon $B' = (a_B - (1 - \zeta)\alpha_B, b_B + (1 - \zeta)\beta_B, \alpha_B, \beta_B, \max(\xi, \lambda_{A'}))$

RÈGLES GRADUELLES

Aucune formalisation efficace et convaincante de règles telles que «plus X est A, alors plus Y est B», n'a encore été réalisée. On peut les représenter par plusieurs règles utilisant une hiérarchie petit, moyen, grand etc pour X ainsi que pour Y, mais on voudrait bien sûr éviter les problèmes de seuil pour passer d'une règle à l'autre. D. Dubois distingue par ailleurs : «plus X est A, plus il est certain que Y est B» avec par exemple «plus on se lève tard, plus on est certain de rater son train» qui est le plus souvent utilisé avec l'implication de Kleene $IK(p, q) = \max(1 - p, q)$, le modus-ponens généralisé donnant une distribution de possibilité. En ce cas, si l'agrégation se fait avec le min, il suffit que l'une des règles donne 0 pour interdire les autres.
Par ailleurs «plus X est A, plus il est possible que Y soit B» avec par exemple : «plus 2 voitures sont semblables, plus il est possible que leurs prix soient identiques», où on utilisera plutôt l'implication de Gödel $IG(p, q) = $ si $(p < q)$ alors 1 sinon q.

Exercice 3.20
On rappelle que l'image d'un ensemble flou A par une relation R est un ensemble flou B définit par :
$\mu_B(y) = \sup \{\min (\mu_A(x), \mu_R(x, y)\}$. Vérifier que pour les α-coupes, $B_\alpha = R_\alpha(A_\alpha)$ et que cela correspond à un ensemble de règles «si $x \in A_\alpha$ alors $y \in B_\alpha$».
En prenant l'implication de Gödel $(p \rightarrow q) = (\text{si } p \leq q \text{ alors } 1 \text{ sinon } q)$, montrer qu'alors $R \subset [A \rightarrow R(A)]$.

INTERPOLATION ENTRE LES RÈGLES

Pour une collection de n règles «si X est A_i, alors Y est B_i», deux points de vue peuvent être défendus.
a) Le point de vue logique qui consiste à prendre la conjonction des règles :
$x R^+ y = T_{(1 \le i \le n)} (\mu_{Ai}(x) \rightarrow \mu_{Bi}(y))$ où T est une t-norme et $\rightarrow$ une définition d'implication. Les règles produisent des degrés de certitudes,
$\mu_{B'}(y) = \min_{1 \le i \le n} (\max (1 - \mu A_i(x), \mu B_i(y))$ et pour tout i on doit avoir $B' \subseteq B_i$, et pour une collection de règles certaines, la relation Rf entre x et y est telle que pour tout i on ait $A_i o R_f \subseteq B_i$ d'où vient $R_f = \cap (\neg A_i \cup B_i)$ encadrement noir de la figure.

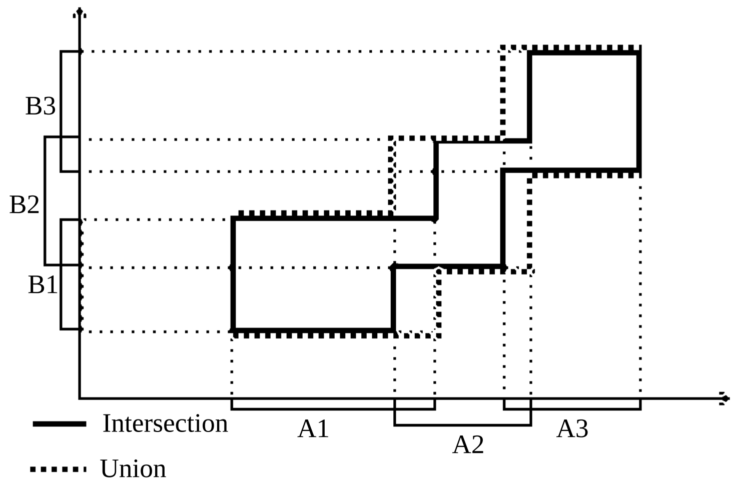

Figure 3.10 Conjonction ou disjonction de règles se chevauchant.

b) Le point de vue de la méthode de Mamdani (c'est le contrôle flou du chapitre suivant) :
$x R^- y = S (\mu_{Ai}(x) \, T \, \mu_{Bi}(y))$ où S est une t-conorme et T = min, consiste à prendre la disjonction des règles, donc l'union des conclusions ce qui peut parfaitement se justifier dans la mesure où il considère que chaque règle peut apporter une réponse partielle et que c'est leur totalité qui doit apporter une réponse consensuelle. Alors les règles produisent des degrés de possibilités $\mu_{B'}(y) = \max_{(1 \le i \le n),} (\min (\mu_{Ai}(x), \mu_{Bi}(y))$ et pour tout i on doit avoir $B_i \subseteq B$. Ce que fait Mamdani est de prendre la relation $R^- = \cup (A_i * B_i)$ (produit cartésien) correspondant au pourtour pointillé.
Pour des entrées précises x :

$$\mu_{Bf} (y) = \sup_{x \in A} (\min(\mu_A(x), \min_{1 \le i \le n} \pi_i (x, y))$$

$$\mu_{Bg}(y) = \sup_{x \in A} (\min(\mu_A(x), \max_{1 \le i \le n} \min (\mu_{Ai}(x), \mu_{Bi}(y)))$$

APPLICATION AUX SYSTÈMES-EXPERTS FLOUS

Dans les systèmes experts flous, si on trouve deux assignations μ et μ' pour une conclusion, et leur compatibilité $h = \sup \min(\mu, \mu')$, si elle n'est pas nulle, permet de prendre par exemple $\mu_{ag} = \min(\mu, \mu') / h$ comme agrégation possible de la conclusion.

En fait, aucun problème concret lié aux systèmes-experts, comme on l'a déjà souligné, n'est résolu de façon universelle. Comment l'utilisateur doit-il donner ses appréciations (une échelle linguistique de 5 ou 7 valeurs est largement suffisante), mais ensuite comment doit se réaliser la conjonction des hypothèses (quelle t-norme choisir ?), comment opérer l'inférence ?, comment doit-on gérer les contradictions et réaliser l'agrégation des différentes conclusions ?