L'ANALYSEUR D'INDICES:
UN SYSTÈME D'AIDE À LA DÉCISION BASÉ SUR LES FONCTIONS DE CONFIANCE DE SHAFER

CLAUDE BOIVIN
Revenu Québec
Louis-Claude Paquin
Centre ATO.CI, UQAM

0. RÉSUMÉ

Dans de nombreuses situations de l'activité humaine, nous colligeons des indices, souvent des informations incertaines et en faisons une analyse pour en tirer des conclusions. Plus souvent qu'autrement, cette analyse s'avère un laborieux et pénible travail de réflexion mentale. Nous pensons que cette réflexion peut être facilitée par un recoupement méthodique des indices que nous manipulons. Idéalement, il faudrait spécifier un modèle probabiliste complet pour traiter l'incertitude, tâche qui peut se révéler trop lourde pour un non spécialiste ou tout simplement non réalisable dans certains cas. Nous proposons ici une solution simple à ce problème, basée sur la théorie des fonctions de confiance de Shafer [1976], qui nous paraît bien adaptée à la manipulation d'information imparfaite. La présente contribution décrit cette solution et une implantation : l'Analyseur d'indices.

1. L'ANALYSE D'ÉVÉNEMENTS

Dans de nombreuses situations de l'activité humaine, nous colligeons des indices (observations, sources d'information) et en faisons une analyse pour en déduire des résultats. Plus souvent qu'autrement, c'est par un laborieux travail de réflexion mentale que nous arrivons à tirer des conclusions. C'est ce que nous appelons ici **l'analyse d'événements**. Une particularité de l'analyse d'événements est qu'elle est effectuée à partir d'informations incertaines. Nous pensons que cette réflexion peut être facilitée par un recoupement méthodique des indices que nous manipulons. Souvent le traitement de l'incertitude par un modèle probabiliste complet représente une tâche trop lourde pour un non spécialiste ou n'est pas réalisable. Nous proposons ici une solution simple basée sur la théorie des fonctions de confiance de Shafer [1976], qui nous paraît bien adaptée à la

manipulation d'information imparfaite. Cette hypothèse est a été vérifiée au moyen d'une implantation de ce modèle appelée **l'Analyseur**.

Les situations (événements) que nous considérons sont celles où un choix doit être effectué parmi plusieurs possibilités mutuellement exclusives [Shafer, 1976: 36]. Pour tirer une conclusion, l'analyste s'appuie sur des indices ou des faits qu'il ou elle observe et met en relation avec l'ensemble des possibilités. Chaque observation permet d'avancer une proposition (un sous-ensemble de l'ensemble des possibilités, une hypothèse). Un nombre entre 0 et 1 est associé à chaque proposition pour indiquer jusqu'à quel point celle-ci est considérée comme vraie.

Comme hypothèse simplificatrice nous supposons que lorsque nous confrontons deux collections d'indices, celles-ci sont considérées entièrement distinctes c'est-à-dire, que ce sont des sources d'information qui ne s'influencent pas entre elles. Cette hypothèse nous permettra d'utiliser la règle de Dempster [Shafer, 1976: 57-61] pour effectuer le recoupement des propositions associées. La règle de Dempster est au coeur de notre système d'analyse d'événements.

Un type de situation qui nous intéresse est celle où le décideur énonce une question en termes généraux (en haut, stable ou en baisse, fort ou faible, chaud ou froid, coupable ou non coupable). Les données relatives à ces questions ne peuvent provenir que d'autres sources, différentes, mais néanmoins reliées à celles-ci. Il est alors difficile de spécifier un modèle probabiliste complet. Le but de l'analyse d'événements est de traiter l'information disponible, même si celle-ci est fragmentaire, considérant qu'une réponse moins précise serait préférable à rien.

Un autre type d'application touche les problèmes de classement d'envergure, dont la solution peut passer par des techniques d'analyse statistique multidimensionnelle. L'analyse d'événements se veut une alternative simple à ces méthodes, en particulier lorsque le savoir-faire nécessaire ou les ressources pour les appliquer ne sont pas disponibles.

Après avoir présenté sommairement les bases de la théorie des fonctions de confiance dans la section 2, nous étudions dans la section 3 comment évaluer l'importance d'un indice ou d'un fait. Notre cadre conceptuel pour l'analyse d'événements est présenté dans la section 4. Nous montrons dans les sections 5 et 6 comment obtenir un coefficient pour une proposition. La section 7 présente l'Analyseur, qui implante l'analyse d'événements.

2. CONCEPTS DE BASE DE LA THÉORIE DES FONCTIONS DE CONFIANCE

La théorie des fonctions de confiance est due à Glenn Shafer [1976] et tire son origine dans les travaux d'Arthur Dempster [1967] sur l'établissement de mesures de probabilité inférieure et supérieure. Nous en présentons brièvement les notions de base ci-après. Pour un exposé complet, on consultera Shafer [1976, 1982, 1987].

2.1. L'ENSEMBLE DES POSSIBILITÉS

Nous considérons un ensemble fini Ω dont chaque élément s'interprète comme une réponse possible à une question qui nous intéresse. Nous assumons qu'un des éléments de l'ensemble constitue la bonne réponse et qu'il n'y en a qu'un seul. Cet ensemble est appelé «l'ensemble des possibilités».

2.2. EXTENSION DU CADRE PROBABILISTE

Si nous étions dans le cadre de la théorie des probabilités, nous pourrions assigner à chaque élément x de l'ensemble Ω une masse $P(x)$ entre 0 et 1, c'est-à-dire une fonction $P: \Omega \rightarrow [0,1]$, représentant la probabilité d'occurence de l'élément x. Malheureusement, il y a des situations où ceci n'est pas réalisable et où il nous est seulement possible d'assigner des probabilités aux éléments d'une autre question, différente de celle qui nous intéresse, mais néanmoins reliée.

Supposons que D est cette autre question, c'est-à dire, D est un ensemble fini de possibilités en relation avec Ω et dont nous connaissons, pour chaque élément d, la probabilité d'occurrence $P(d)$. Le fait d'étudier Ω par l'intermédiaire de D introduit de l'imprécision dans la recherche de la bonne réponse à la question posée par Ω; en effet, il arrivera que la connaissance d'un élément de D ne permettra pas d'identifier exactement un des éléments de Ω mais seulement un sous-ensemble A de Ω comme étant la réponse à la question posée. Nous allons alors mesurer les chances que la bonne réponse soit A.

Étant donné une relation R entre D et Ω, nous dirons qu'un élément d de D est en relation avec un sous-ensemble A de Ω si d est une indication que A est la bonne réponse à la question posée par Ω. De façon formelle on pose la définition suivante :

Étant donné un élément d de D, et un sous ensemble A de Ω,

d R A si d est relié à tous les éléments x de A et seulement à ceux-ci. Si d n'est relié à aucun sous-ensemble de Ω ou à Ω lui-même, d est éliminé et la distribution de probabilités sur D est normalisée.

Nous pouvons alors obtenir un nombre mesurant les chances que le sous-ensemble A de Ω constitue la bonne réponse à notre question ainsi:

Pour toute proposition A 2^{Ω}, on considère la fonction m: 2^{Ω} –> [0,1], telle que :

(1) m(A) = ∑{P(d); d R A, d D},

c'est-à-dire que chaque fois que tous les éléments de A sont reliés à un élément d de D, la valeur P(d) correspondante sera additionnée.

Le nombre m(A) est la somme des probabilités de tous les éléments de D qui indiquent A comme vraie. On dira que m(A) est **le degré d'adhésion direct (la masse)** de la proposition A. Le terme «direct» est utilisé ici pour indiquer qu'un indice tend sans ambiguïté vers une seule proposition. La somme des valeurs m(A) donne 1. De plus, si Ø est l'événement impossible, m(Ø) = 0, par définition.

La mesure m(A) peut aussi s'interpréter comme une allocation globale de probabilité aux événements élémentaires constituant A, sans préjuger de la répartition de cette masse entre ces événements élémentaires [Dubois et Prade, 1987: 17].

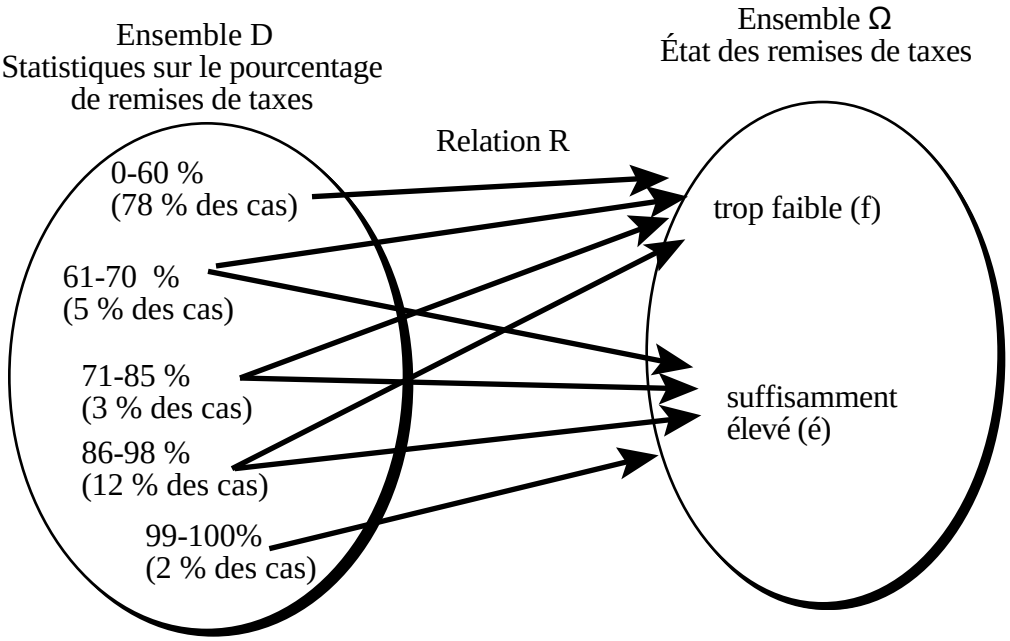
EXEMPLE : EXAMEN D'UN DOSSIER POUR VÉRIFICATION FISCALE.

En préparation de son travail, un vérificateur examine les remises de taxe sur les repas. Il lui faut évaluer si celles-ci sont trop faibles ou suffisamment élevées. Étant donné le genre de commerce qu'il examine, il suppose que si le rapport entre les ventes taxables et les ventes totales est inférieur à 60 %, les remises sont trop faibles et qu'un pourcentage supérieur à 99 % est suffisamment élevé. Entre les deux, il ne peut décider.

Des données sont disponibles sur l'ensemble des dossiers vérifiés antérieurement; elles permettent d'établir que 78 % des dossiers vérifiés avaient un pourcentage de ventes taxables faible (entre 0 et 60 %), 2 % des dossiers étaient corrects (99100 %) et 20 % des

dossiers étaient dans la zone intermédiaire. La figure suivante illustre la relation établie entre l'ensemble des dossiers vérifiés antérieurement et la question à l'étude :

Figure 1 :Relation entre les statistiques sur les remises de taxe et la décision



On obtient la distribution du degré d'adhésion direct :

$$m(\{f\}) = 0,78$$
$$m(\{é\}) = 0,02.$$
$$m(\{f, é\}) = 0,05 + 0,03 + 0,12 = 0,20.$$

$m(A) = 0$, pour toute autre proposition A de Ω. Les propositions ayant une masse supérieure à 0 seront appelée «les propositions d'intérêt». Dans cet exemple simple, toutes les propositions de Ω ont une masse positive.

2.3. LE DEGRÉ D'ADHÉSION ET LA PLAUSIBILITÉ.

La construction d'une mesure m sur l'ensemble des parties de Ω permet de définir deux autres mesures, le degré d'adhésion et la plausibilité.

Le degré d'adhésion envers une proposition A de Ω est défini de la façon suivante :

$$(2) \quad Bel(A) = \sum \{m(B) ; A \quad B\}$$

$Bel(A)$ (de l'anglais «belief», croyance) est obtenue en additionnant les masses de toutes les propositions B qui rendent nécessaire l'occurrence de A. $Bel(A)$ s'interprète comme la confiance que nous avons que A contient la bonne réponse à la question posée. Notons la différence avec le degré d'adhésion direct; le degré d'adhésion direct prend en compte seulement les indications qui pointent exactement vers A, tandis que le degré d'adhésion est augmenté du poids des indications vers des propositions qui entraînent A.

La plausibilité d'une proposition A de Ω est définie ainsi :

$$(3) \qquad Pl(A) = \sum \{m(B) ; B \leftrightarrow A \neq Ø.\}$$

$Pl(A)$ est la somme des masses de toutes les propositions qui rendent possible l'occurence de A, c'est-à-dire, dont l'intersection avec A est non vide.

Une relation importante entre le degré d'adhésion et la plausibilité est la suivante :

$$(4) \qquad Pl(A) = 1 - Bel(\neg A)$$ (¬ indique la négation).

EXEMPLE (SUITE)

À partir de la distribution du degré d'adhésion direct, nous appliquons les formules (2) et (3) pour obtenir, pour chaque proposition d'intérêt, le degré d'adhésion et la plausibilité :

Proposition	Degré d'adhésion	Plausibilité
{f}	0,78	0,98
{é}	0,02	0,22
{f, é}	1	1

La «croyance» que les remises sont faibles est de 0,78 et la croyance que les remises sont suffisamment élevées est de 0,02. $Bel(\{f, é\})$ correspond à la croyance que la réponse est une des deux possibilités, ce dont nous sommes certains.

2.4. LES FONCTIONS D'ADHÉSION SIMPLES

Dans le problème que nous considérons, une indication est toujours reliée à une et une seule proposition. Étant donné un ensemble de possibilités Ω et une proposition A de Ω, on pose :

$m(A) = s, \quad 0 \leq s \leq 1,$
$m(\Omega) = 1\text{-}s,$
$m(B) = 0, \quad B$ différent de $A, \Omega.$

Le degré d'adhésion d'une proposition B quelconque est alors obtenu par la fonction suivante :

$$S(B) = \begin{cases} 0, \text{ si } B \text{ ne contient pas } A, \\ s, \text{ si } B \quad A, B \neq \text{de } \Omega, \\ 1, \text{ si } B = \Omega. \end{cases}$$

S est appelé fonction d'adhésion simple à focus sur A.

Dans une analyse d'événements, une stratégie naturelle consiste à évaluer séparément des collections d'indices indépendantes pour tenter d'en effectuer le recoupement. Nous entendons ici par indépendance la situation où deux sources de renseignement ne s'influencent pas entre elles. La théorie propose avec la règle de Dempster [Shafer, 1976: 58-60] un procédé pour combiner nos sources de renseignement.

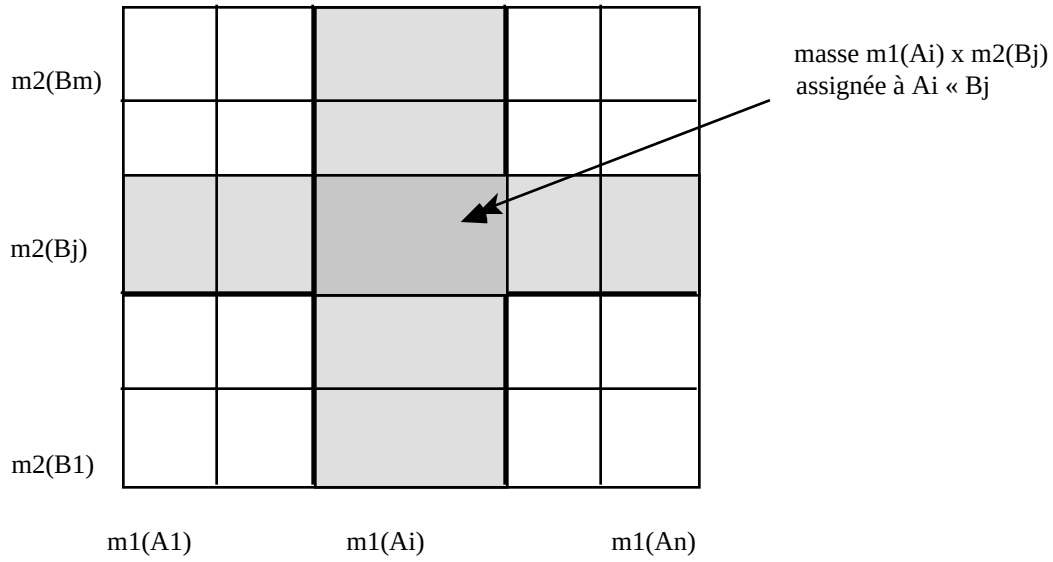
2.5. LA RÈGLE DE DEMPSTER

Nous présentons le recoupement de deux propositions par la règle de Dempster pour le cas de deux fonctions d'adhésion simples, ou de sa résultante avec une autre fonction d'adhésion simple. Selon le vocabulaire de Shafer [1976: 87], la résultante de la combinaison, par la règle de Dempster, de deux ou plusieurs fonctions d'adhésion simples est appelée «fonction d'adhésion séparable».

Étant donné deux distributions m_1 et m_2 du degré d'adhésion direct sur des ensembles Ω_1 et Ω_2, le degré d'adhésion direct combiné m_{12} correspondant à leur recoupement est obtenu par la formule suivante: pour toute proposition X de $\Omega_1 x \Omega_2$,

$$(5) \quad m_{12}(X) = \sum \{ m_1(A) \infty m_2(B); A \leftrightarrow B = X, \Omega_1 \quad A \text{ et } \Omega_2 \quad B \}$$

Figure 2 : Règle de Dempster pour combiner les distributions de degré d'adhésion direct



Dans le recoupement des distributions, il peut arriver que le recoupement de A avec B donne une proposition impossible. Les cas impossibles sont alors éliminés et la distribution résultante est normalisée, soit :

$$(6) \qquad m^*(X) = C^{-1} m_{12}(X),$$
$$\text{avec } C \equiv m_{12}(\emptyset) = \sum\{m_1(A) \infty m_2(B); A \leftrightarrow B = \emptyset\} > 0.$$

L'application de la règle de Dempster permet d'obtenir un degré d'adhésion combiné $Bel_{12}(A)$ qui reflète le recoupement des indices pris en considération.

En pratique, nous aurons couramment à combiner plus de deux indices. Une caractéristique importante de la règle de Dempster est que le résultat final de la combinaison de plusieurs indices est indépendant de l'ordre de combinaison [Shafer, 1976: 62-64].

2.6 L'INDICE DE CONTRADICTION

Un aspect intéressant de la constante C est son interprétation comme un indice de la contradiction entre les propositions qui ont été recoupées. Il est particulièrement important de suivre l'évolution de cet indice lorsque plusieurs indices sont combinés. En effet, posons:

$$C_n = m_{(12\ldots n-1)\,n}(\emptyset),$$

l'indice de contradiction engendré par la combinaison de l'indice n avec le résultat de la combinaison des indices précédents; si on effectue 1, 2, 3, …, n combinaisons, nous obtenons [Shafer, 1976: 66; Almond, 1989: 15] un indice de conflit de l'ensemble d'indices:

$$(7) \quad Con_n = 1 - [1-C_1][1-C_2] \ldots [1-C_n],$$

avec $n \geq 2$,
$Con_1 = 0$ et $C_1 = 0$, par définition.

L'équation 7 peut aussi s'écrire sous forme récursive, soit

$$(7') \quad Con_n = 1-(1-Con_{n-1}) \times (1-C_n), \quad n \geq 2.$$

L'équation (7') nous montre bien qu'à chaque addition d'un nouvel indice, il peut y avoir renforcement de l'indice de conflit. La mesure Con_n nous permettra donc de détecter les situations indécidables.

3. LA DIFFICULTÉ D'ÉNONCER LE DEGRÉ D'ADHÉSION EN FAVEUR D'UNE PROPOSITION

Déterminer le degré d'adhésion en faveur d'une proposition est une opération reconnue par plusieurs comme difficile :

> «A more difficult issue is whether such formal systems can be made meaningful and reliable enough that their numerical assessments of uncertainty can be trusted»; [Dempster, 1989] «where are all the numbers coming from? (…) does the expert have sufficient experience to justify all those numbers, or is he making them up?».[Cheeseman, 1988]

Souvent, il n'y a pas suffisamment de données pour effectuer les analyses statistiques nécessaires. L'analyste doit alors se baser sur son expérience pour effectuer mentalement une évaluation. Par exemple, si un enquêteur policier considère par expérience que des objets peuvent facilement être subtilisés sous des vêtements amples ou cachés dans de grands sacs, les chances qu'un objet puisse être volé dans ces conditions sont alors jugées grandes. Si un des suspects avait comme caractéristique de toujours circuler avec un grand sac dans les magasins, l'enquêteur en déduira que les chances sont grandes que ce suspect soit le coupable. Le nombre qu'il pourrait donner pour cristalliser son évaluation pourrait

être basé sur le nombre de vols à l'étalage dans les grands magasins et le nombre de fois que le coupable avait dissimulé l'objet sous des vêtements amples ou dans un grand sac.

> «The weighting of evidence may be viewed as a mental experiment in which the human mind is used to assess probability much as a pan balance is used to measure weight».[Shafer et Tversky, 1985]

En l'absence de données numériques, nous croyons qu'il serait plus facile de mesurer la certitude d'une proposition sur une échelle ordinale (peu probable, assez probable, presque certain) que sur une échelle à rapport, comme l'intervalle [0,1]. Reprenons l'exemple de la section 2; si un vérificateur n'avait pas accès à des données, il lui faudrait baser son évaluation sur son expérience. Il lui serait plus facile de dire «les remises de ce mandataire sont presque sûrement trop faibles» que «il y a 83 % de chances que les remises soient trop faibles». L'évaluation du degré d'adhésion d'une proposition peut donc être effectuée en deux étapes : d'abord une évaluation sur une échelle ordinale, ensuite une traduction de cette évaluation sur l'intervalle [0,1].

Dans sa monographie de 1976 [chap. 4 et 5], Shafer établit, pour les fonctions d'adhésion simples et séparables, une relation entre le degré d'adhésion d'une proposition et le poids des indices appuyant celle-ci. Le problème d'établir le degré d'adhésion direct d'une proposition est ainsi ramené à celui d'indiquer quelle importance ont les indices à l'appui de cette proposition. Le concept de «l'importance des indices» nous paraît plus facile à manipuler que le degré d'adhésion, car il est courant d'effectuer des analyses par addition d'éléments et par comparaison des résultats accumulés. Les faits sont concrets et prennent tout leur sens lorsqu'ils sont mis en relation avec une proposition. Nous croyons ainsi qu'il est plus facile pour un analyste d'effectuer une pondération des indices que d'établir le degré d'adhésion direct d'une proposition.

Donc, plutôt que de mener la réflexion jusqu'à l'obtention d'un degré d'adhésion ou d'une probabilité, nous proposons de mettre l'accent sur la pondération des indices apportés à l'appui d'une proposition. Cette pondération sera facilitée par l'utilisation d'une échelle à plusieurs modalités (négligeable, peu important, moyennement important, etc.) que nous traduirons sur une échelle continue.

4. UN CADRE CONCEPTUEL POUR L'ANALYSE D'ÉVÉNEMENTS

Les situations à analyser n'ont pas toutes la même envergure; certaines hypothèses peuvent être vérifiées en recoupant trois ou quatre sources d'informations, d'autres nécessitent de recouper 10 ou 15 sources, tandis que des cas complexes vont nécessiter de recouper des centaines de sources. Nous appellerons cela «la taille du problème». Nous pensons que le poids d'un indice à l'appui d'une hypothèse sera inversement proportionnel au nombre total de sources possibles. En d'autres mots, moins il y a d'indices à considérer, plus le poids de chacun est grand. Par exemple, un indice jugé très important dans l'évaluation d'une hypothèse faisant intervenir seulement quelques indices n'aura pas le même poids qu'un autre indice jugé très important, mais relativement à un ensemble de 150 indices. **L'évaluation de l'importance d'un indice est ainsi fonction de la taille du problème.**

De plus, à aucun moment nous pensons que les indices observés constituent la population totale des indices pertinents à l'hypothèse considérée. Nous assumons plutôt qu'un analyste ne peut observer qu'un échantillon d'indices. Le degré d'adhésion réel envers une proposition ou hypothèse ne peut être connu à moins que tous les éléments relatifs à cette proposition aient été évalués, c'est-à-dire que la population totale des indices ait été examinée. **On ne pourra donc en obtenir que des estimations successives, qui s'améliorent à l'ajout de chaque nouvelle information.**

Nous présentons maintenant un modèle pour représenter une population d'indices. D'abord, nous choisissons une échelle ordinale de mesure de l'importance d'un indice; ensuite nous décrivons le modèle retenu. Nous utiliserons ce modèle pour obtenir une estimation du poids résultant de la combinaison de plusieurs indices à l'appui d'une proposition.

4.1. MESURE DE L'IMPORTANCE D'UN INDICE

La première étape de l'évaluation d'une proposition va consister à indiquer l'importance des indices à l'appui de celle-ci. Pour cela nous construisons une échelle de mesure à plusieurs modalités et faisons correspondre à chaque modalité un degré d'adhésion direct, ce qui permet de fixer un poids initial. Ce procédé est arbitraire puisque le choix des modalités et de leur poids dépend uniquement du concepteur.

4.1.1. Établissement d'une échelle de catégories d'importance

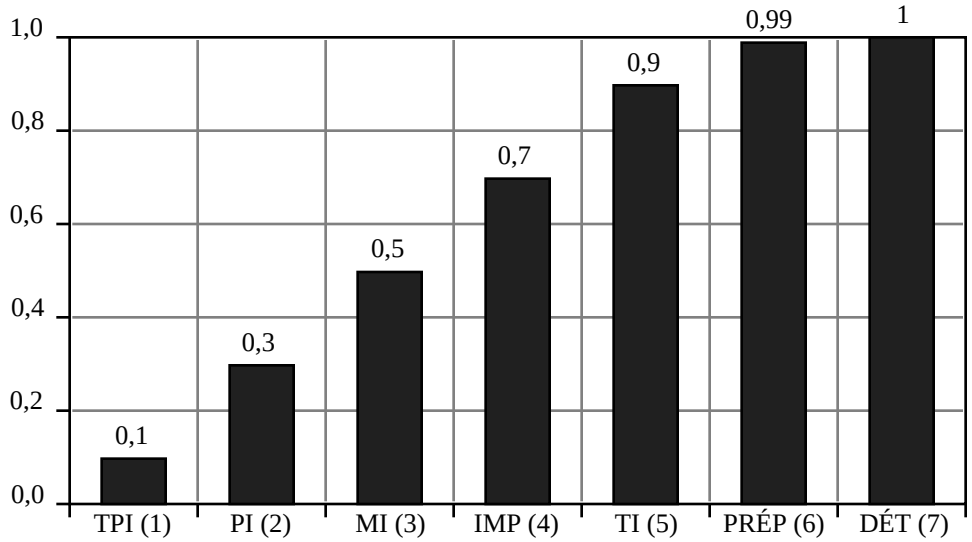
Nous proposons d'indiquer sur une échelle de mesure à plusieurs modalités la catégorie d'importance de l'indice à l'appui d'une proposition. Le tableau 1 décrit l'échelle que nous avons retenue.

Tableau 1: Catégories d'importance des indices

Catégorie	Abréviation
(1) très peu important,	TPI
(2) peu important,	PI
(3) moyennement important,	MI
(4) important,	I
(5) très important,	TI
(6) prépondérant,	PRÉP
(7) déterminant.	DÉT

Pour chaque catégorie d'importance, la valeur désirée du degré d'adhésion direct est fixée. La figure 3 suivante illustre les seuils que nous avons choisis :

Figure 3 : Degré d'adhésion direct en fonction de la catégorie d'importance



Puisque nous nous en tenons toujours aux fonctions d'adhésion simples, le choix d'une catégorie d'importance s'interprète comme l'expression d'un intervalle «degré d'adhésion, plausibilité», avec la valeur de plausibilité égale à 1. Par exemple, si un indice est considéré

très important (TI), cela induit un degré d'adhésion égal à 0,9 envers la proposition considérée, avec un degré de plausibilité de 1.

À titre d'illustration d'un autre choix de catégories, on pourrait ajouter une catégorie «négligeable» avec un degré d'adhésion direct de 0,01. On aurait alors un nombre égal de catégories supérieures et inférieures à 0,5.

4.1.2. Détermination du poids des catégories d'importance

La relation entre les notions de «poids des indices» et de degré d'adhésion (ou probabilité) est décrite par Shafer [1976: 77]. Ainsi le degré d'adhésion direct attaché à une proposition devrait être déterminé par le poids des indices qui attestent cette proposition. Il est montré que dans le cas où un indice pointe exactement vers une seule proposition, le degré d'adhésion direct en faveur de celle-ci est déterminé par une fonction du poids de l'indice. En retenant pour principe qu'il serait naturel que les poids s'additionnent, la fonction désirée est de la forme suivante :

Notons :
$m(A) = s$, le degré d'adhésion direct en faveur de la proposition A
et $w(A)$, le poids des indices à l'appui de la proposition. Alors :

(8) $s = 1 - e^{Kw(A)}$, où :
K est une constante à déterminer
et e est la fonction exponentielle.

Alors que le degré d'adhésion direct $m(A)$ est un nombre entre 0 et 1, le poids $w(A)$ associé peut varier entre 0 et un nombre infiniment grand. Plus le poids d'un indice est grand, plus le degré d'adhésion direct doit être élevé. Si un indice est déterminant, il a un poids infini et l'exponentielle devrait être nulle, tandis qu'un degré d'adhésion direct nul serait la conséquence d'un indice n'ayant aucun poids. La constante K doit donc être négative pour obtenir la relation désirée.

L'équation (8) fournit ainsi une équivalence entre le degré d'adhésion direct s envers une proposition A et son poids associé $w(A)$. En inversant l'équation (8), le poids $w(A)$ est exprimé en fonction du degré d'adhésion direct s :

(9) $w(A) = K^{-1}\ln(1 - s)$,
où $\ln$ est la fonction logarithme en base e.

Comme le fait remarquer Shafer, le choix de la constante de l'équation (9) est arbitraire. Posons donc $K = -1$, ce qui donne :

$$(9') \quad w(A) = -\ln(1 - s).$$

En appliquant l'équation (9') aux catégories de la figure 3, nous pouvons alors déterminer, pour chaque catégorie d'importance, le poids associé à un degré d'adhésion direct fixé a priori, ce qui donne le tableau suivant :

Tableau 2 : Catégories d'importance des indices et poids des catégories

Catégorie	m(A)		w(A)
(1) très peu important	0,1	$-\ln(1 - 0,1) =$	0,105
(2) peu important	0,3	$-\ln(1 - 0,3) =$	0,357
(3) moyennement important	0,5	$-\ln(1 - 0,5) =$	0,693
(4) important	0,7	$-\ln(1 - 0,7) =$	1,204
(5) très important	0,9	$-\ln(1 - 0,9) =$	2,303
(6) prépondérant	0,99	$-\ln(1 - 0,99) =$	4,605
(7) déterminant	1	$-\ln(1 - 1) =$	infini

Par exemple, un indice considéré très important recevra un poids initial égal à 2,303.

4.2. UN MODÈLE POUR DÉCRIRE LA POPULATION DES INDICES ASSOCIÉS À UN ENSEMBLE DE POSSIBILITÉS

Nous supposons l'existence d'un ensemble U de N indices, numérotés de 1 à N :

$$U = \{1,2,\ldots,k,\ldots,N\}.$$

Ces indices peuvent être observés et appuyer différentes propositions de l'ensemble Ω des possibilités. Chaque indice k de U a un poids initial w_k fixé.

Nous supposons que le processus suivant régit la population des indices.

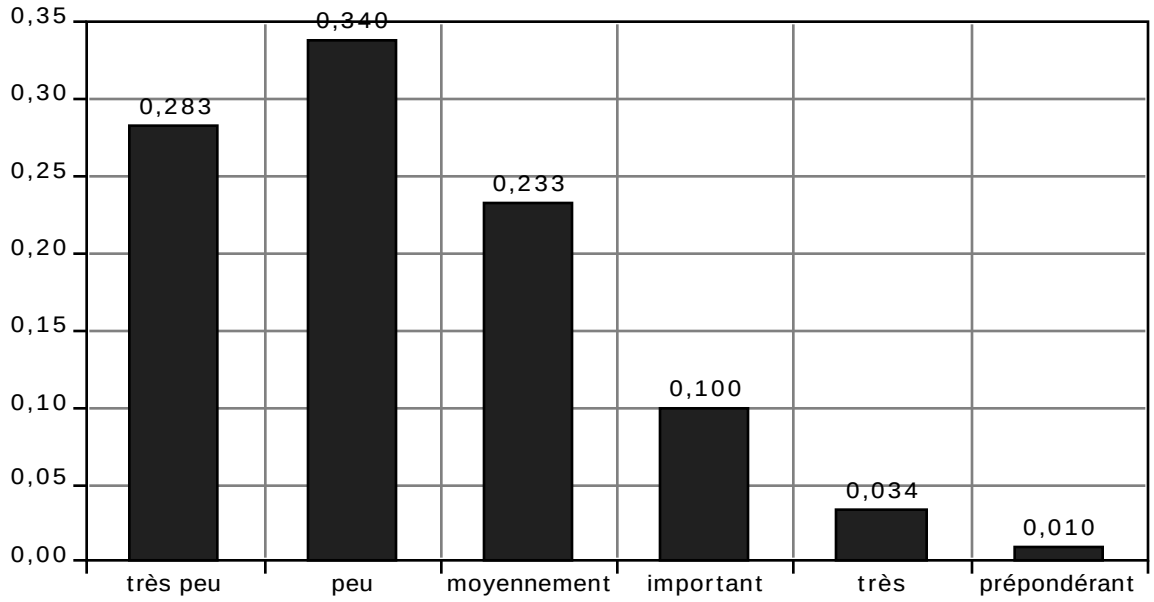
a) Généralement plusieurs indices de peu d'importance interviennent dans l'évaluation d'une proposition, tandis que seulement quelques indices d'importance peuvent intervenir. Ainsi, un analyste aura plus de chances d'observer un indice de peu d'importance qu'un indice très important ou prépondérant. Cette hypothèse s'inspire du principe de Pareto, qui dit que dans tout ensemble d'objets, idées, événements, etc., quelques éléments de l'ensemble

ont plus d'importance que la majorité restante. Autrement dit, **plus les faits sont importants, moins ils sont fréquents.**

b) L'arrivée d'un indice déterminant termine l'évaluation d'une proposition, soit en la rendant certaine, soit en l'éliminant. Notre problème consiste à décrire, pour l'ensemble de toutes les propositions, la population des indices qui ne sont pas déterminants, c'est-à-dire les faits incertains. Nous verrons plus loin (section 6.2) comment ceux-ci se combinent avec les indices déterminants.

En tenant compte des deux remarques précédentes, nous avons conçu un modèle de la forme illustrée par la figure suivante :

Figure 4 Distribution des indices par catégories



Le choix d'un modèle de population d'indices est laissé à la discrétion du concepteur. Il sera d'autant plus convaincant qu'il aura été validé sur un grand nombre d'analyses d'un même type de situations.

L'ensemble U des indices peut alors être décrit en fonction des catégories d'importance, c'est-à-dire, U est la réunion des sous-populations U_g des indices des catégorie g, g= 1, 2,…,6.

Chaque indice de U peut donc être classé dans une et une seule des catégories d'importance. On a :

N_g indices dans la catégorie g, de poids $w=w_g$, pour g=1, 2, …, 6
avec $N = N_1 + N_2 + \ldots + N_6$.

La définition d'une échelle de mesure ordinale du poids d'un indice et le modèle de la population des indices constituent notre schéma pour entreprendre l'évaluation d'une proposition. Nous montrons maintenant comment le poids final d'un ensemble d'indices et le degré d'adhésion direct correspondant sont estimés au moyen d'un échantillon d'indices.

5. L'ESTIMATION DU POIDS ASSOCIÉ À UNE PROPOSITION À PARTIR D'OBSERVATIONS

Les valeurs w_k étant fixées, notre modèle théorique pourrait fournir le poids total réel des indices appuyant directement une proposition, s'ils étaient tous connus de l'analyste. Dans l'analyse d'une situation réelle, on ne peut prétendre avoir eu accès à toute l'information, mais seulement à une partie. Le processus d'évaluation d'une proposition peut alors être décrit de la façon suivante : on commence par calculer l'impact d'un indice, s'en ajoute un second, puis un troisième, etc. À chaque ajout d'un indice, les calculs sont refaits. De façon formelle, cela revient à considérer une suite d'échantillons emboités d'indices E_1, E_2, …, E_n et à évaluer les résultats de chaque échantillon, en calculant le poids total des indices qui appuient la proposition d'intérêt.

Étant donné que notre objectif est de traduire en degrés d'adhésion directs les poids obtenus, la moyenne sera une mesure plus utile que le total car elle peut être convertie au moyen de la formule (8) sans problème de calcul numérique, ce qui ne serait pas le cas d'un total. Nous préférons donc estimer le poids moyen des indices pour cette raison.

Les sections suivantes décrivent le cheminement mathématique conduisant à la formule d'estimation désirée pour le poids moyen des faits appuyant directement une proposition: premièrement, nous construisons un plan d'échantillonnage; deuxièmement nous décrivons le paramètre théorique recherché; troisièmement, nous présentons l'estimateur du poids moyen des faits.

5.1 LE PLAN D'ÉCHANTILLONNAGE

Nous construisons maintenant un plan pour estimer à l'aide d'un échantillon aléatoire E de taille n le poids moyen des indices appuyant exactement une proposition A. Concrètement notre échantillon sera la réunion des sous-échantillons E_g des indices de chaque catégorie g, de taille n_g. Pour chaque indice k de la population U_g, nous adoptons le plan d'échantillonnage suivant :

> $\pi_{kg} \equiv P(k \in E, k \in g)$ est proportionnel à la proportion des indices de U_g qui ont un poids w_g.

En considérant le modèle de population que nous avons adopté précédemment, nous avons en effet seulement un nombre fini de valeurs possibles pour les poids, c'est-à-dire,

> $\{w_1, w_2, \ldots, w_6\}$, d'effectifs $\{N_1, N_2, \ldots, N_6\}$.

Notre plan équivaut donc à considérer pour chaque catégorie d'importance les chances qu'un indice provienne de celle-ci. Cependant un second phénomène intervient dans l'inclusion d'un indice dans l'échantillon. En effet, pour agir avec efficacité il suffit, en corrollaire de la loi de Pareto, de se concentrer sur les quelques éléments les plus importants d'un ensemble. Nous considérons donc qu'un analyste tendra à adopter ce comportement et **à négliger des indices de peu d'importance, même s'ils sont observés**. Ainsi, la probabilité conditionnelle

$$\phi_g \equiv Pr(k \in E \mid k \in g) \equiv n_g / N_g$$

est directement proportionnelle à l'importance de la catégorie d'indices g. Cette hypothèse nous permet de construire un modèle de comportement de l'analyste, que nous illustrons au tableau suivant.

Tableau 3 : modèle de comportement de l'analyste: taux d'échantillonnage de l'analyste par catégories

Catégorie	ϕ_g
(1) très peu important	0,15
(2) peu important	0,30
(3) moyennement important	0,60
(4) important	0,90
(5) très important	1
(6) prépondérant	1

Le plan d'échantillonnage est maintenant complètement défini, soit:

$$\pi_{kg} = \phi_g \times N_g/N,$$

pour chaque indice k de la catégorie g, $g=1, 2, \ldots, 6.$

Le tableau 4 donne les probabilités d'inclusion obtenues de la combinaison du modèle de population des faits de la figure 4 avec le modèle de comportement de l'analyste du tableau 3.

Tableau 4 : Probabilités d'inclusion dans E, par catégories

Catégorie g	$\pi_g = \phi_g \times N_g/N$
(1) très peu important	0,042
(2) peu important	0,102
(3) moyennement important	0,140
(4) important	0,090
(5) très important	0,034
(6) prépondérant	0,010

Ayant établi une probabilité d'occurence pour chaque indice de U, nous identifions maintenant la forme théorique du paramètre à estimer.

5.2 LE POIDS D'UN ENSEMBLE D'INDICES À L'APPUI D'UNE PROPOSITION

Notons $t_{w(A)}$, le poids total des indices qui appuient une proposition A de l'ensemble Ω des possibilités. Le calcul du poids total $t_{w(A)}$ s'obtient ainsi :

(10) $t_{w(A)} = \sum\{w_k(B); B=A; k=1, \ldots,N\},$

Le total $t_{w(A)}$ peut aussi être exprimé en fonction de la répartition des poids en catégories d'importance, c'est-à-dire,

(10') $t_{w(A)}=\sum\{w_g\{\sum c_{lg}; l=1,\ldots,N_g\}; g=1, \ldots,6\}$

avec c_{lg} = 1, si l'indice l du groupe g appuie exactement A, 0 sinon.

Nous pouvons calculer le poids moyen des indices qui appuient exactement A ainsi :

(11) $\mu_{w(A)} = t_{w(A)}/N.$

Procédons maintenant à l'estimation du poids moyen des indices appuyant exactement une proposition.

5.3. ESTIMATION DU PARAMÈTRE

À l'observation d'un indice k, l'analyste identifie son poids w_k et la proposition qu'il appuie. Au fil des observations se constitue un échantillon de mesures

$\{w_1(A_1), w_2(A_2),\ldots,w_k(A_k), \ldots, w_n(A_n)\},$

où $w_k(A_k)$ constitue le poids de l'indice k à l'appui de la proposition A_k.

Le problème de l'estimation du poids moyen résultant de l'addition des indices appuyant exactement A est considéré comme un problème d'estimation de domaine, c'est-à-dire, que l'estimateur est défini sur le sous-ensemble des indices qui pointent exactement vers la proposition A. La formule est la suivante :

(12) $M_{w(A)} = (1/N)\sum_g n_{gA} w_g/\pi_g$

où w_g est le poids de la catégorie g, g=1, …, 6

n_{gA} est le nombre d'indices de l'échantillon E_n qui sont de la catégorie g et qui appuient exactement A,

π_g est la probabilité d'inclusion d'un indice de poids g dans l'échantillon aléatoire E_n.

La formule (12) peut être ré-écrite en fonction du plan d'échantillonnage :

(12') $M_{w(A)}=(1/N)\sum g(w_g/\pi_g)\sum\{c_{lg}D_{lg}; l=1, \ldots, N_g\}$,

avec D_{lg} = 1, si lg est dans l'échantillon E_n, 0 sinon, pour l=1, …, N_g et c_{lg} =1, si le indice l du groupe g appuie exactement A, 0 sinon.

On peut montrer que (12') est un estimateur non biaisé en fonction du plan adopté, c'est-à-dire que son espérance mathématique

$E[M_{w(A)}] = \mu_{w(A)}$,

quelle que soit la valeur de $\mu_{w(A)}$.

Le rapport w_g/π_g dans l'équation (12) peut être interprété comme le coefficient de dilatation du poids initial d'un indice de catégorie g. Ce coefficient augmente avec l'importance de l'indice, comme le montre le tableau suivant :

Tableau 5 : Coefficients de dilatation des poids par catégorie d'importance

Catégorie	dilatation w/π
très peu important	2,47
peu important	3,50
moyennement important	4,96
important	13,38
très important	67,74
prépondérant	460,50

Le coefficient de dilatation constitue le poids final résultant de l'application du modèle de la population des indices. L'estimateur (12) mesure un poids moyen sur le domaine des indices qui pointent exactement vers A.

Le poids final des indices étant obtenu par notre procédure d'estimation, nous montrons comment utiliser l'estimateur pour calculer le degré d'adhésion associé.

6. LA TRANSFORMATION DES POIDS ESTIMÉS EN COEFFICIENTS

6.1. DEGRÉ D'ADHÉSION DIRECT ASSOCIÉ AU POIDS D'UN ENSEMBLE D'INDICES APPUYANT UNE PROPOSITION

En appliquant l'estimateur $M_{w(A)}$ à l'équation (8) avec $K = -1$, le degré d'adhésion direct $m(A)$ est obtenu :

$$(13) \quad m(A) = [1 - e^{-M_{w(A)}}].$$

Ce résultat correspond à la combinaison de tous les indices qui pointent exactement vers A. Nous rappelons que le choix d'estimer le poids moyen, plutôt que le poids total est motivé par des considérations de calcul numérique. Ce choix rend possible en tout temps le calcul de l'équation (13).

Puisque chaque indice (ou collection d'indices) est considéré comme une source d'information indépendante, l'ordre de combinaison des indices n'a pas d'effet sur le résultat final [Shafer, 1976: 62]. Ainsi le même résultat sera obtenu en appliquant l'équation (8) à chaque indice k pris individuellement, soit :

$$(14) \quad m_k(A) = [1 - e^{-(1/N)\, dil(k)}],$$

où $dil(k) = w_g/\pi_g$, si k appartient à la catégorie g, g=1, 2, …, 6,

et en combinant deux à deux les coefficients obtenus, au moyen de la règle de Dempster [Shafer, 1976: 76-78] présentée à la section 2.5.

Le résultat, par la règle de Dempster, de la combinaison deux à deux des degrés d'adhésion associés à chaque indice fournira une estimation du coefficient final. En pratique, il n'est donc pas nécessaire d'utiliser la formule (12). L'estimateur décrit par la formule (12) justifie cependant la forme de l'exponentielle dans la formule (14). C'est là toute l'utilité de la procédure d'estimation du degré d'adhésion direct.

6.2. LA COMBINAISON DES PROPOSITIONS ET LE CALCUL DU DEGRÉ D'ADHÉSION

Puisque l'ordre de combinaison des indices n'a pas d'effet sur le résultat final, la règle de Dempster peut être appliquée à tout moment pour combiner les degrés d'adhésion directs d'indices appuyant une même proposition ou deux propositions différentes. Un fait incertain peut alors être combiné avec un fait déterminant, ce qui aura pour effet, soit de terminer l'évaluation, soit de restreindre l'ensemble des possibilités. De plus, la contradiction entre les propositions est prise en compte par la règle de Dempster via l'équation (6) et l'indice de conflit est calculé selon l'équation (7).

Finalement le degré d'adhésion d'une proposition de Ω est obtenu en appliquant la formule (2).

6.3. DÉTERMINATION DE LA TAILLE DE LA POPULATION DES INDICES

Pour appliquer l'équation (14), il nous faut déterminer la taille de la population des indices. Bien qu'il soit impossible de connaître avec exactitude cette taille, en pratique, nous pouvons cependant en déterminer l'ordre de grandeur pour un type de problème particulier; en effet, puisqu'un analyste accorde plus d'attention aux indices d'importance, il lui sera plus facile de dénombrer seulement l'ensemble des éléments importants d'un type de problème plutôt que le nombre total d'éléments, c'est-à-dire les indices des trois catégories les plus importantes.

Notons n_g le nombre d'indices de la catégorie g qui sont observés et n_{imp} le nombre de faits dénombrés dans les trois catégories les plus importantes; étant donné que, par définition,

$$\phi_g \equiv Pr(k \quad E \mid k \quad g) \equiv n_g/N_g,$$

nous avons:

$$(15) \quad \pi_{imp} = \sum\{\pi_g ; g = 4,5,6\} = n_{imp} /N$$

En utilisant le plan d'échantillonnage du tableau 4 de la section 5.1, nous obtenons :

$$\pi_{imp} = (0,09+0,034+0,01) = 0,134.$$

Une évaluation de la taille de la population des indices est obtenue, soit :

$$(16) \quad N = n_{imp}/\pi_{imp} = n_{imp}/0,134.$$

Par exemple, une personne expérimentée dans la sélection des dossiers de vérification pourrait considérer que généralement 4 ou 5 indices d'importance interviennent dans l'analyse d'un dossier. En utilisant l'équation (16), la taille de la population des indices serait établie entre 30 et 37.

7. APPLICATION DE LA MÉTHODE

7.1. UN PROGRAMME, «L'ANALYSEUR»

Le modèle de représentation d'une population d'indices que nous avons présenté (tableaux 1 et 2, figures 3 et 4), la méthode d'estimation des poids des indices et leur traitement par la règle de Dempster ont été implantés dans un programme, l'Analyseur.

Le langage d'implantation est ICON version 8.6 [Griswold, 1990]. Il s'agit d'un langage du domaine public supporté par le Département d'informatique de l'Université de l'Arizona. Ce langage tourne présentement sur les plateformes suivantes : Amiga, Atari, MS-DOS, Macintosh, OS/2, CMS, VMS, Unix. L'Analyseur est disponible présentement dans sa version ß-test moyennement entente avec les auteurs. La présente implantation a pour but de vérifier l'intérêt du modèle dans le cadre des systèmes d'aide à la décision, de sorte que l'interface personne-machine minimal.

Dans la version actuelle du programme, l'utilisateur de l'Analyseur, que nous appelons l'analyste, peut situer la taille de son problème parmi trois classes qui lui sont proposées : petit problème, problème de moyenne envergure et problème faisant intervenir un grand nombre d'indices. Enfin, l'analyste a le choix d'ignorer le modèle pour fournir directement un coefficient numérique, ce qui revient selons nous à faire, pour chaque indice, l'exercice mental de construction d'un modèle.

Dans la prochaine version que nous prévoyons, l'utilisateur pourra aussi indiquer le nombre d'indices d'importance présumé, auquel cas la taille de population sera calculée au moyen de l'équation (16). L'analyste pourra aussi construire son propre modèle de population à partir d'une banque de cas déjà résolus.

La pondération des faits étant déterminée, l'analyste décrit ensuite l'ensemble des possibilités à prendre en considération. Il est alors prêt à présenter les indices au programme. Chaque indice est associé à une ou plusieurs hypothèses et son poids est indiqué. Dès l'introduction d'un second indice, la combinaison est effectuée au moyen de la règle de Dempster. Cette opération est répétée tant que des indices sont présentés au programme. L'analyste voit alors évoluer ses hypothèses à chaque nouvel indice qui est ajouté. L'Analyseur va fournir les résultats suivants :

- une description de chaque indice, le lien effectué avec une proposition ainsi que le poids accordé à cet indice;
- le résultat de la combinaison, par la règle de Dempster, d'une proposition avec le résultat précédent;
- un rapport contenant la liste des indices soumis au programme et le résultat final.

À tout moment l'analyste peut écrire le rapport de sa consultation ou encore intégrer un ou plusieurs rapports produits précédemment. L'intégration de rapports peut servir à «instruire» l'analyseur pour que les combinaisons à venir tiennent compte de certains facteurs.

7.2 UNE SESSION DE TRAVAIL AVEC «L'ANALYSEUR»

Voici la trace commentée d'une session de travail avec «l'analyseur» appliqué à un problème de fiscalité : dans le cadre de l'application de la Loi sur la Régie de l'assurance-maladie du Québec ou de l'application de la Loi sur la Régie des rentes, il s'agit de déterminer si un travailleur est salarié ou autonome. à titre d'exemple, nous avons extrait d'un jugement de la Cour la partie exposant les faits et un vérificateur d'expérience en fait l'évaluation au moyen de l'Analyseur.

_____

ANALYSEUR D'INDICES
Conception : C. Boivin et L.-C. Paquin
Implantatation en ICON : L.-C. Paquin

_____

PRÉFÉRENCES

1. Affichage des résultats : développement
2. Type de coefficient : catégorie assez d'indices (15 à 20)
3. Trace sur fichier : non
4. Valider

Votre choix : **4**

CHOIX DU FICHIER D'HYPOTHESES

1. Nouveau fichier
2. hyp_test.HYP
3. Quitter

Votre choix : **1**

Nom du fichier : **STATUT**

TAPER <n> POUR TERMINER D'AJOUTER

Identificateur d'hypothèse : **autonome**
Identificateur d'hypothèse : **salarié**
Identificateur d'hypothèse : **n**

STATUT.HYP

1. Afficher
2. Ajouter
3. Enregistrer
4. Retirer
5. Trier
6. Terminer

Votre choix : **6**

MENU PRINCIPAL

1. Indices
2. Hypothèses
3. Préférences
4. Quitter

Votre choix : **1**

INDICES

1. Intégrer un rapport
2. Localiser un fichier pour saisie
3. Saisie
4. Rapport
5. Terminer

Votre choix : **3**

TAPER <F> POUR UNE SAISIE SUR FICHIER;
<N> POUR TERMINER

Entrer un indice : **Le camion utilisé aux fins du transport des produits en question était la propriété du travailleur;**
Choix d'hypothèse(s)
1. autonome 2. salarié
Votre choix (- si négatif) : **1**
Catégorie d'importance :
1. Très peu important 2. Peu important
3. Moyennement important 4. Important
5. Très important 6. Prépondérant
7. Déterminant
Votre choix : **3**

Indice #1 : Le camion utilisé aux fins du transport des produits en question était la propriété du travailleur;
Tend vers : autonome
Poids : Moyennement important

TAPER <F> POUR UNE SAISIE SUR FICHIER;
<N> POUR TERMINER

Entrer un indice : **Aux termes du contrat, le travailleur s'est** engagé **à payer toutes les dépenses d'opération, d'entretien et d'assurances de son camion;**
Choix d'hypothèse(s)
1. autonome 2. salarié
Votre choix (- si négatif) : **1**
Catégorie d'importance :
1. Très peu important 2. Peu important
3. Moyennement important 4. Important
5. Très important 6. Prépondérant
7. Déterminant
Votre choix : **3**

Indice #2 : Aux termes du contrat, le travailleur s'est engagé à payer toutes les dépenses d'opération, d'entretien et d'assurances de son camion;
Tend vers : autonome
Poids : Moyennement important

	Masse	Degré d'adhésion	Plausibilité
1 0	0.048361	0.048361	1.0
1 1	0.951639	1.0	1.0

Indice de contradiction : 0

TAPER <F> POUR UNE SAISIE SUR FICHIER;
<N> POUR TERMINER

Entrer un indice : **En contrepartie, la requérante s'est engagée à payer un maximum de 675$ / semaine;**
Choix d'hypothèse(s)
1. autonome 2. salarié
Choix (- si négatif) : **2**
Catégorie d'importance :
1. Très peu important 2. Peu important
3. Moyennement important 4. Important
5. Très important 6. Prépondérant
7. Déterminant
Votre choix : **4**

Indice #3 : En contrepartie, la requérante s'est engagée à payer un maximum de
675$ / semaine;
Tend vers : salarié
Poids : Important

	Masse	Degré d'adhésion	Plausibilité
1 0	0.043947	0.043947	0.908735
0 1	0.091265	0.091265	0.956053
1 1	0.864788	1.0	1.0

Indice de contradiction : 0.004617

TAPER <F> POUR UNE SAISIE SUR FICHIER;
<N> POUR TERMINER

Entrer un indice : **n**

7.3. INTERPRÉTATION DES RÉSULTATS

Un logiciel d'aide à la décision doit bien sûr fournir au décideur une règle pour utiliser les résultats. Nous adoptons ici une proposition de Williams [1990] et utilisons comme critère de décision le rapport entre les valeurs de plausibilité de deux hypothèses A et B:

$$(17) \quad R(A,B) = Pl(A)/Pl(B).$$

Étant donné la relation entre le degré d'adhésion et la plausibilité, établie par l'équation (4), nous pouvons dire que la plausibilité d'une hypothèse décroît avec l'accumulation d'indices contre celle-ci. Il s'agit d'un raisonnement par élimination. Au début d'une analyse, toutes les hypothèses sont considérées comme certainement plausibles. L'accumulation d'indices en faveur d'une hypothèse a alors pour effet de rendre moins plausible son contraire, tend à l'éliminer. Le rapport de plausibilité $R(A,\neg A)$ sera d'autant plus élevé que A est plausible relativement à son contraire.

Il peut cependant arriver qu'un rapport de plausibilité entre deux hypothèses soit élevé sans pour cela qu'une décision puisse être prise. C'est le cas des situations où plusieurs indices contradictoires sont présents. Il convient d'utiliser l'indice de conflit Con_n (non implanté) obtenu au moyen de l'équation (7) pour construire le rapport suivant [Almond, 1989: 15]:

$$(18) \quad Rcon = Con_n/(1-Con_n).$$

Ce rapport nous permet d'évaluer l'importance des faits contradictoires relativement aux indices qui se renforcent. Il sera utile de suivre l'évolution du rapport de conflit lorsque des révisions d'une analyse sont effectuées.

7.5. INTÉRÊT DE L'ANALYSEUR

L'Analyseur peut être utilisé comme un système d'identification de connaissances ou comme un outil d'aide à la décision. En tant que système d'aide à la décision, l'Analyseur facilite la lecture des indices et montre pas à pas l'évolution des hypothèses. Il est actuellement expérimenté pour efectuer le classement de textes dans le cadre d'un projet de *Conception d'un système expert pour l'aide à l'analyse (tri, classification et indexation) des documents de jurisprudence* [Bertrand-Gastaldy 1992]. Cette recherche, en cours de réalisation, bénéficie d'une subvention du CEFRIO (Centre francophone de recherche en informatisation des organisations) à laquelle contribuent à la fois SOQUIJ (Société

québécoise d'information juridique) et le ministère des Communications du Québec dans le cadre du projet Delta.

Une enquête cognitive effectuée auprès des conseillers juridiques de la SOQUIJ nous a révélé que le modèle cognitif à l'oeuvre lors de l'opération de classement d'une décision est celui d'une lecture rapide (peu de temps est en effet consacré à cette tâche). Ceux-ci cherchent à retrouver dans le texte certains indices qui leur permettent de classer les décisions selon un plan de classification qui comporte 57 domaines de droit, subdivisés de façon inégale en raison du nombre inégal de jugements traitant des différents sujets. Les indices discriminants sont de valeur différente. Certains sont prépondérants comme le tribunal où la décision est rendue (par ex.: *La Cour supérieure division criminelle*), l'intitulé du jugement (par ex.: *Déclaration de culpabilité*) ou les lois citées par le juge (par. ex.: la *Loi sur les stupéfiants*). Toutefois, certains de ces indices pointent vers un sous-ensemble de domaines (par ex.: Le *Code de la route* pointe vers les domaines de responsabilité, d'assurance et de transport). Par contre, en l'absence de tels indices, le recours à des indices terminologiques est nécessaire. Ces indices terminologiques sont ordonnés en quatre classes d'importance selon qu'ils sont composés d'un seul ou de plusieurs mots et qu'ils appartiennent aux outils documentaires (thesaurus et plan de classification) ou qu'ils ont été identifiés comme termes du domaine par les experts eux-mêmes.

Deux ordres de difficultés plaident en faveur d'un système de classification basé sur la théorie des fonctions de confiance. D'une part, l'absence de distribution égale des textes dans les classes rend difficile les analyses classiques de distribution. D'autre part, il nous est impossible *a priori* de constituer un univers de référence qui permettrait de distribuer la certitude sur l'ensemble des indices. La solution qui reste est de traiter indépendamment chacun des indices et d'opérer un cumul rapporté sur le ou les descripteurs vers lequel pointe l'indice. Le design est ainsi fait qu'il est possible de calibrer par essai erreur l'attibution d'un poids aux différentes classes indices.

Comme système d'identification de connaissances, il permet de rassembler efficacement des informations provenant de multiples sources (interviews, observations personnelles, textes longs, etc.) pour documenter une question. En utilisant le système sur un grand nombre de situations d'un même type (ex. : sélection d'un dossier pour vérification fiscale), on disposera d'une banque de rapports qui pourront être étudiés afin d'en tirer des règles à inclure dans un système expert. Une analyse permettra alors d'évaluer si un développement de système expert est concevable; s'il y a convergence d'informations vers des règles, le

système est possible. Sinon, l'analyseur aura montré qu'un système expert ne peut être envisagé.

Une première expérience est commencée en fiscalité pour évaluer la faisabilité d'un système expert de détermination du statut de travailleur (salarié ou autonome). Pour cela nous avons extrait de 25 jugements de la Cour les parties exposant les faits. Quatre vérificateurs d'expérience en font l'évaluation au moyen de l'Analyseur. Ces évaluations constituent la banque de connaissances qui sera étudiée par la suite.

L'Analyseur est aussi un outil de formation. Dans le domaine fiscal, par exemple, il pourrait servir à former le personnel moins expérimenté, en préparant des exercices sur des points nébuleux de la Loi.

Nos croyons que l'Analyseur sera applicable à une foule de situations : indexation des textes, diagnostic de maladies, décisions administratives à caractère répétitif (sélectionner un dossier pour vérification, établir s'il y a eu négligence, choisir parmi différents scénarios de gestion, etc.), enquêtes policières, analyse de marchés boursiers, etc. Rappelons que, pour que la méthode soit applicable, les hypothèses doivent être mutuellement exclusives et les faits indépendants.

8. CONCLUSION

Nous avons proposé un procédé pour évaluer la certitude à accorder à une proposition. En utilisant la notion «d'importance d'un indice» associée à la notion de «coefficient», une première simplification est effectuée. En utilisant des expressions verbales pour qualifier le poids d'un indice, une seconde simplification est réalisée.

Puisqu'on n'a pas explicitement recours à des données, chaque évaluation devient subjective. La théorie de fonctions de confiance a été retenue comme cadre théorique pour représenter ces évaluations et les combiner. Le choix de cette théorie plutôt que la théorie des probabilités est motivé par notre connaissance du milieu de travail où nous évoluons; en effet, bien que des données soient disponibles, c'est souvent à un coût et avec des délais jugés trop élevés. Entre disposer d'une information complète et d'aucune information, les fonctions de confiance apparaissent comme un compromis intéressant. C'est en somme la loi du moindre effort pour un rendement maximum qui est appliquée.

Suivant Shafer et Tversky [1985], nous avons envisagé la théorie de fonctions de confiance comme un langage formel, avec un vocabulaire, une syntaxe et une sémantique. L'échelle des catégories d'importance des indices et le modèle de population constituent notre vocabulaire; la méthode d'estimation des poids et les règles de calcul de la théorie de Dempster-Shafer font la syntaxe. La sémantique demeure à développer, afin d'assurer une homogénéité dans l'interprétation des indices.

La méthode d'analyse d'événements, et son implantation dans l'Analyseur, peut être étendue à l'étude d'un ensemble de concepts qu'un expert aura mis en relation. Les concepts deviennent les noeuds d'un réseau tissé par les règles d'association établies entre ceux-ci. L'Analyseur permettrait alors de constituer un journal des informations pertinentes à l'évaluation de chaque noeud. Nous envisageons en quelque sorte un réseau d'analyseurs, dont les résultats de chaque noeud pourront être combinés au moyen d'un algorithme de propagation des coefficients dans l'ensemble du réseau [Shenoy et Hsia, 1989; Almond, 1991]. Dans le cas du calcul exact des coefficients, le défi consiste à améliorer la performance des algorithmes existants. En pratique cependant, la solution pourrait résider dans une méthode de calcul par approximation ou dans l'exploitation du parrallélisme.

RÉFÉRENCES

Almond, R. G. (1989). *Fusion and Propagation of Graphical Belief Models: an Implementation and an Example*, Ph.D. Thesis and Technical Report S-130, Harvard University, Department of Statistics.

Almond, R., (1991). "Building blocks for graphical belief models". *Journal of Applied Statistics*, (18), 1.

Bertrand-Gastaldy, S., Daoust, F., Meunier, J.-G., Pagola, G. et Paquin, L.-C. 1992 "Un prototype de système expert pour l'aide à l'analyse des jugements", *Actes du congrès Informatique et droit* de l'AQDIJ, volume C1.3.

Cheeseman, P., Critique de l'article de Lauritzen, S. and Spiegelhalter, D. J. (1988). Local computation with probabilities on graphical structures and their application to expert systems (with discussion). J. R. Statistist. Soc. B, No. 2, pp. 157-224.

Dempster, A. P., (1967). "Upper and Lower probabilities induced by a multivalued mapping". *Annals of Math. Statistics*, (38), 2 : 325-339.

Dempster, A. P., (1989). Commentary on the two precedings papers : commentary on (1) "subjective probability and causality assessment" by David A. Lane, and (2) "Current research directions in the development of expert systems based on belief network" by Gregory F. Cooper. *Applied Stochastic Models and Data Analysis*, (5): 77-81.

Remerciements à Serge Garceau et Robert Paré pour leurs commentaires sur une version antérieure de ce texte, ainsi qu'à Jacques Fortin et Jean Lamontagne pour l'exemple d'application.

Griswold, R. et Griswold, M., (1990) *The Icon Programming Language*, Prentice Hall, Englewood Cliffs, New Jersy.

Dubois, D. et Prade, H., (1987). *Théorie des possibilités. Applications à la représentation des connaissances en informatique*, Masson, Paris.

Shafer, G., (1976). *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, New Jersey. 296 p.

Shafer, G., (1982). "Belief functions and parametric models". *Journal of Royal Statistical Society B*, (44), 3 :. 322-652.

Shafer, G., (1987). "Probability Judgement in Artificial Intelligence and Expert Systems" (with discussion). *Statistical Science*, (2), 1 : 3-44.

Shafer, G., et Tversky, A., (1985). "Languages and Design for Probability Judgements". *Cognitive Science*, (9) : 309-339.

Shenoy, P. P. et Hsia, Y.-T., (1989). "An evidential language for expert systems". *Methodologies for Intelligent Systems*, 4 (Z. Ras, ed.), Elsevier Science Publishing co.

Williams, P., (1990). "An interpretation of Shenoy and Shafer's axioms for local computation". I*nternational Journal of Approximate Reasoning* 4, pp. 225-232.