

Mathématique *et* *Pédagogie*

Sommaire

- *J. Navez, Éditorial* 2
- *GEM 1er degré, Représenter et se représenter des situations dans l'espace* 3
- *R. Gérardy, Approche de la théorie de la communication. Hasard ou certitude : l'information* 9
- *D. Odiet, L'horloge de la gare de Mons* 27
- *A. Bajart, Sous-ensembles flous, logique floue et applications* 36
- *G. Haesbroeck, Modélisation du comportement d'un individu face à des revenus aléatoires* 54
- *F. Clette, J.-R. Gabryl, Une courte nuit pour un grand jour : le 11 août 1999!* 68
- *M. Fremal, Revue des revues* 76
- *C. Festraets, Olympiades* 80
- *C. Festraets, Des problèmes et des jeux* 86

Éditorial

J. Navez,

L'actualité en Belgique est évidemment concentrée sur les prochaines élections législatives. Tous les partis politiques peaufinent leur programme dans lequel on retrouve forcément le volet enseignement. Leurs promesses me semblent si belles qu'il faut croire que la situation va nettement s'améliorer à partir du 14 juin. Si, dans nos lecteurs, certains ont quelque influence politique, je les prie de rappeler aux décideurs que le problème qui nous paraît être particulièrement épineux dans l'enseignement des mathématiques au secondaire est le programme commun au second degré.

L'actualité à l'étranger, c'est surtout la guerre dans l'ex-Yougoslavie, l'exode des réfugiés, les massacres d'innocents, les bombardements et leurs "dommages collatéraux". On installe des écoles dans les camps de réfugiés ; c'est peut-être, comme le disent certains médias, pour "occuper" les enfants, mais c'est aussi parce que le besoin d'éducation et de formation est primordial, même dans les pires situations.

Jacques NAVEZ

Représenter et se représenter des situations dans l'espace

GEM 1er degré, Université Catholique de Louvain

Mots-clés : ligne de visée, perspective cavalière, fuyantes, orientation, projections orthogonales.

Le sous groupe 1er degré du GEM a conçu une séquence d'activités sur la géométrie dans l'espace, qui a été expérimentée en partie dans des classes de 2ème durant l'année scolaire 97-98. Cette séquence a été proposée aux participants du congrès de Floreffe en août 98.

L'objectif principal de ces activités est l'élaboration progressive de deux techniques de représentation plane : la perspective cavalière et les projections orthogonales coordonnées.

Dans un premier temps, pour favoriser le développement de l'intuition spatiale, il est souhaitable de proposer aux élèves du matériau concret, tangible.

1. Reconstituer un paysage

Les élèves reçoivent plusieurs développements présentés (à échelle réduite) à la Figure 1 et sont invités à construire les maisonnettes correspondantes. Ils disposent de deux photographies d'un bâtiment construit par assemblage de 5 maisonnettes et doivent reconstituer ce bâtiment ainsi que des éléments du décor (arbre, mare, boîte aux lettres) sur un support fourni (Fig. 2).

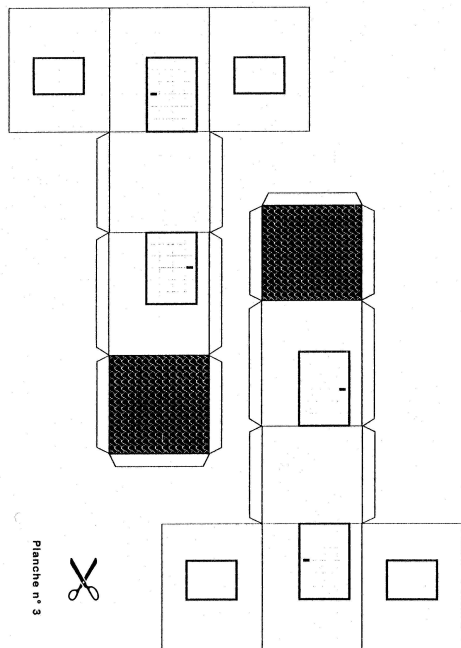


Fig. 1

Les élèves passent donc de représentations planes familières (photos) à une maquette. Les difficultés résident dans l'orientation de chaque maisonnette (remarquez le décentrage des portes) et dans le positionnement des éléments du décor. On achemine alors les élèves vers les concepts de lignes de visée et d'angle de visée : ils repèrent sur leur maquette les différentes positions à partir desquelles un personnage (matérialisé par un pion) pourrait voir un élément de décor à l'exclusion d'autres. Ils représentent ensuite ces zones sur des exemplaires du support.

2. Représenter un paysage

L'objectif de cette deuxième activité, basée sur un échange de messages, est de faire naître un mode de représentation plane facile à réaliser et facile à décrypter. Les élèves, répartis en groupe, sont invités à construire

un nouveau bâtiment à l'aide de 5 maisonnettes sur un nouveau support (Fig. 3), et à réaliser les dessins nécessaires pour qu'un autre groupe puisse construire un bâtiment identique. Comme le support est fixé, les variations ne résident que dans les orientations respectives des maisonnettes. On pourrait ouvrir davantage la situation en ne fournissant pas de support.

Les représentations esquissées par les élèves s'inspirent souvent de la technique de perspective cavalière et quelquefois de celle des projections orthogonales. L'échange entre les groupes favorise une prise de conscience de lacunes ou ambiguïtés dans les dessins.

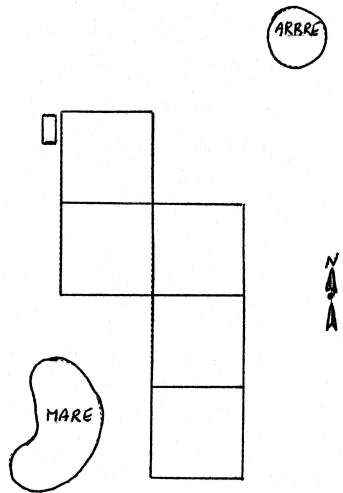


Fig. 3

On peut alors faire une petite synthèse à propos de la différence principale entre le fonctionnement de notre perception visuelle, dont se rapprochent celui des prises photographiques et la technique des représentations planes utilisées : dans le premier cas, le parallélisme n'est pas conservé, dans le second oui. On peut préciser la technique de représentation par perspective cavalière : pour représenter un cube, un dessinateur propose un dessin comme celui de la Figure 4.

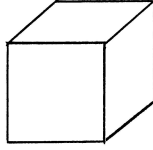


Fig. 4

Il représente

- les arêtes situées dans le plan vertical avant en respectant les longueurs et les angles formés entre elles,
- les arêtes perpendiculaires à ce plan vertical par des segments obliques parallèles entre eux, qui portent le nom de fuyantes.

Dans ce cas, le cube est représenté par 3 faces : l'une est représentée par un carré et les deux autres par des parallélogrammes. (Voici une bonne occasion de revoir les quadrilatères, dans un autre contexte).

3. Représenter une maisonnette en perspective cavalière

L'objectif de cette activité est

- de faire naître un questionnement sur la direction des fuyantes et la longueur de celles-ci,
- de faire prendre conscience de la conservation de la proportionnalité sur ces fuyantes.

On demande aux élèves de représenter deux fois une maisonnette en perspective cavalière, la première avec une porte de face, la seconde avec une fenêtre vue de face.

Lors de la synthèse, on peut confronter différentes représentations et aboutir à un consensus quant à la direction des fuyantes et à leur facteur de rétrécissement. Dans les classes, nous avons choisi 45° pour l'inclinaison des fuyantes par rapport à l'horizontale : ce choix facilite les dessins sur papier quadrillé (où on repère aisément les diagonales inclinées à 45°) mais il peut parfois amener des ambiguïtés dans la superposition de certains traits.

Pour la longueur des fuyantes, après avoir examiné différentes possibilités, nous avons choisi de représenter un cm de la réalité par 0,7 cm (plus précisément $\frac{1}{\sqrt{2}}$, c'est-à-dire la demi-diagonale d'un carré de côté 1). De cette manière, la représentation d'un cube est assez vraisemblable et facile à réaliser.

Une fois cette unité choisie, les élèves doivent donc tenir compte des dimensions des portes et des fenêtres de la maisonnette pour les transformer adéquatement sur les dessins des faces latérales. En réalité, les hauteurs ne sont pas modifiées ; seules les largeurs doivent être adaptées proportionnellement à leurs mesures initiales. Les solutions sont présentées à la Figure 5.

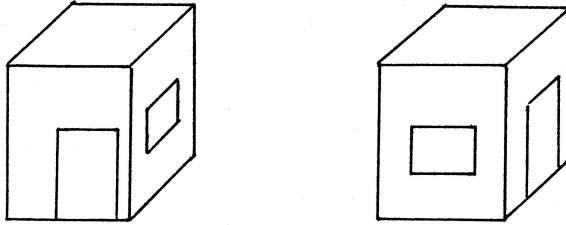


Fig. 5

4. Voir une boîte dans sa tête et la représenter

Ici, les élèves ne disposent plus de matériau concret. Ils reçoivent une représentation (Figure 6) et le message suivant :

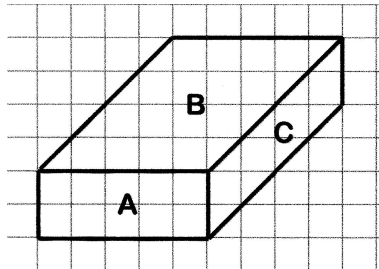


Fig. 6

“Voici une boîte de dimensions $4 \times 2,5 \times 1$ (en cm). Elle est dessinée suivant les conventions établies. De la même façon, dessinez d’autres représentations de cette boîte en prenant les rectangles B ou C comme vues

de face, en vraie grandeur.” L’objectif est de provoquer un décryptage des conventions établies juste avant, suivi d’une nouvelle application de celles-ci.

Dans cette activité, intervient aussi la prise en compte de l’orientation intrinsèque de l’objet : les faces A , B et C sont disposées dans un certain ordre : autour du sommet commun à ces 3 faces, on passe successivement de A à B et de B à C en tournant dans le sens des aiguilles d’une montre. Il doit en être ainsi sur les autres représentations. Ceci suscitera la mise au point des conventions relatives aux arêtes cachées.

5. Projections orthogonales

D’autres activités ont été conçues pour les projections orthogonales : lecture de plans d’architecte, construction de projections orthogonales (dites coordonnées) d’un objet disponible ; analyse de 3 projections orthogonales d’un objet non disponible pour y repérer des points, des arêtes, des faces.

Si vous êtes intéressés par cette séquence sur la géométrie dans l’espace, un fascicule, reprenant les fiches d’activité et les commentaires à leur sujet, est disponible au prix de 100 BEF augmenté des frais de port à l’adresse des auteurs :

GEM (sous groupe 1er degré)
Département de mathématiques
Chemin du Cyclotron 2
1348 Louvain-la-Neuve

Approche de la théorie de la communication. Hasard ou certitude : l'information

R. Gérardy,

Texte relatif à la conférence donnée lors du Congrès SBPMef 1997 à Ciney.

On peut définir l'information par tout ce qui est échangé dans une communication sous quelque forme que se soit : suite de symboles, lettres, chiffres, sons, images, odeurs, gestes, ... En principe, cette suite est porteuse de sens, tout au moins pour le destinataire du message transmis ; toutefois, comme le dit Claude Shannon "les aspects sémantiques de la communication ne relèvent pas de ses aspects technologiques" et ce sont précisément ces derniers qu'il a étudiés dans sa "Théorie mathématique de la communication" publiée en 1946 sous forme de trois articles dans le *Bell System Technical Journal* puis repris en 1949 avec Warren Weaver sous le titre général de *The Mathematical Theory of Communication*. L'information réduit donc l'incertitude initiale en apportant une quantité de données nouvelles, quant au fond ou à la forme. Cette quantité d'information fournie par un signal ou un message émanant d'une source déterminée se mesure en "bits" et est désignée par Shannon par le terme "entropie". Celle-ci augmente lorsque le nombre des éventualités augmente (le jeu de pile ou face donne une entropie de 1 bit tandis qu'un jeu de 32 cartes en fournit 5) ; d'autre part, l'entropie diminue si les probabilités sont inégales et à la limite, l'entropie sera nulle si une des éventualités a une probabilité égale à 1 (il n'y a aucune incertitude, donc aucune information !). S'il y a répétition de données ou simplement corrélation entre les éléments du message transmis, on parlera alors de redondance, ce qui aura pour effet de diminuer la quantité d'information reçue (la présence de la lettre "u" après la lettre "q" n'a pas le même poids informationnel que si le "u" suit un "m" ou un "r" ; dans le premier cas, la probabilité d'occurrence du "u" est proche de 1, il n'y a donc pas d'effet de surprise à trouver la lettre "u" après la lettre "q") ; la redondance est en quelque sorte l'anti-entropie, c'est ce qu'il faut éviter dans un télégramme par exemple, compte tenu du coût de transmission : "nous arriverons demain vendredi 12 décembre" sera avantageusement remplacé, vu les indications de service se trouvant sur tout télégramme, par "arriverons demain". Par contre, les codes et systèmes redondants sont utilisés afin d'éviter les erreurs dans les circuits électroniques ou autres afin de parer aux pannes éventuelles. Signalons par exemple le procédé de détection d'erreurs CRC (Contrôle de redondance cyclique) utilisé tant par un modem émetteur que par un modem récepteur et s'appliquant à chaque bloc

de données transmis ; pour les modems les plus rapides, c'est le protocole ARQ (Automatic repeat request) qui est utilisé et qui refuse toutes les données supplémentaires jusqu'à ce que le bloc défectueux soit corrigé. La redondance est indispensable dans le langage oral car celui-ci ne permet pas un retour en arrière ou une vérification a posteriori, elle permet donc une meilleure compréhension du discours. D'autre part, lorsqu'un texte est mal écrit, c'est également la redondance, propre à chaque langue, qui aide à la compréhension du message (pensez à l'ordonnance du médecin dont l'écriture laisse souvent à désirer!).

Un message transmis par un canal peut aussi être perturbé par certains bruits ou modifications altérant le message lui-même et pour lequel une incertitude apparaîtra au double point de vue de l'émetteur ou du récepteur. L'incertitude que l'émetteur a que son message ait été correctement transmis s'appelle l'ambiguïté (il s'agit d'une probabilité conditionnelle désignée par $H(B/A)$, A étant le message émis et B le message reçu); quant à l'équivoque, représentée par $H(A/B)$, elle correspond à l'incertitude qui demeure sur le signal après qu'il ait été reçu. Il s'agit donc de mesures de l'imperfection du canal.

La mesure de l'information a déjà été formulée par Ralph Hartley dès 1928; mais Ronald Fisher (biologiste et statisticien anglais : 1890-1962) s'est intéressé, dès 1920, à la notion d'information en tant que concept de précision des estimations des paramètres utilisés; toutefois ses travaux ne furent publiés qu'en 1934 (Probability, likelihood and quantity of information in the logic of uncertain inference - Royal Society) et le point de vue de Fisher était différent de celui de Shannon.

1. Mesure de la quantité d'information

Dans un système disposant de S signes équiprobables, un message comprenant n de ces signes, contient une quantité d'information égale à :

$$H = n \log s$$

formule que l'on peut généraliser à un ensemble E comprenant n signes équiprobables, de probabilité $p = 1/n$, et dont l'émission de chaque signe apportera une quantité d'information égale à :

$$H = K \log n = -K \log(1/n) = -K \log p$$

la constante K dépend notamment de la base des logarithmes choisie : si K est la constante de Boltzmann, on retrouve au signe près la mesure de l'entropie thermodynamique.

Claude Shannon a ainsi introduit le terme d'entropie (informationnelle) ou de négentropie (entropie négative) pour mesurer la quantité d'information transmise par une ligne télégraphique.

C'est John W. Tukey qui a proposé à Claude Shannon le terme de "bit" dont la double définition est :

- a) symbole numérique (binary digit : 0 ou 1) ;
- b) unité d'information (binary unit : logon : shannon : hartley).

Le bit est donc à la fois la plus petite unité d'information et le chiffre élémentaire de la numération binaire (permettant précisément de mesurer cette information!).

Une suite finie de bits s'appelle un byte (multiplet comprenant 4, 6 ou 8 bits) dont le plus couramment utilisé est l'octet (8 bits) comprenant parfois un bit de parité.

Remarquons que l'apport d'information est surtout lié à un problème de combinatoire car au départ d'un alphabet donné, connu de l'émetteur comme du récepteur, ce sont les différentes façons d'assembler les lettres qui constituent les éléments d'un message, dans une langue déterminée, selon en particulier les règles de la sémantique et de la syntaxe, mais aussi selon les aptitudes et la personnalité de chacun.

2. Codes et information

Pour identifier un objet, un élément d'un ensemble, pour transmettre un message ou simplement pour se faire comprendre, nous utilisons un code (répertoire de signes conventionnels) qui peut, selon le cas, paraître hermétique à certains et qui parfois l'est effectivement afin d'en limiter son usage ou sa compréhension.

Quelques exemples relevant de différents domaines

1. ouvrage d'une bibliothèque : COM 1 WEA 6796008-11

(il s'agit en l'occurrence de "Théorie mathématique de la communication" de Shannon et Weaver, édition CEPL, se trouvant à l'Université de Liège);

la CDU (classification décimale universelle de Melvil Dewey) est encore utilisée dans les fichiers analytiques des bibliothèques.

2. numéros de sécurité sociale; numéros du registre national; plaque d'immatriculation; etc ...
3. numéro de téléphone : 04/382 19 33 (code national)

le détail donnera :	04	zone de Liège
	3	rive droite de la Meuse
	82	secteur de Sprimont
	1933	numéro de l'abonné relié au central n° 1

toutefois certaines variantes existent, par exemple si l'abonné souhaite disposer de plusieurs lignes (voire de plusieurs centraux) avec éventuellement des implantations différentes.

4. des codes que tout un chacun est censé connaître ou utiliser : le code de la route, le code postal, le code pénal, le code civil, etc ...
5. le "langage silbo" des bergers des Iles Canaries (sifflements conventionnels).
6. le "langage" des sourds-muets, l'alphabet dactylogique de Juan Pablo Bonnet (1620).
7. - différents alphabets : arabe : 30 lettres, braille : 64 signes plus les annonceurs et les notations du code de Nemeth, cyrillique : 33 lettres en russe, mais les symboles chimiques sont désignés par les lettres de l'alphabet latin, espéranto : 28 lettres, grec classique : 24 lettres, hébreu : 22 lettres, latin avec parfois des graphies supplémentaires (26 lettres en français et en anglais mais 28 en espagnol, 21 en italien et 27 en allemand), phénicien : 22 lettres;
- écriture chinoise (idéographique);
- écriture japonaise utilisant à la fois les idéogrammes (il faut en connaître environ 2000 pour lire un journal), les syllabes hiragana (mots japonais) et katakana (mots étrangers) et les lettres rōmaji (lettres romanes : k, m); les chiffres sont les mêmes que les nôtres;
- des alphabets télégraphiques : de Chappe, de Morse, de Baudot, de Gauss, de Murray;
- l'alphabet phonétique international : A = alpha, B = bravo, C = Charlie, ...

-
-
8. le code des couleurs utilisé pour identifier les condensateurs et les résistances en électronique : noir = 0, marron = 1, rouge = 2, orange = 3, ...
9. le code “Q” utilisé en radiotélégraphie, en radiotéléphonie et en Scott (signaux optiques) afin d’éviter les transmissions trop longues résultant de l’usage de l’alphabet Morse (presque totalement abandonné) et de créer en outre un langage maritime et aérien international :
 QAN : direction et vitesse du vent, QBA : visibilité horizontale, QCX : indicatif d’appel complet, etc ...
10. le code SINPO :
 Code servant à noter sur 5 la qualité d’une liaison en télécommunications selon les notes suivantes :
 5 = excellent, 4 = bon, 3 = assez bon, 2 = passable, 1 = médiocre, 0 = mauvais.
 La note *S* juge la force du signal reçu (Strength), le *I* : l’absence de brouillage (Interference), le *N* : l’absence de bruit (Noise), le *P* : l’absence de défauts dus à la propagation : évanouissements, parasites (Propagation disturbance) et le 0 l’intelligibilité globale (Overall readability).
 Le code SINPO est parfois remplacé par le code SINPFEMO où s’ajoutent des notes sur la fréquence du fading, la qualité et la profondeur de la modulation.
11. le MIRACODE (Microfilm Information Retrieval Access Code de Kodak) consistant en une indexation d’un microfilm par un damier de rectangles noirs et blancs lu par un jeu de cellules photoélectriques et interprété en numération binaire (plus guère usité).
12. le code ASCII (American Standard Code for Information Interchange) permettant l’échange d’informations traitées par un ordinateur ou par un système de transmissions et s’effectuant sur 8 bits dont 7 réservés à l’information (128 combinaisons) plus un huitième de contrôle ; si le 8ème n’est pas utilisé comme bit de parité, on dispose alors de 256 combinaisons, il s’agit du code ASCII étendu. Ce code créé en 1963 par Robert Bemer, a évolué depuis pour s’adapter à d’autres langues que l’anglais.
 A l’origine, la répartition était la suivante :

Code ASCII	groupes de codes
de 0 à 31	codes de contrôle
de 32 à 173	caractères de type “texte”
de 174 à 255	caractères semi-graphiques

Actuellement, d'après les standards internationaux, le groupe de caractères de type texte s'arrête au code 127, les codes 128 à 255 étant réservés aux caractères graphiques.

Les applications sous "DOS" (Disk Operating System) utilisent le code ASCII tandis que le code ANSI (American National Standard Institute) est réservé aux applications Windows. La différence entre les deux tables n'apparaît qu'à partir du code 128, notamment pour les caractères accentués (le "à" est codé 133 en ASCII [Alt +133] et 224 en ANSI [Alt +0+224] - [cf aussi le code EBCDIC (Extended Binary Coded Decimal Interchange Code), il s'agit d'un code à 8 bits utilisé par IBM pour ses gros ordinateurs (il diffère du code ASCII et n'est pas utilisé en micro-informatique)].

13. les codes détecteurs et/ou correcteurs ; codes bips - POST (mise en route d'un PC).
14. le code de Huffman permettant la compression des fichiers (codes courts pour les caractères fréquents et codes longs pour les caractères rares) – code de Fano-Shannon – etc ...
15. le code à barres : constitué d'une série de barres verticales parallèles dont l'épaisseur et l'espacement sont variables, cette série étant lue par un scanner qui transmet l'information alphanumérique à un ordinateur. Plusieurs normes coexistent, notamment : UPC (Universal Product Code) servant à l'identification des produits dans la grande distribution ; EAN (European Article Number) comprenant 13 chiffres correspondant au numéro du pays, à celui du fabricant et à un code de contrôle.
16. différents codes binaires, pondérés ou non, comme le code de Gray utilisé notamment pour la correction automatique des erreurs dans les télécommunications (ce code est dû à Frank Gray des Laboratoires Bell et fut introduit dans les années 40).

Les suites de chiffres utilisées dans certains codes ne sont évidemment pas des nombres et il est inutile de leur appliquer une quelconque propriété ou une opération usuelle (l'addition de 2 numéros de téléphone n'a guère de sens!). Par contre, il peut être intéressant de savoir combien de nombres il est possible d'écrire en utilisant 1, 2, 3, 4, ... n chiffres parmi les dix chiffres disponibles de notre système décimal (ou dans un autre système de numération) ; ou, d'autre part, d'établir combien de mots il est théoriquement possible de construire en utilisant 1, 2, 3, 4, 5, 6, ... n lettres parmi les 26 de notre alphabet (mots ayant ou non un sens en français ou

dans une autre langue), ce qui nous permettra d'approcher la notion de quantité d'information transmise en utilisant tel alphabet ou tel système de numération.

3. Recherche sur les nombres

Combien de nombres N de 1, de 2, de 3, de 4, ... de p chiffres peut-on écrire en utilisant tous les n chiffres disponibles dans un système de numération déterminé ?

$$N = n^p \quad \text{et} \quad p = \log_n N.$$

Par exemple, dans le système binaire : $n = 2$ (chiffres disponibles : le 0 et le 1); d'où $N = 2^p$ et $p = \log_2 N$.

4. Recherche semblable avec un système alphabétique

Avec les 26 lettres de notre alphabet, contrairement aux nombres, très peu de solutions d'écriture de "mots" sont valables, que ce soit en français ou dans une autre langue (le dictionnaire de l'Académie française ne comprend que 35 000 mots!), alors que l'on pourrait théoriquement écrire 26 "mots" d'une lettre, 676 "mots" de 2 lettres, 17576 mots de 3 lettres, ..., 308915776 "mots" de 6 lettres, ...

Ceci nous amène au problème de la fréquence d'utilisation de chaque lettre dans une langue déterminée, recherche particulièrement utile en cryptographie.

Dès la plus haute antiquité, le secret couvrait déjà la transmission d'un message d'une personne à une autre, que ce soit oralement ou par écrit.

La lecture des hiéroglyphes n'était-elle pas réservée à quelques initiés ?

Certains renseignements sur les Egyptiens, les Phéniciens, les Grecs et les Romains nous sont parvenus notamment grâce à Hérodote (450 avt J.C.) et à Plutarque (50 avt J.C.) qui nous décrit le premier la scytale des Lacédémoniens (il s'agit d'un bâton calibré autour duquel on enroulait

en oblique une lanière de cuir ou un ruban de papyrus sur lequel le message était inscrit ; pour reconstituer le texte, il fallait disposer d'un bâton semblable).

Ce système constitue vraisemblablement la première machine à *transposition* (changement de l'ordre des lettres d'un message).

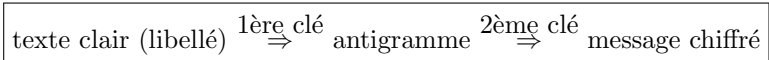
Le second système utilisé est celui dû à Jules César (selon Suétone), il se base sur le remplacement d'une lettre par une autre (principe de substitution).

Dans la manière de dissimuler un message à transmettre, nous pouvons distinguer la stéganographie qui vise à cacher l'existence même du texte rédigé en clair, comme par exemple l'usage des encres sympathiques ou des micropoints ; il en est de même pour les conventions qui consistent à relever la première lettre d'une suite de mots formant un texte apparemment anodin (pensez aux acrostiches chers à François Villon !) ou encore la transmission à vitesse élevée d'un texte donné. Ce procédé se retrouve également dans des dessins où des traits longs ou courts (code Morse) ont des apparences anodines, d'autres recourent au jeu de la perspective (anamorphose : *Les Ambassadeurs* de Hans Holbein, 1533) et à notre époque la transmission par Internet d'images dont certains pixels ont été modifiés ! Il en va tout autrement du chiffrement qui ne cache nullement l'existence d'un message secret mais qui le rend inintelligible en transformant un texte clair ou libellé, notamment par transposition ou par substitution, en un texte chiffré ou cryptogramme.

Le déchiffrement consiste à reconstituer le libellé tout en connaissant la clé utilisée.

Si cette recherche s'effectue sans connaissance préalable de la clé, on parle alors de décryptement : travail effectué par un cryptologue-décrypteur ou cryptanalyste.

Si l'on veut effectuer un surchiffrement, c'est-à-dire chiffrer à nouveau le message déjà chiffré, ce dernier sera appelé antigramme et donnera naissance, à l'aide d'une deuxième clé, à un nouveau message chiffré :



La clé utilisée n'est pas nécessairement numérique, elle peut être formée d'un ou de plusieurs mots. Les lettres utilisées seront alors numérotées dans l'ordre alphabétique et dans l'ordre d'apparition dans le mot.

4.1. Chiffrement par transposition

Ce procédé consiste à changer l'ordre des lettres d'un message selon une clé donnée. En guise d'exemple, considérons le message en clair ou libellé suivant :

LES MEMBRES DE LA COMMISSION SONT TENUS AU DEVOIR DE RESERVE.

Clé numérique directe utilisée (7 chiffres) : 5 2 7 1 4 6 3

Disposition pratique :

lecture en clair \Rightarrow

5	2	7	1	4	6	3
L	E	S	M	E	M	B
R	E	S	D	E	L	A
C	O	M	M	I	S	S
I	O	N	S	O	N	T
T	E	N	U	S	A	U
D	E	V	O	I	R	D
E	R	E	S	E	R	V
E						

Chiffre :
- colonne 1
 \Downarrow - colonne 2
- colonne 3
etc ...

4.2. Chiffrement par substitution

Ce procédé consiste à remplacer une lettre du libellé par une autre en utilisant par exemple un alphabet décalé (méthode de Jules César), une "roue à chiffrer" formée de deux disques concentriques (alphabet en clair à l'extérieur ; alphabet aléatoire à l'intérieur [mobile ou non] : système dû à Charles Wheatstone [redécouverte d'un instrument conçu à la Renaissance] et présenté en 1867 à l'Exposition Universelle de Paris) ou encore le carré "indéchiffrable" de Blaise de Vigenère (1525-1596). Ce carré de Vigenère (table à double entrée) est en fait une invention de l'abbé Trithème.

Exemple :

Message en clair : QUINZE FUSEES

Clé utilisée : MARINE

Disposition pratique en utilisant la grille de Vigenère ci-dessous :

clé	M A R I N E M A R I N E
message clair	Q U I N Z E F U S E E S
message chiffré	C U Z V M I R U J M R W

A l'intersection de la colonne *M* (alphabet de la clé de chiffrement) et de la ligne *Q* (alphabet du message en clair) se trouve la lettre *C* du message chiffré (recherche semblable pour les autres lettres); le déchiffrement s'effectuera en sens inverse.

Tableau de l'abbé Trithème ou grille de Vigenère

	alphabet du message en clair												alphabet de la clé de chiffrement													
	↓												←													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

En raison de la symétrie du tableau, les alphabets du texte clair et de la clé peuvent être permutés.

Ces procédés de transposition et de substitution, auxquels s'ajoutait l'usage de répertoires ou de codes, se sont perfectionnés au fil des siècles en ayant notamment recours à des machines de plus en plus complexes dont la plus célèbre est certes l'ENIGMA, brevetée en 1919 et due au hollandais Hugo Alexandre Koch de Delft. Celui-ci céda ses droits à un mécanicien berlinois Arthur Scherbius qui y apporta de nombreuses améliorations, la dénomma Enigma puis céda les brevets à Alexander von Kryha, ingénieur ukrainien établi à Berlin ; cette machine devait permettre de garder secrètes les transactions financières et commerciales ; mais le colonel Erich Fellgiebel du service des transmissions de la Reichswehr comprit que ses fabuleuses possibilités (22 millions de combinaisons différentes avant de se répéter)

pourraient être utiles à l'armée allemande qui l'utilisa durant la deuxième guerre mondiale.

C'est à Bletchley Park (à la Government Code & Cypher School), près de Londres, qu'une équipe de mathématiciens et de techniciens anglais, polonais et français, sous la direction de Turing et de Welchman, arriva à décrypter les messages que l'armée allemande transmettait grâce aux Enigmas ; c'est là que furent construits les fameux Colossus, ancêtres de nos ordinateurs.

A l'heure actuelle, la cryptologie se base essentiellement sur l'étude des nombres premiers et la factorisation des grands nombres, tels les systèmes PGP (Pretty Good Privacy de Philip Zimmermann : clé de 128 caractères), DES (Data Encryption Standard : clé unique) ou RSA (Rivest, Shamir, Adleman : système asymétrique à 2 clés).

5. Vers une définition

INFORMATION :

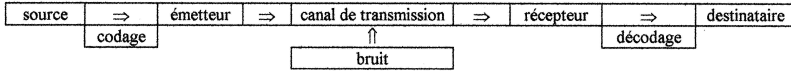
FORMULATION ABSRAITE DECRIVANT UN ELEMENT
(SITUATION - PROPRIETE - EVENEMENT) D'UN
SYSTEME STRUCTURE

CE QUI PEUT ETRE TRANSMIS PAR UN SIGNAL OU UNE
COMBINAISON DE SIGNAUX

CE QUI REDUIT L'INCERTITUDE

UNE AUTRE FACE DU HASARD

6. Schéma de la communication proposé par Shannon



Ce schéma fut élaboré en tenant compte des problèmes particuliers qui se posaient à la Bell Company, c'est-à-dire de transmettre avec une rentabilité satisfaisante tout type de messages, ce qui amena Claude Shannon à opter pour une position probabiliste où la fréquence des éléments constitutifs des messages transmis prenait toute son importance. Ce point de vue est présent notamment dans le cas de la compression de données : méthode de Fano - Shannon ou méthode de Huffman, par exemple. Ce schéma évolua par la suite selon les applications envisagées (en biologie, en économie, en informatique, en sociologie, en psychologie et en linguistique notamment) ; il a suscité de nombreuses controverses et encore maintenant est parfois malmené : voir par exemple "La science de l'information" de Le Coadic (PUF) pour qui celle-ci est devenue une "science sociale rigoureuse" et qui n'hésite pas à parler de "confusion conceptuelle" ou "d'imposture" pour ce qui était au départ et selon son auteur une "Théorie mathématique de la Communication", en particulier télégraphique ! Claude Shannon était ingénieur électricien, il devait résoudre des problèmes d'ordre technique et concevoir un modèle mathématique correspondant, ce qu'il a fait avec toutes les précautions nécessaires quant à l'aspect sémantique, voire épistémologique, relatif à toute communication. Ne lui reprochons pas ce que d'autres ont compris (ou non) de sa théorie, ni l'évolution ou les adaptations qui sont intervenues au fil des années et dont il n'est nullement responsable ; quant au choix du terme "information", certains ont pensé qu'il était mal choisi mais comme le signalait Benoit Mandelbrot aux "Colloques philosophiques de Royaumont" [1965], il est trop tard pour changer de terminologie et d'autre part "on peut se demander si la théorie de l'information aurait joué le grand rôle qui a été le sien si elle avait porté un nom moins alléchant". Claude Shannon fut également professeur au M.I.T. et publia des articles sur la cryptographie, la construction de codes télégraphiques, la compression des données ou la correction automatique des erreurs dans un message ; il a étendu sa théorie au cas des variables continues qui, en généralisant, peut être obtenu par un passage à la limite du cas discret.

Une autre conception est notamment celle d'Andrei Kolmogorov qui adopte une démarche algorithmique et introduit la notion de complexité : l'information pouvant s'envisager comme une mesure de celle-ci. La complexité d'une suite (aléatoire) de chiffres ou de lettres correspond au nombre minimal de bits qu'il faudra utiliser dans un programme déterminé pour obtenir cette suite à l'aide d'une machine universelle (type Turing). Par exemple, obtenir 2400 décimales du nombre π avec un programme de 158 caractères en langage *C* laisse rêveur quant à la notion de complexité au sens de Kolmogorov.

7. Remarques complémentaires

1) Pour Claude Shannon

INFORMATION = LOGARITHME DU NOMBRE DE CHOIX
= ENTROPIE

Référentiel	Entropie
jeu de pile ou face	1 bit // ($2^1 = 2$)
jeu de 32 cartes	5 bits // ($2^5 = 32$)
damier de 64 cases	6 bits // ($2^6 = 64$)
un dé (6 faces)	2,58496250072... bits // ($2^{2,58496250072} = 6$)
jeu de 52 cartes	5,7004397181... bits
système binaire	1 bit // ($2^1 = 2$)
système ternaire	1,5849625007... bit
système octal	3 bits // ($2^3 = 8$)
système décimal	3,32192809489... bits
système duodécimal	3,58496250072... bits
système hexadécimal	4 bits // ($2^4 = 16$)
alphabet de 26 lettres	4,70043971814... bits
alphabet de 33 lettres	5,04439411936... bits

Ce tableau nous donne la quantité d'information nécessaire pour déterminer un élément appartenant au référentiel choisi [quantités discrètes ou discontinues] (ensemble des possibilités d'obtenir pile ou face, de choisir une carte dans un jeu de 32 ou de 52 cartes, une case d'un damier ou une face d'un dé; de désigner un nombre appartenant à un système de

numération déterminé ou une lettre d'un alphabet ...) en tenant compte d'une équiprobabilité théorique pour chaque élément (ce qui n'est pas toujours le cas).

Le calcul du logarithme d'une probabilité amena Shannon à faire précéder sa formule du signe "moins" afin de rendre positive la valeur de l'entropie calculée pour des quantités discontinues.

2) Comparaison entre l'alphabet français et l'alphabet russe

- alphabet russe de 33 lettres (supposées équiprobables) :
l'entropie sera égale à :

$$H_R = -\log_2(1/33) = \log_2 33 = 5,04439411936 \text{ bits};$$

- alphabet français de 26 lettres (supposées également équiprobables!) :
l'entropie sera égale à :

$$H_F = -\log_2(1/26) = \log_2 26 = 4,70043971814 \text{ bits}.$$

Rapport $H_R/H_F = 5,04439411936/4,70043971814 = 1,07317494146$, soit un coût de transmission de 7 % plus élevé en russe!

3) Application au départ d'un code dont on connaît les probabilités d'occurrence respectives : code bilitère : A et B

Probabilité d'occurrence (p) :	A	0,6
	B	0,4

Quantité d'information (I) liée à l'utilisation de chaque lettre :

$I(A) = -\log_2 0,6 = 0,737 \text{ bit}$
--

$I(B) = -\log_2 0,4 = 1,322 \text{ bit}$
--

Rassemblons ces éléments :

	A	B
Probabilité d'occurrence :	0,6	0,4
Quantité d'information :	0,737	1,322

Le “poids” informationnel du B est plus important que celui du A ; ce qui signifie que la réalisation d’un événement rare a une valeur informationnelle plus grande, ou l’événement le plus fréquent est le moins informatif !

Recherche de l’information moyenne (H) ou de l’entropie (néguentropie de Shannon) :

$$H = p_A \cdot (-\log_2 p_A) + p_B \cdot (-\log_2 p_B).$$

Pour l’exemple ci-dessus, on a :

$$H = 0,6 \cdot 0,737 + 0,4 \cdot 1,322 = 0,971 \text{ bit}.$$

D’une manière plus générale, pour n signes, nous mesurons l’entropie de la source par :

$$H = - \sum_{i=1}^n p_i \cdot \log p_i.$$

4) Comme l’écrit Aimé Fuchs (La loi normale et la théorie de l’information in “Le hasard” - Dossier Pour la Science - Belin - page 57) :

“La théorie de l’information, qui a pour vocation la mesure du degré de complexité d’un système, s’est développée dans les années 1940, à partir des travaux d’avant-garde de Claude Shannon et de Norbert Wiener ... Supposons que l’on étudie un système caractérisé par une variable aléatoire X , admettant un nombre n de valeurs discrètes avec les probabilités $p_1, p_2, p_3, \dots, p_n$. L’entropie statistique mesure le degré d’indétermination du système, ou la quantité d’information à laquelle on peut accéder. Pour remplir ce rôle, l’entropie doit posséder quelques propriétés particulières, dont les suivantes :

a) elle est nulle quand l’un des événements est certain (on sait tout sur le système), c’est-à-dire quand $p_i = 1$, les autres probabilités étant nulles ;

b) elle est maximale dans le cas de l’équipartition, c’est-à-dire lorsque toutes les probabilités p_i sont égales (on ne sait rien sur le système) ;

c) l’entropie d’un couple de variables aléatoires indépendantes est égale à la somme des entropies des variables individuelles (lorsqu’un événement est composé de 2 événements indépendants, sa probabilité est égale au produit des 2 probabilités individuelles)”.

5) De nombreux points de vue se sont affrontés depuis la parution des travaux de Claude Shannon, mais une mention particulière doit être faite aux notions d’algorithme et de complexité dues notamment au mathématicien soviétique Andreï Kolmogorov (le père de l’axiomatisation de la théorie des probabilités) (cf aussi Bennett C. H. et Chaitin G. J. in Delahaye J. P. - Information, complexité et hasard - Hermès - 1994).

Dans “Le fascinant nombre π ” (Belin - 1997 - page 95), Jean-Paul Delahaye illustre cette conception par un programme de 158 caractères en langage C, d’auteur inconnu, permettant d’obtenir 2400 décimales :

```
int a=10000,b,c=8400,d,e,f[8401],g;main(){for(;b-c;)
f[b++]=a/5;for(;d=0,g=c*2;c=14,printf(“<<%4d>>,e+d/a),
e=d%a)for(b=c;d+=f[b]*a,f[b]=d%-g,d/=g--,--b;d*=b);}
```

$\pi = 3, 141\ 592\ 653\ 589\ 793\ 238\ 462\ 643\ 383\ 279\ 502\ 884\ 197 \dots$

ce qui peut s’écrire en binaire :

11,001 001 000 011 111 101 101 010 100 010 001 000 101 101 ...

8. Quelques protagonistes :

- Ralph Hartley (1888 - ?) - ingénieur à la Bell Company ; a montré, en 1927, comment mesurer l’information à l’aide d’une unité baptisée d’abord “hartley” puis bit (contraction de binary unit) - en 1950, Denis Gabor a introduit le terme de “logon” : unité d’information mesurant la capacité des canaux de transmission utilisés.
- Andreï Kolmogorov (1903-1987) - mathématicien soviétique, axiomatisation de la théorie des probabilités - notion de complexité - démarche algorithmique.
- Harry Nyquist (1889-1976) - théorie mathématique des amplificateurs électroniques ; première idée de la notion d’information sélective (1924).
- Claude Elwood Shannon (Gaylord, Michigan - 1916 - ?) - ingénieur électricien aux Laboratoires Bell à New York où il étudie les problèmes de rendement des lignes télégraphiques - professeur au M.I.T. - démarche probabiliste.
- Warren Weaver (1894-1978) - coauteur avec Shannon de la “Théorie mathématique de la Communication”.

– Norbert Wiener (1894-1964) - professeur de mathématique au
Massachusetts Institute of Technology - créateur de la cybernétique.

Adresse de l'auteur :
Robert GERARDY
rue Cléchêne 12
4140 SPRIMONT

L'horloge de la gare de Mons

D. Odiet,

Cet article est également paru dans la revue suisse Math-Ecole, 185, décembre 1998, pp. 35–40.

“Une grande part du temps scolaire est consacré à apprendre de nouvelles connaissances et à s'exercer à leur utilisation. Reste-t-il du temps pour développer le plaisir de chercher, pour lancer aux élèves des sortes de défis intellectuels dont l'objectif se trouve dans l'activité plus que dans le résultat ? Du temps aussi pour développer les attitudes et les connaissances nécessaires à cette activité, du temps pour essayer, pour se tromper, pour discuter une solution avec d'autres, pour tenter de convaincre, pour mettre en doute une méthode ou une réponse ?” (1)

Dans le numéro 179 de *Math-Jeunes* (2), Claude Villers propose, sous le joli titre de “Quelle heure est-elle ?”, un problème présentant les aspects essentiels cités plus haut. Le voici :

Au moment de pénétrer dans le vaste bâtiment de la gare, Maxime jeta rapidement un regard interrogatif sur la grande façade vitrée. C'est qu'elle était dotée d'une grande horloge au design très simplifié. Douze gros points marquaient les emplacements des heures et deux aiguilles lumineuses se chargeaient de renseigner les voyageurs. Bien que grossière, la lecture de l'heure du moment pouvait se faire à la minute près.

Maxime constata ainsi qu'il disposait encore de six minutes avant le départ de son train. Tout de suite, il se retrouva instantanément (3) dans le hall d'attente où il jeta un coup d'oeil vers l'horloge visible par transparence.

“Tiens, se dit-il, si je n'y prenais garde, je croirais disposer maintenant de trois heures et demie d'attente”.

Quelle heure est-il à l'horloge de la gare ?

Quelle est l'heure de départ du train que Maxime veut emprunter ?

1. “Pourquoi des mathématiques à l'école ?”, Roland Charnay, ESF Editeur, 1996.

2. MATH-JEUNES, revue trimestrielle publiée par la Société Belge des Professeurs de Mathématique d'expression française.

3. *derechef* dans le texte original

Tout problème mathématique lié à une montre, notamment aux positions des aiguilles sur son cadran, plus précisément à l'angle formé entre celles-ci, me ramène inexorablement à ma période scolaire lycéenne. Mon professeur de mathématique d'alors, fervent adepte d'un enseignement de type purement transmissif, en était particulièrement friand, tellement d'ailleurs qu'il avait réussi à installer parmi plusieurs de ses élèves un véritable sentiment de dégoût. Il avait beau s'évertuer à asséner son schéma de résolution, sa méthode infaillible, son "Il suffit de faire comme ça!", peine perdue, chaque nouveau problème d'horloge suscitait systématiquement pour beaucoup d'entre nous angoisse, impossibilité de mobiliser la recette du prof, donc invitation à passer directement au problème suivant.

Bientôt vingt ans ont passé . . . Trouvé lors de quelques paisibles lectures de vacances, un extrait de Marc Legrand ⁽⁴⁾ résume à lui seul ce pittoresque professeur : "Dans le Supérieur plus encore que dans le Secondaire, on rencontre de telles personnes parfois passionnées par leur discipline ; des élèves ou des étudiants, ou plus exactement des auditeurs, sont nécessaires à leurs discours, mais au fond ce sont plus des conférenciers que des professeurs, car ils n'ont aucun souci de ce qui se passe réellement dans la tête de leurs interlocuteurs. Certains d'entre eux d'ailleurs déclarent sans ambages ne pas vouloir le savoir ; ils sont là, disent-ils, pour enseigner leur discipline...".

Grâce à l'horloge de la gare de Mons, l'heure à la réconciliation avec de tels énoncés a enfin sonné, si j'ose dire . . . Entendons-nous : il me paraît absolument évident que ce type de problème bien spécifique invite à la mise en oeuvre d'outils de résolution bien personnels, reflets de la compréhension, de la perception et de la conceptualisation ô combien différentes que chacun d'entre nous s'en fait. Il est donc illusoire de vouloir imposer un type de méthode à ses élèves. L'enseignant doit en être conscient. Les différentes productions d'élèves de degré 9, n'ayant pas les mêmes compétences mathématiques, en sont une illustration tangible.

Au fait, avez-vous trouvé l'heure de départ du train ?

Nul doute que la précision de l'énoncé du problème – c'est là aussi un de ses côtés remarquables – n'est pas étrangère à votre réussite. Une seule petite ambiguïté réside éventuellement dans "l'expression" et la notation de l'heure : est-ce exact de dire que si l'on retranche trois heures à 1h de l'après-midi (13h), il est 10h du matin ? Ceci n'a en tous les cas aucunement perturbé les élèves, bien désireux de ne pas ériger un obstacle supplémentaire.

4. "La crise de l'enseignement : un problème de qualité", Marc Legrand, ALEAS Editeur, 1989.

Pour eux, la condition d'infériorité d'une heure par rapport à une autre rime avec le recul des aiguilles.

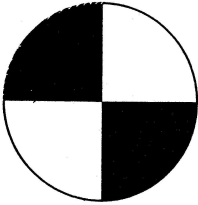
Vous allez probablement retrouver votre démarche parmi celles proposées ci-dessous ⁽⁵⁾.

L'horloge de la gare de Mons		Julien et Sébastien	
On s'est dit qu'il fallait utiliser du papier calque. C'est plus facile pour s'imaginer l'heure depuis l'intérieur du hall. On est parti dans le côté droit en haut de l'horloge. Après quelques essais, on s'est aperçu que l'heure soit se situer entre 13h00 et 14h00.			
Heure devant la gare	Départ du train	Heure hall	Différence
13h40	13h46	10h20	3h26
13h41	13h47	10h19	3h28
13h42	13h48	10h18	3h30
<u>Réponse</u> : il est 13h42 ; le train part à 13h48.			

L'emploi du papier calque prend ici une dimension non dépourvue de sens : n'est-ce pas le support inévitable de l'introduction de la symétrie axiale ?

5. Les travaux présentés ont été "relookés", mais sont l'exact reflet des productions des élèves.

La démarche de Yohann et de Luc est plus détaillée ...

<u>L'horloge de la gare de Mons</u>	Yohann et Luc																																				
<p>1) Nous avons cherché les zones impossibles pour la petite aiguille pour l'heure actuelle et extérieure. Nous avons découvert deux zones impossibles : ■</p> <p>2) Nous avons cherché les zones possibles pour la grande aiguille. Nous avons remarqué qu'elle doit être près du quart (15) ou du trois-quart (45). La moitié de six minutes : 3 minutes. $15 - 3 = 12 \dots$ $45 - 3 = 42 \dots$ symétrie axiale ... 48 symétrie axiale... 18 $48 + 30 = 1h18$ $18 + 30 = 48$ $12 + 6 = 18$ $42 + 6 = 48$ Cela prouve que l'aiguille des minutes peut se trouver sur 12 (heure actuelle) Cela prouve que la grande aiguille peut se trouver sur 42 (heure actuelle)</p>																																					
<p>3)</p> <table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left; padding: 2px;">heure actuelle (1)</th> <th style="text-align: left; padding: 2px;">heure par par symétrie (2)</th> <th style="text-align: left; padding: 2px;">(2) + 3h30</th> <th style="text-align: left; padding: 2px;">(1) + 6 min</th> </tr> </thead> <tbody> <tr><td>1h12</td><td>10h48</td><td>2h18</td><td>1h18</td></tr> <tr><td>1h42</td><td>10h18</td><td>1h48</td><td>1h48</td></tr> <tr><td>2h12</td><td>9h48</td><td>1h16</td><td>2h18</td></tr> <tr><td>2h42</td><td>9h18</td><td>0h48</td><td>2h48</td></tr> <tr><td>7h12</td><td>4h48</td><td>6h16</td><td>7h18</td></tr> <tr><td>7h42</td><td>4h18</td><td>7h48</td><td>7h48</td></tr> <tr><td>8h12</td><td>3h48</td><td>7h18</td><td>8h18</td></tr> <tr><td>8h42</td><td>3h18</td><td>6h48</td><td>8h48</td></tr> </tbody> </table>	heure actuelle (1)	heure par par symétrie (2)	(2) + 3h30	(1) + 6 min	1h12	10h48	2h18	1h18	1h42	10h18	1h48	1h48	2h12	9h48	1h16	2h18	2h42	9h18	0h48	2h48	7h12	4h48	6h16	7h18	7h42	4h18	7h48	7h48	8h12	3h48	7h18	8h18	8h42	3h18	6h48	8h48	
heure actuelle (1)	heure par par symétrie (2)	(2) + 3h30	(1) + 6 min																																		
1h12	10h48	2h18	1h18																																		
1h42	10h18	1h48	1h48																																		
2h12	9h48	1h16	2h18																																		
2h42	9h18	0h48	2h48																																		
7h12	4h48	6h16	7h18																																		
7h42	4h18	7h48	7h48																																		
8h12	3h48	7h18	8h18																																		
8h42	3h18	6h48	8h48																																		
<p>4) Les heures de départ sont : 1h48 et 7h48.</p>																																					

Celle de Dominik et Nicolas est très subtile. Comment ne pas la rapprocher de celle très gracieusement présentée par un ancien membre du comité de Math-Ecole ?

L'horloge de la gare de Mons

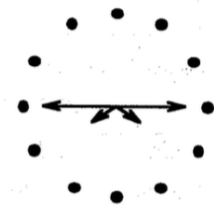
Nous avons fait un axe de symétrie d'axe $12h - 6h$.

Ensuite, nous avons tracé un écart de $1h45$ de chaque côté.

En enlevant 3 minutes, nous obtenons l'heure de la gare : $7h42$. Le train part à $7h48$.

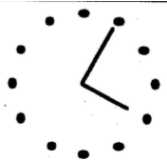
De l'autre côté, il est $4h18$.

Dominik et Nicolas



Quand il est
4h05, après midi,
devant la gare

il "est"
4h05, avant midi,
dans le hall



$$\begin{aligned} \text{l'écart horaire est de} \\ \mathbf{4h05 + 4h05} \\ = \\ \mathbf{8h10} \end{aligned}$$

inversement

$$\begin{aligned} \text{Si l'écart considéré est de} \\ \mathbf{3h24} \\ = \\ \mathbf{1h42 + 1h42} \end{aligned}$$

Il est donc
1h42, après midi,
devant la gare

et il "est"
1h42, avant midi,
dans le hall



L'emploi de l'outil équation était évidemment attendu.

Ici, équation à une variable

MC	<u>L'horloge de la gare de Mons</u>	Brice et Raphaël
	Heure lue depuis l'extérieur	Heure lue depuis l'intérieur
	1	11
	2	10
	3	9
	...	
	x	$-1(x - 12)$
Nous avons trouvé une équation.		
	unité heure	unité minute
heure lue depuis l'extérieur	x	x
heure lue depuis l'intérieur	$-1(x - 12)$	$-1(x - 720)$
heure de départ du train	$x + 0,1$	$x + 6$
	$x + 0,1 = -1(x - 12) + 3,5$	$x + 6 = -1(x - 720) + 210$
	$x + 0,1 = -x + 12 + 3,5$	$x + 6 = -x + 720 + 210$
	$2x = 15,4$	$2x = 924$
	$x = 7,7$	$x = 462$
$7,7\text{h} = 7\text{h}42\text{ min} = 462\text{ min}$		
Réponse : il est 7h42 min et le train part à 7h48 min.		

...ici, à deux variables :

L'heure de la gare de Mons

Tarik et Ignace

Quand on regarde l'horloge par transparence, c'est en fait une symétrie axiale. Si l'aiguille des heures est placée sur "h", elle sera par transparence sur "11 - h". L'aiguille des minutes "m" sera placée sur "60 - m".

$h;m \longrightarrow 11 - h; 60 - m$

Ensuite, l'heure indiquée sur la 1ère montre doit donner la même heure moins 3h24 par transparence.

$h;m \longrightarrow h - 3; m - 24$

On peut en déduire les deux équations suivantes :

$$\begin{array}{l} 11 - h = h - 3 \quad \text{et} \quad 60 - m = m - 24 \\ 7 = h \quad \quad \quad 42 = m \end{array}$$

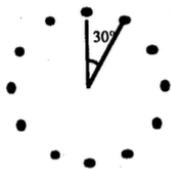
$$7h + 42m = 7h \ 42\text{min}$$

Il est 7h42, le train part à 7h48. De l'autre côté, il semble être 4h18.

Le travail de Pauline et Lorraine est vraiment surprenant et bien révélateur de ce qui peut se passer dans la tête de certains de nos élèves.

L'horloge de la gare de Mons

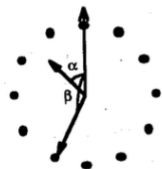
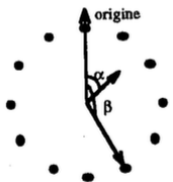
Pauline et Lorraine



En 1 h, la petite aiguille avance de 30° ;
 En 1 min, la petite aiguille avance de $0,5^\circ$ ($30 : 60$).
 En 1 h, la grande aiguille avance de 360° ;
 En 1 min, la grande aiguille avance de 6° .

A l'extérieur de la gare

A l'intérieur de la gare



Heure lue : $(\alpha ; \beta)$

Heure lue : $(360^\circ - \alpha ; 360^\circ - \beta)$

Heure de départ du train : $(\alpha ; \beta) + 6 \text{ min}$
 $(\alpha + 6 \cdot 0,5^\circ ; \beta + 6 \cdot 6^\circ)$
 $(\alpha + 3^\circ ; \beta + 36^\circ)$

Heure lue de l'intérieur + 3h30 = heure de départ.

1) Petite aiguille : $(360^\circ - \alpha) + 3,5h \cdot 30^\circ = \alpha + 3^\circ$
 $360^\circ - \alpha + 105^\circ = \alpha + 3^\circ$
 $\alpha = 231^\circ$

$$2) \text{ Grande aiguille : } (360^\circ - \beta) + 180^\circ = \beta + 36^\circ$$

$$\beta = 252^\circ$$

(en 3h30, la grande aiguille parcourt 1260° , elle se “déplace” donc de 180°).



Petite aiguille (231°) : il est entre 7h et 8h.

Grande aiguille (252°) :

en 1 min $\rightarrow 6^\circ$

en ? min $\rightarrow 252^\circ$

$$256 : 6 = 42$$

Vérification : en 42 min, la petite aiguille parcourt 21°

$$7 \cdot 30^\circ + 21^\circ = 231^\circ$$

Conclusion : il est 7h42, Maxime doit prendre son train à 7h48.

Ah oui, j’allais oublier ... L’horloge de la gare de Mons, la voici !



Adresse de l’auteur :

Denis ODIET

Marchands 2

CH-2900 Porrentruy

Suisse

Sous-ensembles flous, logique floue et applications

A. Bajart, *Electronique de Commutation*
Université de Liège

Mots-clés : logique floue, sous-ensembles flous.

La logique floue est une théorie assez récente, extension de la logique binaire classique. Les valeurs de vérité qui appartiennent traditionnellement à l'ensemble $\{0, 1\}$ sont étendues à l'intervalle $[0, 1]$ tout entier, permettant ainsi une modulation de la connaissance. Une proposition peut dès lors être à la fois vraie et fausse tout comme un élément d'un ensemble A peut appartenir à A et à son complémentaire A_C . Un exemple particulièrement illustratif est celui du classement d'une population en deux groupes : les petits et les grands. A quelle hauteur fixer la limite ? De plus, la connaissance de la taille exacte des individus à classer est requise. Pourtant, l'être humain est capable de formuler une réponse à la question "x est-il petit ou grand ?", et plus encore de la moduler, sans disposer précisément de ces deux informations. C'est cette richesse de raisonnement que l'on cherche à mimer avec la logique floue. De nombreuses applications témoignent aujourd'hui de son intérêt, dans des domaines aussi variés que les sciences humaines, exactes ou appliquées. Cet article reprend les grandes lignes de l'exposé présenté lors du Congrès de la SBPMef, le lundi 24 août 1998.

1. Introduction à la logique floue

1.1. Introduction

Une des caractéristiques du raisonnement humain est qu'il est généralement fondé sur des données imprécises ou même incomplètes. Ainsi, déterminer si une personne est de petite ou grande taille est aisé pour n'importe lequel d'entre nous, et cela sans nécessairement connaître sa taille exacte. De plus, il nous est facile de moduler notre réponse : assez grand, assez petit... Pour procéder à cette même classification, un ordinateur, lui,

doit non seulement connaître la taille exacte mais également disposer d'un algorithme, algorithme divisant inmanquablement une population en deux groupes rigides : les grands et les petits. Supposons que la limite soit à 1,70 m. Un individu mesurant 1,69 m peut-il vraiment être qualifié de petit ?

Le premier à avoir souligné ces possibilités de développement du raisonnement binaire traditionnel est Lofti A. Zadeh, professeur d'origine iranienne à l'Université de Berkeley (Californie, USA), aux environs de 1965. En 1973, il publie dans l'IEEE Transactions on Systems, Man and Cybernetics, un de ses articles qui fait date [Zadeh, 1973]. Il y mentionne pour la première fois le terme de variable linguistique (dont la valeur est un mot et non un nombre) et ce concept ouvrait la voie à l'utilisation de la logique floue dans de nombreuses applications. Dès 1974, Mandani, professeur d'origine iranienne lui aussi, mais travaillant à l'Université de Londres, concevait la première application en logique floue [Mandani, 1975]. Il faut ensuite attendre les années 80 et les Japonais qui s'emparent de cette théorie. Le développement est alors rapide, avec la création de laboratoires spécialisés et les travaux de nombreux chercheurs. Depuis, l'engouement est devenu mondial et chaque mois, on voit paraître nombre d'articles traitant du sujet et d'applications nouvelles.

1.2. Logique floue et pensée occidentale

Si la logique floue est déjà bien implantée au Japon, il n'en est pas de même dans notre monde occidental. Le concept du flou est difficile à admettre pour des esprits bercés de philosophie grecque ou cartésienne que nous sommes. Le concept du flou viole en effet un des fondements du raisonnement occidental classique qu'est *le principe du tiers exclu d'Aristote*. Aristote, au 4ème siècle avant J.-C., formula le premier la règle des syllogismes (Tous les hommes sont mortels. Socrate est un homme, donc Socrate est mortel). Cette théorie, et plus tard celle des ensembles, est responsable d'une notion fondamentale dans l'élaboration de nos raisonnements mathématiques ; un objet est ou n'est pas, un élément appartient ou n'appartient pas à un ensemble. L'union d'un ensemble A et de son complément A^C est ainsi l'univers de référence X . L'intersection d'un ensemble et de son complémentaire est nécessairement vide.

$$A \cup A^C = X$$

$$A \cap A^C = \emptyset$$

Dans notre exemple, X étant l'ensemble des individus, A sera l'ensemble des gens petits et A^C celui des gens qui ne le sont pas, c'est-à-dire des grands. Si un élément de X n'appartient pas à A , il ne peut qu'appartenir à A^C . Et si un élément de X appartient à A , il n'appartient pas à A^C .

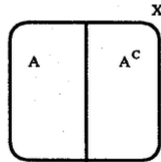


Figure 1.

Il existe toutefois des *paradoxes* qui mettent en défaut cette “théorie du blanc et du noir”. Celui d’Epiménide (6ème siècle avant J.-C.) fut le premier. Ce penseur crétois affirmait : “tous les Crétois sont des menteurs”. Qu’en déduire ? S’il dit vrai, c’est qu’il ment. Et s’il ment, il n’est pas nécessairement un menteur. . . De nombreux autres exemples suivirent, dont les plus fameux sont dus au mathématicien B. Russel au début de ce siècle. Nous nageons avec ces paradoxes en pleine contradiction et la logique classique ne peut pas nous aider. Nous verrons plus loin que la logique floue peut intégrer ces curieux problèmes.

En logique floue, une proposition peut à la fois être vraie et fausse, tout comme un élément peut appartenir à un ensemble et à son complémentaire. Une personne de 1m69 pourra donc à la fois être grande et petite. Cette violation de la loi du tiers exclu par la logique floue rend son acceptation difficile en Occident. De plus, le terme flou (fuzzy) lui confère une connotation négative, comme si elle n’était pas bien fondée, ce qui est totalement faux. Certains auteurs ont même proposé de la rebaptiser logique continue pour mieux illustrer son absence de “discontinuité” entre les ensembles A et A^C . (1)

La pensée orientale par contre intègre plus facilement ces nuances. Que le lecteur pense par exemple aux principes du Yin et du Yang. Il n’est par conséquent pas étonnant que les recherches et les applications dans le domaine de la logique floue fleurissent autant dans l’Asie tout entière : au Japon bien sûr, mais également en Chine, en Corée, à Taiwan . . .

1. Nous garderons toutefois l’appellation “logique floue” pour des raisons de cohérence avec les divers travaux déjà publiés.

1.3. Le concept du flou

Soit X l'univers de référence, aussi appelé univers de discours, de la variable x .

Un sous-ensemble flou ⁽²⁾ A est défini par une fonction d'appartenance $\mu_A(x)$, définie sur X et de co-domaine $[0, 1]$. La valeur de la fonction d'appartenance décrit le degré avec lequel l'élément x appartient à A . Ainsi, dans notre exemple, PETIT et GRAND sont des sous-ensembles flous définis sur l'ensemble des tailles de tous les individus. Deux fonctions d'appartenance μ_{petit} et μ_{grand} les définissent. Elles peuvent par exemple être du type de celles de la figure 2.

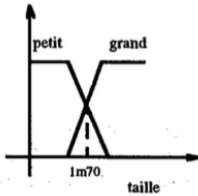


Figure 2.

Si je mesure 1,70 m, j'appartiens à la fois à l'ensemble "petit" et à l'ensemble "grand" avec des degrés d'appartenance respectifs de 0.5.

Si je mesure 1,80 m, je n'appartiens qu'à l'ensemble "grand", avec un degré 1.

L'appellation d'un sous-ensemble flou est d'ordinaire en relation avec la signification que l'on souhaite lui donner. Cette appellation est souvent désignée par le terme label ou étiquette du sous-ensemble. Ce mot correspond à la valeur de la variable linguistique x appartenant à X (en abscisse).

PETIT et GRAND sont ainsi deux valeurs linguistiques associées à la variable linguistique taille.

1.4. Et Dieu créa les ensembles ...

(Ou du moins les mathématiciens qui vont avec.)

La lecture de ce qui précède pourrait laisser croire que la théorie des ensembles est une "erreur historique". Le bon sens aurait-il été que la logique floue et les ensembles flous constituent la base de notre mode de raisonnement, la logique aristotélicienne et les ensembles classiques n'étant que des

2. On parle souvent d'ensemble flou et non pas de sous-ensemble flou. Le terme correct est sous-ensemble flou car ce dernier sera toujours affecté d'un ensemble de référence. C'est par abus de langage qu'on utilise ensemble au lieu de sous-ensemble.

cas particuliers ? Quelques adeptes fervents de la logique floue le pensent, l'écrivent parfois. Les applications qui découlent de la théorie classique attestent cependant du contraire.

Les constructions essentielles de la théorie des ensembles sont dues à Georg Cantor, mathématicien allemand (1845–1918) et à Bertrand Russell, savant anglais (1872–1970). Selon les termes mêmes de Cantor, “un ensemble résulte de la réunion, dans une même entité, de certains objets bien déterminés”. Savoir comment est rassemblée cette collection d'objets est difficile à préciser et est souvent passé sous silence dans l'approche mathématique de la question. Les mathématiciens sont pourtant tout à fait conscients de l'abstraction naissant de cette classification rigide, sous-jacente à la théorie des ensembles. Référons-nous à quelques citations pour illustrer ce propos.

“Le mathématicien n'est pas obligé, en principe, d'aller au delà du cadre de la théorie formelle des ensembles ; ceci s'applique naturellement seulement aux recherches mathématiques comme telles et non à leurs applications physiques ou autres processus extramathématiques. *La question du modèle mathématique adapté aux processus* de cette sorte n'est pas, à strictement parler, un problème mathématique”, in [Petite Encyclopédie des Mathématiques ; p. 786].

“Et en effet, en quoi consiste une déduction, par exemple une déduction syllogistique ? Elle ne consiste pas simplement à constater et à affirmer une relation d'application entre les prémisses et la conclusion : elle consiste en outre à constater que les prémisses sont vraies (ou supposées telles), et à *affirmer la conclusion absolument*”, in [Couturat, 1903 ; p. 11].

“Les mathématiques font croire que tout l'univers est divisible en deux parties distinctes, qu'à chaque question il existe deux sortes de réponses : les bonnes et les mauvaises, les vraies et les fausses, ou plus simplement : *oui ou non.*”, in [Pourquoi la Mathématique ? ; p. 133].

Et Bertrand Russell lui-même soulève l'ambiguïté de cette classification inflexible [Russel, 1923] :

“All traditional logic habitually assumes that precise symbols are being employed. It is therefore not applicable to this terrestrial life, but only to an imagined celestial one. The law of excluded middle (A or not- A) is true when precise symbols are

employed but is not true when symbols are vague as, in fact, all symbols are”.

Remplacez vague par flou dans sa phrase, et les fondements de la logique floue sont jetés, ou presque. C’est donc dire combien, déjà aux balbutiements de la théorie des ensembles, on se rendait compte des problèmes potentiels qu’elle soulevait.

La logique floue n’est pas plus épargnée. Certains l’accusent aussi de rigidité :

“Les logiques à n dimensions ou éléments ou composantes (je ne sais plus trop). Là encore, après discussion, les tenants de cet argument ont dû reconnaître qu’en définitive, il y a aussi dichotomie : entre ce qui est conforme à la logique et ce qui ne l’est pas”, in [Pourquoi la Mathématique ? ; p. 133].

Je pense, quant à moi, que la conformité à un mode de pensée est une nécessité dans l’élaboration d’un schéma de raisonnement parfaitement reconnaissable et maîtrisable par tout un chacun. La logique classique et la logique floue poursuivent des buts voisins mais différents. Il est de nombreux problèmes pour lesquels la logique traditionnelle s’avère tout à fait satisfaisante et utile. La classification d’objets définis précisément et selon un critère déterminé est de ceux-là : par exemple, l’ensemble S de toutes les personnes habitant un même immeuble B à un certain instant t [Petite Encyclopédie des Mathématiques, 1980], ou celui des boules noires tirées d’une urne. D’autres cas par contre sont plus délicats à traiter : l’ensemble des fleurs rouges et celui des fleurs roses par exemple [Scholten, 1994] ou ceux des petits et des grands. L’intérêt de la logique floue se marque surtout dans sa proximité avec la réalité physique et humaine (d’où l’engouement suscité par cette technique dans les milieux techniques ou des sciences humaines) mais la logique traditionnelle reste et restera une absolue nécessité.

1.5. Logique floue et probabilités

On l’a compris, la logique floue introduit une notion d’imprécision dans les raisonnements. Une question vient alors à l’esprit : la théorie des probabilités ne suffit-elle pas à décrire cet incertain ?

La réponse est non mais la confusion est facile. Les deux théories décrivent une notion de doute et ce, à l’aide de nombres compris entre 0 et 1. Toutes deux jouent avec des ensembles et traitent des propositions. Il

existe pourtant une différence fondamentale entre elles deux : elle concerne les ensembles et leurs complémentaires.

La probabilité qu'un élément appartienne à A et A^C est nulle : $P(A \cap A^C) = P(\emptyset) = 0$.

En logique floue, par contre, un élément pourra appartenir et à A et à A^C .

Pour mieux saisir cette nuance, reprenons un exemple de B. Kosko [Kosko, 1989]. Supposons avoir une pomme. Elle peut se trouver dans le frigo ou hors du frigo. Il y a donc un pourcentage de chances pour qu'elle soit à l'intérieur : c'est la probabilité avec laquelle la pomme appartient à l'ensemble "est dans le frigo". Coupons la pomme en deux. Nous pouvons mettre une demi-pomme dans le frigo et poser l'autre moitié sur la table. La pomme, dans son unité, est à la fois dans et hors du frigo. Le degré avec lequel la pomme appartient à l'ensemble $\{x \mid x \text{ est dans le frigo}\}$ ou $\{x \mid x \text{ n'est pas dans le frigo}\}$ est le degré d'appartenance de la variable pomme à ces deux sous-ensembles flous.

La logique floue décrit l'ambiguïté d'un événement. Les probabilités décrivent le doute quant à l'occurrence d'un événement.

Une probabilité est toujours relative à un ensemble classique, jamais à un sous-ensemble flou. Il existe une théorie particulière qui regroupe à la fois la logique floue et les probabilités ; c'est *la théorie des possibilités* [Zadeh, 1978 et Dubois-Prade, 1988]. Elle permet de raisonner sur des connaissances imprécises ou vagues tout en introduisant un moyen de prendre en compte l'incertitude sur ces connaissances.

2. Bases théoriques de la logique floue

2.1. Le sous-ensemble flou

2.1.1. Définition

Un sous-ensemble flou est défini par sa fonction d'appartenance :

$$\mu_A(x) : X \rightarrow [0, 1]$$

Un ensemble classique (on le qualifie aussi parfois d'ordinaire) est un ensemble flou dont la fonction d'appartenance ne peut prendre que deux

valeurs : 0 ou 1 :

$$\begin{aligned}\mu_A &= 1 \quad \text{si } x \in A \\ &= 0 \quad \text{sinon.}\end{aligned}$$

C'est le cas de la première fonction de la figure 3.

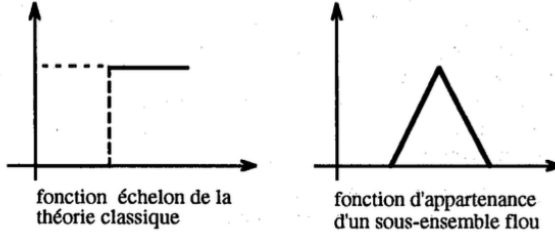


Figure 3.

2.1.2. Particularités

On définit la *hauteur* (*height*) d'un sous-ensemble flou comme étant la valeur maximale que prend sa fonction d'appartenance :

$$\text{hgt}(A) = \sup_{x \in X} \mu_A(x).$$

On dit qu'un sous-ensemble est *normal* lorsque sa hauteur vaut 1.

Le *support* de A est l'ensemble des éléments de X qui appartiennent au moins un peu à A :

$$\text{supp}(A) = \{x \in X \mid \mu_A(x) \neq 0\}.$$

Le *noyau* de A est l'ensemble des éléments appartenant de façon absolue à A :

$$\text{noy}(x) = \{x \in X \mid \mu_A(x) = 1\}.$$

Un sous-ensemble classique est un sous-ensemble flou normal, identique à son support et à son noyau.

La cardinalité d'un sous-ensemble flou est définie par

$$\begin{aligned} M(A) &= \sum_{x \in X} \mu_A(x) \quad \text{si } A \text{ est discret.} \\ &= \int_x \mu_A(x) dx \quad \text{si } A \text{ est continu.} \end{aligned}$$

On retrouve la définition usuelle pour les ensembles classiques : si l'ensemble est discret, la cardinalité est le nombre d'éléments que contient l'ensemble.

2.2. Opérations en logique floue

Il s'agit d'étendre aux sous-ensembles flous les opérations de la théorie des ensembles classiques : l'union, l'intersection et la négation. On utilise pour ce faire les fonctions d'appartenance des sous-ensembles flous, qu'on combinera de manière adéquate. On doit retrouver dans le cas classique les propriétés habituelles. Les solutions proposées doivent en outre répondre à certaines propriétés (croissance, associativité, ...) spécifiques aux fonctions considérées.

2.2.1. Complément d'un ensemble (fonction négation)

a) *Critères*

Une fonction négation *non* doit vérifier les conditions suivantes :

- $non(\mu(x)) : [0, 1] \rightarrow [0, 1]$
- $non(\mu(x)) \downarrow$ si $\mu(x) \uparrow$ (croissance)
- $non(1) = 0$ et $non(0) = 1$ (conditions aux limites)

b) *Choix*

Il existe plusieurs fonctions répondant aux critères énoncés. On choisit généralement d'adopter la fonction complément à 1, très simple d'utilisation :

$$\text{non}(\mu_A(x)) = 1 - \mu_A(x).$$

c) *Illustration*

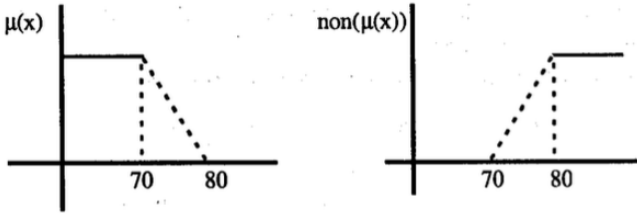


Figure 4.

2.2.2. Intersection de deux ensembles (fonction *et*)

a) *Critères*

- $et(\mu_A(x), \mu_B(x)) : [0, 1] \times [0, 1] \rightarrow [0, 1]$
- $et(\mu_A(x), \mu_B(x)) = et(\mu_B(x), \mu_A(x))$ (commutativité)
- si $\mu_A(x) < \mu_B(x)$ et $\mu_C(x) < \mu_D(x) : et(\mu_A(x), \mu_C(x)) < et(\mu_B(x), \mu_D(x))$ (croissance)
- $et(\mu_A(x), et(\mu_B(x), \mu_C(x))) = et(et(\mu_A(x), \mu_B(x)), \mu_C(x))$ (associativité)
- $et(0, \mu_A(x)) = 0$, $et(1, \mu_A(x)) = \mu_A(x)$ (conditions aux limites).

Une fonction répondant à de tels critères est dite une *t-norme*.

b) *Choix*

De nombreuses possibilités existent (voir tableau récapitulatif). La première proposition, due à Zadeh, est encore aujourd'hui la plus utilisée : la fonction minimum $et(\mu_A(x), \mu_B(x)) = \min(\mu_A(x), \mu_B(x))$. Son interprétation est simple : un élément ne peut pas appartenir plus à l'intersection de deux ensembles qu'à chacun de ceux-ci. Cette fonction est illustrée ci-après.

c) *Illustration*

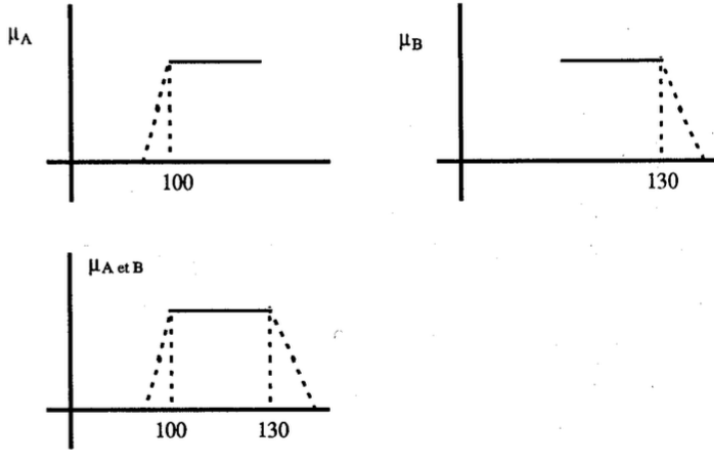


Figure 5.

2.2.3. Union de deux ensembles (fonction *ou*)

a) Critères

La fonction doit être telle que :

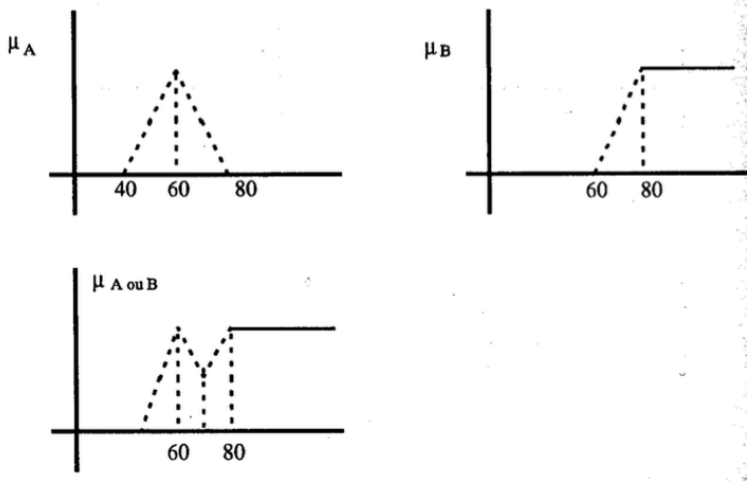
- $ou(\mu_A(x), \mu_B(x)) : [0, 1] \times [0, 1] \rightarrow [0, 1]$
- $ou(\mu_A(x), \mu_B(x)) = ou(\mu_B(x), \mu_A(x))$ (commutativité)
- si $\mu_A(x) < \mu_B(x)$ et $\mu_C(x) < \mu_D(x) : ou(\mu_A(x), \mu_C(x)) < ou(\mu_B(x), \mu_D(x))$ (croissance)
- $ou(\mu_A(x), ou(\mu_B(x), \mu_C(x))) = ou(ou(\mu_A(x), \mu_B(x)), \mu_C(x))$ (associativité)
- $ou(0, \mu_A(x)) = \mu_A(x)$, $ou(1, \mu_A(x)) = 1$ (conditions aux limites).

Une telle fonction est une *s-norme*.

b) *Choix*

Plusieurs choix sont également possibles, dont les plus courants sont repris dans le tableau illustratif ci-après. C'est généralement la fonction maximum qu'on sélectionne ou $(\mu_A(x), \mu_B(x)) = \max(\mu_A(x), \mu_B(x))$. Elle est illustrée ci-dessous.

c) *Illustration*



2.2.4. Tableau récapitulatif des fonctions floues les plus usitées

Appellation	INTERSECTION	UNION	COMPLEMENT
Zadeh	$\min(\mu_A(x), \mu_B(x))$	$\max(\mu_A(x), \mu_B(x))$	$1 - \mu(x)$
Probabilistique	$\mu_A(x)\mu_B(x)$	$\mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$	$1 - \mu(x)$
Lukasiewicz	$\max(\mu_A(x) + \mu_B(x) - 1, 0)$	$\min(\mu_A(x) + \mu_B(x), 1)$	$1 - \mu(x)$
Hanacer	$\mu_A(x)\mu_B(x)$	$\mu_A(x) + \mu_B(x) + \mu_A(x)\mu_B(x) - (1, \beta)\mu_A(x)\mu_B(x)$	$1 - \mu(x)$
$(\beta > 0)$	$\beta + (1 - \beta)(\mu_A(x) + \mu_B(x) + \mu_A(x)\mu_B(x))$	$1 - (1 - \beta)\mu_A(x)\mu_B(x)$	
Weber	$\mu_A(x)$ si $\mu_B(x) = 1$ $\mu_B(x)$ si $\mu_A(x) = 1$ 0 sinon	$\mu_A(x)$ si $\mu_B(x) = 0$, $\mu_B(x)$ si $\mu_A(x) = 0$ 1 sinon	$1 - \mu(x)$

2.3. Des sous-ensembles flous à la logique floue

Tout comme la théorie des ensembles classiques est intimement liée à la logique des propositions, les sous-ensembles flous donnent naissance à une logique qualifiée de floue (bien que, rappelons-le, elle soit tout à fait déterministe!).

Un sous-ensemble flou A et une proposition floue p_F peuvent être mis en rapport lorsque tous deux sont définis en terme d'un même ensemble d'objets et que chacun représente le même sens de l'expression linguistique.

Soit une proposition floue $p_F : 'x \text{ est } F'$ où F décrit une valeur linguistique de l'élément x , telle que petit, grand, rouge, etc.

Soit A un sous-ensemble flou défini par sa fonction d'appartenance $\mu_A(x)$ définie sur X .

On peut définir une proposition floue $p_A : 'x \text{ est élément de } A'$ dont la valeur de vérité sera identique à $\mu_A(x)$, $\forall x \in X$. De même, on peut définir un sous-ensemble flou F , de fonction d'appartenance $\mu_F(x)$, relatif à p_F . L'équivalence est donc prouvée. Dès lors, les opérations logiques définies pour les sous-ensembles flous peuvent s'appliquer aux propositions floues et on parlera indifféremment de complément ou de fonction négation, d'intersection ou de fonction ET, d'union ou de fonction OU.

2.4. L'implication floue et la relation floue

L'étape suivante dans la généralisation est l'implication. La table de vérité de l'implication booléenne peut être obtenue en la décomposant préalablement en une combinaison de fonctions fondamentales.

p	q	$p \Rightarrow q$	
0	0	1	
0	1	1	$p \Rightarrow q \Leftrightarrow \text{non}(\text{et}(p, \text{non}(q))) \Leftrightarrow \text{ou}(\text{non}(p), q)$
1	0	0	
1	1	1	

L'implication est toujours vraie sauf si le fait p est vérifié alors que la conclusion q ne l'est pas.

Une implication floue est de la forme $x \in A \Rightarrow y \in B$. La généralisation de l'implication concernera donc les fonctions $\mu_A(x)$ et $\mu_B(x)$, relatives aux sous-ensembles flous A et B entrant en jeu. Le degré de véracité $\mu_{B/A}(x, y)$

de cette proposition sera une fonction $I(\mu_A(x), \mu_B(y))$. De nouveau, plusieurs possibilités existent.

Citons (la liste est loin d'être exhaustive) :

Appellation	$\mu_{B/A}(x, y)$
Lukasiewicz	$\min(1 - \mu_A(x) + \mu_B(y), 1)$
Larsen	$\mu_A(x) \cdot \mu_B(y)$
Zadeh ou Mandani	$\min(\mu_A(x), \mu_B(y))$
Reichenbach	$1 - \mu_A(x) + \mu_A(x)\mu_B(y)$
Brouwen-Gödel	1 si $\mu_A(x) \leq \mu_B(y)$; $\mu_B(y)$ sinon

Les trois premières possibilités sont les plus courantes. Celles de Larsen et Mandani sont largement utilisées pour le contrôle flou des processus. Elles ne généralisent pas strictement l'implication booléenne (cas $\mu_A(x) = \mu_B(y) = 0$) mais sont très simples à implémenter.

2.5. Le raisonnement flou

Le problème tel qu'il se pose en pratique n'est généralement pas de mesurer le degré de véracité d'une implication mais bien de déduire, à l'aide de faits et de diverses règles implicatives, des événements potentiels. En logique classique, un tel raisonnement porte le nom de MODUS PONENS.

Si $p \Rightarrow q$ vrai
 et p vrai
 alors q vrai.

Il faut donc le généraliser. Vont intervenir dans cette étape la fonction d'implication de l'implication floue, image du lien, et les fonctions d'appartenance des sous-ensembles des prémisses et conclusions, images de l'importance des événements. Le MODUS PONENS GENERALISE sera :

Si x est A alors y est B .
 x est A'
 alors y est B' .

A' et B' sont les sous-ensembles flous constatés dans le cas que l'on traite et ne sont pas nécessairement strictement égaux à A et B . B' est le sous-ensemble flou résultant de A' par l'application de l'implication. Pour un x donné, le degré d'appartenance de l'élément résultant de B' est intimement

lié à celui de x à A' et à la fonction d'implication $\mu_{B/A}(x, y)$ puisqu'il en découle. C'est donc à la fonction "et" floue qu'on a affaire, traditionnellement représentée par un minimum (Zadeh). L'ensemble B' est obtenu dans ce cas par

$$\mu_{B'}(y) = \max_x \min\{\mu_{A'}(x), \mu_{B/A}(x, y)\}.$$

L'opération que l'on vient de réaliser est en fait la composition de deux relations (appartenance à A' et fonction d'implication) et porte le nom de composition max/min. Voir à ce sujet [Bouchon-Meunier, 1995] ou [Klir G., 1996] pour de plus amples renseignements.

Le choix de la fonction d'implication déterminera donc la fonction $\mu_{B'}(y)$. Les deux fonctions les plus usitées sont celles de Mandani et de Larsen qui donnent respectivement :

- $\mu_{B'}(y) = \max_x \min\{\mu_{A'}(x), \min(\mu_A(x), \mu_B(y))\}$
 $= \max_x \min(\mu_A(x), \mu_B(y))$
- $\mu_{B'}(y) = \max_x \min\{\mu_{A'}(x), \mu_A(x)\mu_B(y)\}$

2.6. Paradoxes et logique floue

Revenons un instant aux problèmes que soulevaient les paradoxes en logique classique. Nous avons dit que la logique floue les intégrait aisément. C'est en effet le cas. La proposition p "tous les Crétois sont des menteurs" est tout bonnement "maximally fuzzy". Ses degrés de véracité $t(p)$ et d'erreur $t(\text{non}p)$ sont identiques, car a priori on ne peut pas savoir si Epiménide dit la vérité ou non.

$$t(p) = \text{non}(t(p)) = 1 - t(p) \Rightarrow t(p) = 0.5.$$

Cette constatation avait déjà conduit Lukasiewicz à introduire une logique multivariée (0, 0.5 et 1) dès 1920. Cette relation entre $t(p)$ et $\text{non}(t(p))$ est applicable à de nombreux autres paradoxes. Elle illustre combien l'introduction de la logique floue peut finalement être considérée comme une évolution naturelle de la théorie classique.

3. Applications de la logique floue

Ces dernières années, l'engouement pour la logique floue a pris une ampleur exceptionnelle. On la rencontre aujourd'hui à chaque coin de rue, ou presque ... Elle est présente dans des domaines variés (sciences exactes, sociales, économiques, appliquées, etc.) et a très souvent prouvé son efficacité dans la résolution de problèmes jusque là qualifiés d'épineux.

3.1. Application à la commande de processus

Une des principales applications de la logique floue est la conduite ou la commande de processus. Pouvoir commander un système est essentiel, que ce soit par exemple en aéronautique (commande des ailes d'un avion, ...), en robotique (commande d'un bras de robot, ...) ou en chimie (contrôle de la production d'une substance chimique, ...). Les raisons du succès de la logique floue dans ce domaine sont principalement une grande souplesse et une relative facilité de conception. Un contrôleur flou intègre en fait la connaissance sous une forme plus naturelle et plus proche du raisonnement humain qu'un algorithme numérique. La connaissance "intuitive" d'un expert s'avère d'ailleurs très précieuse lors de sa conception.

La commande floue a ainsi été appliquée avec succès à de nombreux systèmes :

- ABS [Nissan - Mercedes Benz].
- Contrôle de la production d'éthylène [Hoechat, Allemagne - Rubber, Canada].
- Réglage de la ventilation d'un tunnel [Toshiba].
- Auto-focus d'appareils photographiques [Canon - Minolta].
- Conduite automatique de métros urbains [Sendai - Tokyo].
- etc.

3.2. Application à l'aide à la décision

Un autre domaine où fleurissent les applications en logique floue est celui de l'aide à la décision. Ici aussi, les exemples sont nombreux.

- Machines à laver intelligentes [AEG, Philips].
- Reconnaissance vocale en logique floue [Ricon].
- Reconnaissance de l'écriture manuscrite [Intel, Sony].
- Contrôle du trafic automobile [Autoroute Osaka-Sakai, Japon].

-
-
- Réduction du déclenchement intempestif d’alarme incendie [Cerberus, Suisse].
 - Détermination de cartes de crédit frauduleuses [XTEC, Floride].
 - etc.

Citons aussi, car c’est important, une réalisation belge dans le domaine : la tondeuse TORTUE, fonctionnant à l’énergie solaire et gérant la tonte automatique de la pelouse selon sa hauteur et le degré d’ensoleillement.

3.3. Divers

Citons enfin quelques autres domaines où la logique floue a également été appliquée :

- Météorologie
prévision du temps par logique floue.
- Médecine
diagnostic informatique des maladies par logique floue.
- Finance
prédiction des marchés financiers.
- etc.

4. En guise de conclusion

La logique floue est une discipline relativement neuve mais prometteuse à bien des égards. Son intérêt principal réside dans la modulation de la connaissance, avec l’extension de la logique binaire traditionnelle à l’intervalle $[0,1]$ tout entier. Cette généralisation des valeurs de vérité entraîne tout naturellement celle des opérations logiques traditionnelles. Il est à noter que la formalisation théorique de la logique floue est encore aujourd’hui en plein développement. On pourrait presque parler de “logiques floues” plus que de logique floue, chacune d’elle correspondant à un type précis de recherches et d’applications.

Sur un plan plus pratique, les résultats obtenus parlent le plus souvent d’eux-mêmes. Ainsi, les nombreuses réalisations pratiques déjà commercialisées témoignent des avantages qu’elle peut apporter. Lofti Zadeh a coutume de conclure ses exposés par la maxime “Clearly, the future is fuzzy”. C’est peut-être en effet là la meilleure manière de le faire.

Je terminerai en renvoyant le lecteur intéressé aux références de B. Bouchon-Meunier, G. Klir et L. Zadeh pour de plus amples informations.

Bibliographie

- [1] Bouchon-Meunier B., *La logique floue*, Collection Que Sais-Je ?, Presses Universitaires de France, 1994.
- [2] Bouchon-Meunier B., *Logique Floue et Applications*, Hermes, 1995.
- [3] Couturat L., *Les Principes des Mathématiques*, Librairie Scientifique et Technique, 1903.
- [4] Cox E., Fuzzy fundamentals, *IEEE Spectrum*, October 1992, 58–61.
- [5] Dubois D. et Prade H., *Théorie des Possibilités*, Masson, seconde édition, 1988.
- [6] Kosko B., *Fuzzy Systems and Neural Networks*, Prentice-Hall, 1989.
- [7] Klir G., *Fuzzy Sets*, Diderot Publishers, 1996.
- [8] Handani E., An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller, *International Journal on Man-Machine Studies*, vol. 7, 1–13, 1975.
- [9] *Petite Encyclopédie des Mathématiques*, Didier, 1980.
- [10] *Pourquoi la Mathématique ?* 1974. Union Générale d'Éditions (Paris), 10/18.
- [11] Russel B., *Vagueness*, Australian Journal of Philosophy, vol. I, 1923.
- [12] Scholten H., *Logique floue et régulation PID*, Publitronic, 1994.
- [13] Yukatovitch P., *Fuzzy Control*, Wiley, 1998.
- [14] Zadeh L.A., Outline of a New Approach to the Analysis Complex Systems and Decision Processes, *IEEE Trans. Syst. Man Cybern.*, vol. 3, 28–44, 1973.
- [15] Zadeh L.A., Fuzzy Sets as a basis for a theory of possibility, *Fuzzy Sets and Systems*, vol. 1, 3–28, 1978.

Adresse de l'auteur :

Anne BAJART

Ingénieur de recherches F.R.I.A. (Fonds pour la
Recherche dans l'Industrie et l'Agriculture)
Electronique de Commutation – Université de Liège
Institut Montéfiore - Bât. B28 – 4000 Liège
04/366 2618 - bajart@montefiore.ulg.ac.be

Modélisation du comportement d'un individu face à des revenus aléatoires

G. Haesbroeck, Université de Liège

Mots-clé : fonction d'utilité, risque, espérance mathématique, convexité

1. Critère de l'espérance de l'utilité

Lorsqu'un individu est amené à comparer des revenus aléatoires, tout se passe comme s'il était en présence d'une *loterie* (ou variable aléatoire) $\tilde{x} = [x; \Pi]$ où x est un vecteur (x_1, \dots, x_n) de gains monétaires x_i pour chaque possibilité, souvent baptisée *état de la nature*, $i = 1, \dots, n$ et Π désigne un vecteur (Π_1, \dots, Π_n) de probabilités Π_i associées à chaque gain monétaire x_i , avec, bien entendu,

$$\sum_{i=1}^n \Pi_i = 1.$$

De quelle façon un agent rationnel évalue-t-il une telle loterie ?

Pour répondre à cette question, considérons en guise d'exemple un billet de loterie qui fait gagner 100 F dans le tiers des cas et 70 F dans les deux tiers des cas restants : beaucoup de personnes sont prêtes à payer un tel billet au prix égal à l'espérance mathématique du gain, soit à

$$\frac{1}{3} \times 100 + \frac{2}{3} \times 70 = 80 \text{ F.}$$

Plus généralement, le billet de loterie $\tilde{x} = [x; \Pi]$ sera évalué par l'*espérance de gain* (ou *valeur attendue*)

$$E(\tilde{x}) = \sum_{i=1}^n \Pi_i x_i.$$

Cette évaluation intuitive se heurte au *paradoxe de Saint-Pétersbourg* qui peut être introduit simplement de la manière suivante. Considérons un jeu J de "pile ou face" dans lequel le joueur gagnera 2^n francs si face sort pour

la première fois au n -ième lancer. Pour une pièce parfaitement équilibrée, la probabilité de cet événement est égale à $\frac{1}{2^n}$ (puisque les lancers successifs sont indépendants et que la pièce a autant de chance de tomber sur pile que sur face). Ce jeu J possède donc une espérance de gain égale à

$$E(J) = \sum_{n=1}^{+\infty} \Pi_n x_n = \sum_{n=1}^{+\infty} \left(\frac{1}{2^n} \right) 2^n = \sum_{n=1}^{+\infty} 1 = +\infty!$$

Cette situation est paradoxale car, bien que l'espérance du gain soit "infinie", il n'existe évidemment aucun individu prêt à parier une somme très importante (quoique finie) pour participer à un tel jeu!

Afin de lever ce paradoxe, D. Bernoulli (en 1738) a émis l'hypothèse que les individus ont une utilité marginale décroissante pour la monnaie. Par exemple, si on propose à une personne de jouer 10 francs à pile ou face, elle acceptera vraisemblablement; par contre, si la mise est de 100.000 francs, la même personne déclinera probablement l'offre. Alors qu'un gain de 10 francs semble équilibré par une perte de 10 francs, pourquoi un gain de 100.000 francs ne contrebalance-t-il pas une perte de 100.000 francs? La réponse est simple: dans la deuxième situation, la plupart des gens sont effrayés par le risque à courir; en d'autres termes, ils pensent que la perte de 100.000 francs est plus pénible que le bénéfice obtenu par le gain de la même valeur.

La théorie de la décision examine ce problème en introduisant une fonction qui "mesure" l'attraction de l'argent: il s'agit de la *fonction d'utilité* u (ou encore fonction de satisfaction) qui, à une somme x , associe l'utilité $u(x)$. Cette fonction d'utilité doit être strictement croissante, mais à dérivée décroissante; de plus, elle est souvent définie à une fonction affine croissante près et porte alors le nom de fonction d'utilité de Von Neumann et Morgenstern.

Selon Bernoulli, le comportement rationnel d'un agent face au risque consiste à mesurer l'*espérance de l'utilité* (ou *utilité attendue*), à savoir, pour la loterie $\tilde{x} = [x; \Pi]$, l'expression notée $E[u(\tilde{x})]$ et définie par

$$E[u(\tilde{x})] = \sum_{i=1}^n \Pi_i u(x_i).$$

Par exemple, pour le jeu J ci-dessus, prenons $u(x) = \ln x$. On trouve alors

$$\begin{aligned}
 E[u(J)] &= \sum_{n=1}^{+\infty} \Pi_n u(x_n) = \sum_{n=1}^{+\infty} \left(\frac{1}{2^n}\right) \ln(2^n) \\
 &= (\ln 2) \sum_{n=1}^{+\infty} \frac{n}{2^n} = (\ln 2) \sum_{n=1}^{+\infty} \sum_{k=1}^n \frac{1}{2^n} \\
 &= (\ln 2) \sum_{k=1}^{+\infty} \sum_{n=k}^{+\infty} \frac{1}{2^n} \\
 &= (\ln 2) \sum_{k=1}^{+\infty} \frac{2}{2^k} = 2 \ln 2 = \ln 4 = u(4).
 \end{aligned}$$

Ainsi, l'individu est prêt à payer 4 francs pour jouer puisque l'espérance de l'utilité du jeu est égale à l'utilité de l'obtention d'une somme certaine de 4 francs.

2. Comportement vis-à-vis du risque

Soient u une fonction d'utilité de Von Neumann et Morgenstern définie sur \mathbb{R}_+ et $\tilde{x} = [x_1, \dots, x_n; \Pi_1, \dots, \Pi_n]$ une variable aléatoire (ou loterie).

Plusieurs comportements sont possibles face au risque.

a) *Neutralité pour le risque.* Un individu est *neutre* pour le risque lorsque l'utilité de la valeur attendue est égale à l'espérance de l'utilité, c'est-à-dire lorsque

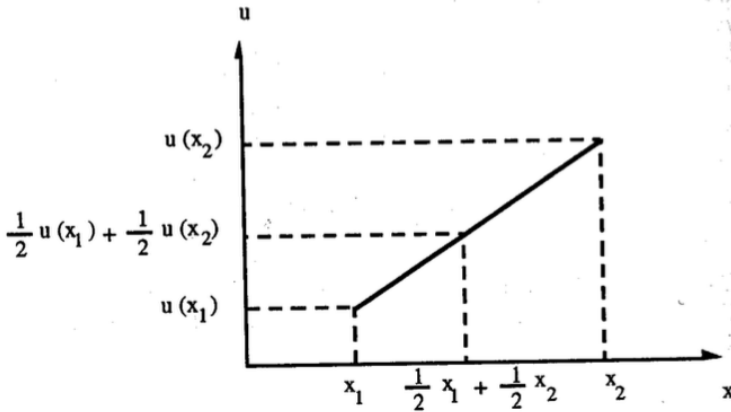
$$u[E(\tilde{x})] = u\left(\sum_{i=1}^n \Pi_i x_i\right) = E[u(\tilde{x})] = \sum_{i=1}^n \Pi_i u(x_i).$$

Par exemple, considérons un individu qui a le choix entre, d'une part, un billet de loterie qui donne un gain de x_1 avec une probabilité $\frac{1}{2}$ et un gain de x_2 avec une probabilité $\frac{1}{2}$ et, d'autre part, un revenu certain égal à la moyenne arithmétique $(\frac{1}{2} x_1 + \frac{1}{2} x_2)$ des gains ; s'il est neutre face au

risque, alors il est indifférent entre ces deux possibilités, soit

$$u\left(\frac{1}{2}x_1 + \frac{1}{2}x_2\right) = \frac{1}{2}u(x_1) + \frac{1}{2}u(x_2).$$

Il y a neutralité pour le risque lorsque la fonction d'utilité est affine et croissante [1], c'est-à-dire quand son graphique est une droite qui "monte" (en la parcourant de la gauche vers la droite).



b) *Goût pour le risque.* Un individu a *du goût pour le risque* lorsque l'utilité de la valeur attendue est inférieure à l'espérance de l'utilité, c'est-à-dire lorsque

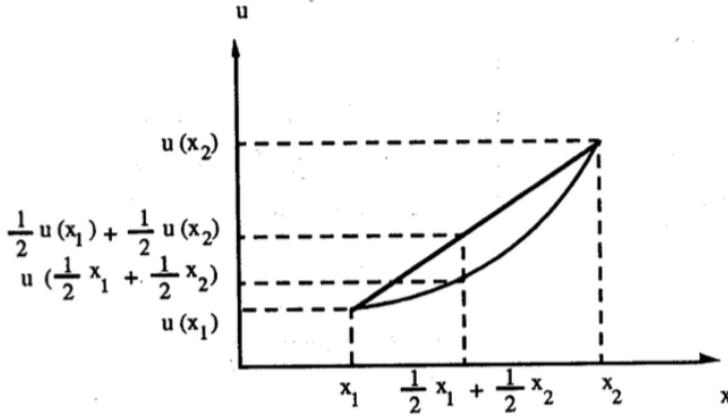
$$u[E(\tilde{x})] = u\left(\sum_{i=1}^n \Pi_i x_i\right) < E[u(\tilde{x})] = \sum_{i=1}^n \Pi_i u(x_i).$$

C'est le cas lorsque la fonction d'utilité est strictement convexe, notamment lorsque

$$u''(x) = \frac{d^2u}{dx^2} > 0, \quad \forall x > 0,$$

puisqu'alors l'inégalité de Jensen est vérifiée [2].

Si l'individu, considéré dans l'exemple ci-dessus, a le goût pour le risque, on se trouve dans une situation représentée sur un graphique de ce type :



c) *Aversion pour le risque.* Un individu a de l'*aversion pour le risque* lorsque l'utilité de la valeur attendue est supérieure à l'espérance de l'utilité, c'est-à-dire lorsque

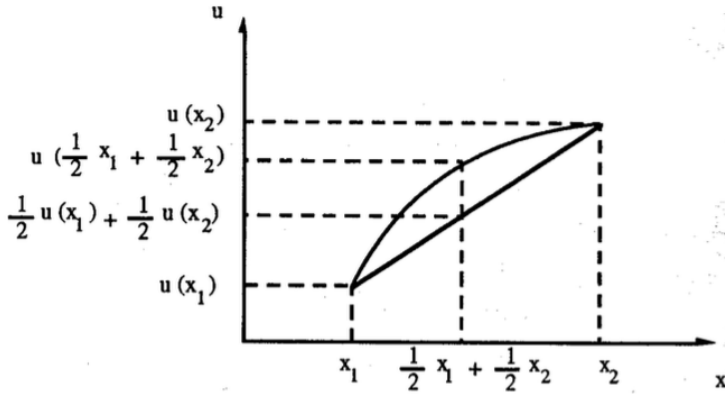
$$u[E(\tilde{x})] = u\left(\sum_{i=1}^n \Pi_i x_i\right) > E[u(\tilde{x})] = \sum_{i=1}^n \Pi_i u(x_i).$$

C'est le cas lorsque la fonction d'utilité est strictement concave [2], notamment lorsque

$$u''(x) = \frac{d^2u}{dx^2} < 0, \quad \forall x > 0,$$

puisqu'alors l'inégalité de Jensen relative aux fonctions strictement concaves est vérifiée.

Si l'individu de l'exemple ci-dessus a de l'*aversion pour le risque*, la figure correspondante se présente comme suit :



3. Prime de risque

On appelle *équivalent certain* de la variable aléatoire \tilde{x} la valeur certaine, notée $EC(\tilde{x})$, “équivalente”, pour l’agent, à \tilde{x} , c’est-à-dire telle que

$$u[EC(\tilde{x})] = E[u(\tilde{x})]. \quad (1)$$

Par définition, la *prime de risque* de \tilde{x} vaut la différence, notée $\rho(\tilde{x})$, entre la valeur attendue $E(\tilde{x})$ et l’équivalent certain $EC(\tilde{x})$. On a donc

$$\rho(\tilde{x}) = E(\tilde{x}) - EC(\tilde{x}) \Leftrightarrow EC(\tilde{x}) = E(\tilde{x}) - \rho(\tilde{x}), \quad (2)$$

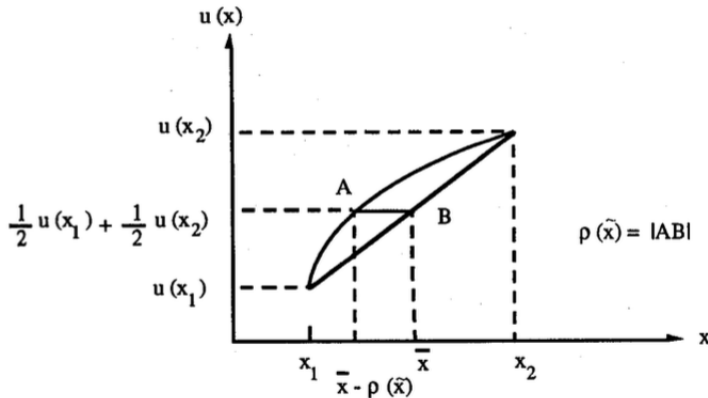
d’où l’on déduit :

$$u[E(\tilde{x}) - \rho(\tilde{x})] = E[u(\tilde{x})]. \quad (3)$$

Le caractère strictement croissant de la fonction d’utilité u permet de déterminer le signe de la prime de risque $\rho(\tilde{x})$. En effet, dans le cas d’aversion pour le risque, l’inégalité $u[E(\tilde{x})] > E[u(\tilde{x})] = u[E(\tilde{x}) - \rho(\tilde{x})]$ entraîne $\rho(\tilde{x}) > 0$; dans le cas de goût pour le risque, $u[E(\tilde{x})] < E[u(\tilde{x})] = u[E(\tilde{x}) - \rho(\tilde{x})]$ entraîne $\rho(\tilde{x}) < 0$.

En guise d’exemple, considérons un billet qui rapporte ou bien x_1 avec la probabilité $\frac{1}{2}$ et x_2 avec la probabilité $\frac{1}{2}$, ou bien un revenu certain égal à $\bar{x} = \frac{1}{2}x_1 + \frac{1}{2}x_2$. La prime de risque peut être visualisée géométriquement : de fait, si l’on trace le graphique de la fonction d’utilité entre les

abscisses x_1, x_2 , la droite joignant les points $(x_1, u(x_1)), (x_2, u(x_2))$, puis la droite horizontale passant par $B = (\bar{x}, E[u(\bar{x})])$ et coupant le graphique de u au point A , alors $\rho(\tilde{x})$ vaut la mesure algébrique du segment de droite d'extrémités A et B . Dans le cas d'aversion pour le risque, on obtient une figure de ce type :



4. Indice absolu d'aversion pour le risque

Nous allons considérer une variable aléatoire $\tilde{x} = \bar{x} + \tilde{\varepsilon}$ de moyenne \bar{x} (ce qui entraîne évidemment $E(\tilde{\varepsilon}) = 0$), et de variance $\sigma_{\tilde{\varepsilon}}^2$ peu élevée (d'où intuitivement, la variation par rapport à la moyenne est petite), avec des moments d'ordre supérieur négligeables par rapport à $\sigma_{\tilde{\varepsilon}}^2$. Supposons de plus la fonction d'utilité u deux fois différentiable, concave et strictement croissante sur $[0, +\infty[$, ce qui entraîne $u'(x) > 0$ pour tout $x > 0$ ⁽¹⁾. La prime de risque $\rho(\tilde{x})$ sera encore notée $\rho(\bar{x}, \tilde{\varepsilon})$ puisqu'elle dépend de la constante \bar{x} et de la variable aléatoire $\tilde{\varepsilon}$: on a donc

$$E[u(\tilde{x})] = E[u(\bar{x} + \tilde{\varepsilon})] = u[\bar{x} - \rho(\bar{x}, \tilde{\varepsilon})]. \tag{4}$$

Pour toute valeur ε prise par la variable aléatoire $\tilde{\varepsilon}$, on peut écrire, en exploitant le développement de Taylor d'ordre 2 et en négligeant le reste de

1. En effet, la croissance de u entraîne $u'(x) \geq 0$ pour tout $x > 0$. S'il existe $y > 0$ tel que $u'(y) = 0$, alors, en vertu de la concavité de u , $u(y)$ est le maximum global de u sur $[0, +\infty[$, ce qui est absurde puisque u est strictement croissante.

Lagrange obtenu :

$$u(\bar{x} + \varepsilon) \simeq u(\bar{x}) + \varepsilon u'(\bar{x}) + \frac{\varepsilon^2}{2} u''(\bar{x}), \quad (5)$$

d'où

$$\begin{aligned} E[u(\bar{x} + \tilde{\varepsilon})] &\simeq u(\bar{x}) + u'(\bar{x}) E(\tilde{\varepsilon}) + \frac{u''(\bar{x})}{2} E(\tilde{\varepsilon}^2) \\ &= u(\bar{x}) + u'(\bar{x}) 0 + \frac{u''(\bar{x})}{2} \text{Var}(\tilde{\varepsilon}) \\ &= u(\bar{x}) + \frac{u''(\bar{x})}{2} \sigma_\varepsilon^2. \end{aligned} \quad (6)$$

Par ailleurs, on obtient cette nouvelle approximation linéaire de u :

$$u[\bar{x} - \rho(\bar{x}, \tilde{\varepsilon})] \simeq u(\bar{x}) - \rho(\bar{x}, \tilde{\varepsilon}) u'(\bar{x}). \quad (7)$$

En conclusion, grâce à (4), (6) et (7), on obtient

$$\rho(\bar{x}, \tilde{\varepsilon}) \simeq -\frac{1}{2} \sigma_\varepsilon^2 \frac{u''(\bar{x})}{u'(\bar{x})}. \quad (8)$$

L'expression $-\frac{u''(\bar{x})}{u'(\bar{x})}$ est appelée l'*indice absolu d'aversion pour le risque* au niveau de richesse \bar{x} et est notée $r_a(\bar{x})$; on a donc

$$r_a(\bar{x}) = -\frac{u''(\bar{x})}{u'(\bar{x})} \simeq \frac{2\rho(\bar{x}, \tilde{\varepsilon})}{\sigma_\varepsilon^2}. \quad (9)$$

Notons que $r_a(\bar{x})$ est une mesure locale de l'aversion pour le risque (car elle est valable pour une variation $\tilde{\varepsilon}$ petite) et que cet indice est préférable aux paramètres "classiques" utilisés pour mesurer le degré de concavité de la courbe u au voisinage de \bar{x} , à savoir la dérivée seconde $u''(\bar{x})$ et le rayon de courbure

$$\frac{u''(\bar{x})}{[1 + (u'(\bar{x}))^2]^{\frac{3}{2}}}.$$

En effet, la fonction d'utilité est définie à une fonction affine croissante près, ce qui n'affecte pas $r_a(\bar{x})$, mais bien les deux autres paramètres puisque, si $v(x) = \alpha u(x) + \beta$, alors $v'(x) = \alpha u'(x)$ et $v''(x) = \alpha u''(x)$, d'où

$$\frac{v''(x)}{v'(x)} = \frac{u''(x)}{u'(x)}.$$

L'indice absolu d'aversion pour le risque permet de comparer le "degré de concavité" de la fonction d'utilité considérée. Pour deux fonctions d'utilité u_1 et u_2 supposées aimables, concaves et strictement croissantes sur $I =]0, +\infty[$, et deux fois continûment dérivables sur $\overset{\circ}{I} =]0, +\infty[$, on dispose de ce théorème (où $r_a^1(\bar{x})$, $r_a^2(\bar{x})$ désignent les indices absolus et $\rho_1(\bar{x}, \tilde{\varepsilon})$, $\rho_2(\bar{x}, \tilde{\varepsilon})$ les primes de risque correspondants) :

Théorème de Pratt (1964). *Soient les trois propositions suivantes :*

- (i) $r_a^1(\bar{x}) \geq r_a^2(\bar{x})$, $\forall \bar{x} > 0$;
- (ii) $\rho_1(\bar{x}, \tilde{\varepsilon}) \geq \rho_2(\bar{x}, \tilde{\varepsilon})$, $\forall \bar{x} > 0$, $\forall \tilde{\varepsilon}$ tel que $E(\tilde{\varepsilon}) = 0$ et $E(\tilde{\varepsilon}^2)$ petit ;
- (iii) il existe une fonction f concave sur I telle que $u_1 = f \circ u_2$.

On a les implications suivantes : (i) \Leftrightarrow (iii) \Rightarrow (ii).

Preuve. Posons $f = u_1 \circ u_2^{-1}$, fonction qui existe puisque u_2 admet une réciproque en tant que fonction strictement croissante sur I . Pour tout x de $\overset{\circ}{I}$, on a $f[u_2(x)] = u_1(x)$, d'où

$$f'[u_2(x)] = \frac{u_1'(x)}{u_2'(x)} \quad \text{et} \quad f''[u_2(x)] = \frac{u_1'(x)}{(u_2'(x))^2} \left[\frac{u_1''(x)}{u_1'(x)} - \frac{u_2''(x)}{u_2'(x)} \right].$$

Comme $u_1'(x) > 0$ et $u_2'(x) > 0$, on a donc

$$f''[u_2(x)] \leq 0 \quad \text{si et seulement si} \quad \frac{u_1''(x)}{u_1'(x)} \leq \frac{u_2''(x)}{u_2'(x)},$$

(c'est-à-dire si et seulement si $r_a^1(x) \geq r_a^2(x)$), d'où l'on déduit l'équivalence entre (i) et (iii).

Montrons à présent que (iii) implique (ii). L'égalité

$$u_1[\bar{x} - \rho_1(\bar{x}, \tilde{\varepsilon})] = E[u_1(\bar{x} + \tilde{\varepsilon})] = E[f[u_2(\bar{x} + \tilde{\varepsilon})]] \quad (10)$$

entraîne, par concavité de f (et l'inégalité de Jensen)

$$E[f[u_2(\bar{x} + \tilde{\varepsilon})]] \leq f[E[u_2(\bar{x} + \tilde{\varepsilon})]] = f[u_2[\bar{x} - \rho_2(\bar{x}, \tilde{\varepsilon})]] = u_1[\bar{x} - \rho_2(\bar{x}, \tilde{\varepsilon})]. \quad (11)$$

De (10) et (11), on tire

$$u_1[\bar{x} - \rho_1(\bar{x}, \tilde{\varepsilon})] \leq u_1[\bar{x} - \rho_2(\bar{x}, \tilde{\varepsilon})] \quad (12)$$

et, comme u_1 est strictement croissante

$$\bar{x} - \rho_1(\bar{x}, \tilde{\varepsilon}) \leq \bar{x} - \rho_2(\bar{x}, \tilde{\varepsilon})$$

ou encore

$$\rho_1(\bar{x}, \tilde{\varepsilon}) \geq \rho_2(\bar{x}, \tilde{\varepsilon}). \quad (13)$$

5. Indice relatif d'aversion pour le risque

Au lieu de considérer un risque “additif” (qui fait appel à une variable aléatoire de la forme $\tilde{x} = \bar{x} + \tilde{\varepsilon}$), on peut introduire un risque “proportionnel” qui donne naissance à une variable aléatoire $\tilde{x} = \bar{x}(1 + \tilde{\varepsilon})$, avec toujours $E(\tilde{\varepsilon}) = 0$ et $\sigma_{\tilde{\varepsilon}}^2$ petit, ainsi qu’une *prime de risque relative* $\rho^*(\bar{x}, \tilde{\varepsilon})$ définie par

$$E[u(\bar{x}(1 + \tilde{\varepsilon}))] = u[\bar{x}(1 - \rho^*(\bar{x}, \tilde{\varepsilon}))]. \quad (14)$$

Le produit $\bar{x}\rho^*(\bar{x}, \tilde{\varepsilon})$, qui coïncide avec $\rho(\bar{x}, \bar{x}\tilde{\varepsilon})$, est donc égal à la différence entre la valeur attendue $E(\tilde{x}) = \bar{x}$ et l'équivalent certain $EC(\tilde{x})$, d'où

$$\rho^*(\bar{x}, \tilde{\varepsilon}) = \frac{E(\tilde{x}) - EC(\tilde{x})}{\bar{x}}. \quad (15)$$

Des égalités

$$E[u(\bar{x}(1 + \tilde{\varepsilon}))] = E[u(\bar{x} + \bar{x}\tilde{\varepsilon})] = u[\bar{x} - \rho(\bar{x}, \bar{x}\tilde{\varepsilon})], \quad (16)$$

et par un raisonnement semblable à celui formulé dans le cas absolu, on trouve

$$\rho(\bar{x}, \tilde{\varepsilon}) = \bar{x}\rho^*(\bar{x}, \tilde{\varepsilon}) \simeq -\frac{1}{2}\bar{x}^2\sigma_{\tilde{\varepsilon}}^2\frac{u''(\bar{x})}{u'(\bar{x})}, \quad (17)$$

d'où l'on extrait

$$\rho^*(\bar{x}, \tilde{\varepsilon}) \simeq -\frac{1}{2}\sigma_{\tilde{\varepsilon}}^2\bar{x}\frac{u''(\bar{x})}{u'(\bar{x})}. \quad (18)$$

L'expression $\bar{x}\frac{u''(\bar{x})}{u'(\bar{x})}$, qui n'est rien d'autre que l'élasticité de l'utilité marginale en \bar{x} , permet de définir l'indice relatif d'aversion pour le risque au niveau de richesse \bar{x} , à savoir

$$r_r(\bar{x}) = -\bar{x}\frac{u''(\bar{x})}{u'(\bar{x})}. \quad (19)$$

De ce qui précède, il découle

$$r_r(\bar{x}) = \bar{x}r_a(\bar{x}) \simeq 2\frac{\rho^*(\bar{x}, \tilde{\varepsilon})}{\sigma_{\tilde{\varepsilon}}^2}. \quad (20)$$

De cette approximation (20), on déduit facilement que, pour deux fonctions d'utilité u_1 et u_2 aimables, concaves et strictement croissantes sur

$I = [0, +\infty[$ et deux fois continûment dérivables sur $\overset{\circ}{I}$, les conditions suivantes sont équivalentes :

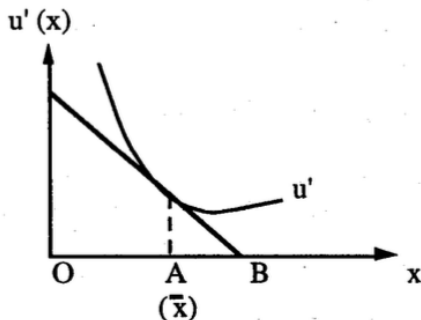
- (i) $r_r^1(\bar{x}) > r_r^2(\bar{x}), \forall \bar{x} > 0$
- (ii) $\rho_1^*(\bar{x}, \tilde{\varepsilon}) > \rho_2^*(\bar{x}, \tilde{\varepsilon}), \forall \bar{x} > 0, \forall \tilde{\varepsilon}$ tel que $E(\tilde{\varepsilon}) = 0$ et $E(\tilde{\varepsilon}^2)$ petit.

Tout comme l'indice absolu $r_a(\bar{x})$, l'indice relatif $r_r(\bar{x})$ est invariant par rapport à une transformation affine croissante. Par ailleurs, on peut donner de ces deux indices une interprétation géométrique à partir du graphique de l'utilité marginale u' . Par exemple, sur la figure qui suit, où la fonction u' est prise convexe, on voit aisément que

$$r_a(\bar{x}) = \frac{1}{|AB|},$$

tandis que

$$r_r(\bar{x}) = \frac{|OA|}{|AB|}.$$



6. Hypothèses d'Arrow et Pratt

Arrow (en 1965) et Pratt (1964) ont formulé deux hypothèses à propos du comportement des individus face au risque.

(i) *Hypothèse de décroissance de l'indice absolu.* On suppose que l'indice absolu d'aversion pour le risque détermine une fonction décroissante sur $[0, +\infty[$. On exige même souvent que $\frac{d}{dx} r_a(x) < 0$ pour tout $x > 0$. De l'approximation (9), on déduit que cette dernière condition est rencontrée lorsque $\frac{\partial}{\partial x} \rho(x, \tilde{\varepsilon}) < 0$ pour tout $x > 0$: cela revient à dire que la prime de

risque est une fonction décroissante de x ou, intuitivement, que le riche est plus tolérant vis-à-vis du risque que le pauvre.

La condition $\frac{d}{dx} r_a(x) < 0$ impose $u'''(x) > 0$ pour tout x , c'est-à-dire la stricte convexité de l'utilité marginale; en effet,

$$\frac{d}{dx} r_a(x) = \frac{[u''(x)]^2 - u'(x) u'''(x)}{[u'(x)]^2},$$

et $\frac{d}{dx} r_a(x) < 0$ entraîne $[u''(x)]^2 < u'(x) u'''(x)$, d'où $u'''(x) > 0$ puisque $u'(x) > 0$ (voir note 1 en bas de la page 63).

(ii) *Hypothèse de croissance de l'indice relatif.* On admet également que l'indice relatif d'aversion pour le risque est une fonction croissante sur $[0, +\infty[$ et même que $\frac{d}{dx} r_r(x) > 0$ pour tout $x > 0$. D'après (20), la dernière hypothèse équivaut à $\frac{\partial}{\partial x} \rho^*(x, \tilde{\varepsilon}) > 0$ pour tout $x > 0$: la prime de risque relative est une fonction croissante de x ou, intuitivement, le risque relatif grandit en même temps que la richesse risquée.

On vérifie sans peine que $u''(x) < 0$ et $u'''(x) < 0$ (d'où u et u' sont strictement concaves) entraînent $\frac{d}{dx} r_r(x) > 0$ pour tout $x > 0$.

7. Courbes d'indifférence

Considérons une loterie $\tilde{x} = [x_1, x_2; \Pi, 1 - \Pi]$ ou Π désigne un nombre donné (compris strictement entre 0 et 1); l'espérance de l'utilité vaut $E[u(\tilde{x})] = \Pi u(x_1) + (1 - \Pi) u(x_2)$. Les couples (x_1, x_2) pour lesquels l'espérance de l'utilité est constante forment l'ensemble

$$E = \{(x_1, x_2) \in \mathbb{R}^2 : \Pi u(x_1) + (1 - \Pi) u(x_2) = u_0\},$$

où u_0 désigne une constante connue; cet ensemble se représente, dans le plan $x_1 O x_2$ par la courbe d'équation implicite $\Pi u(x_1) + (1 - \Pi) u(x_2) = u_0$: cette courbe est appelée *courbe d'indifférence sans risque*.

La forme de telles courbes d'indifférence dépend, pour chaque individu, du goût ou de l'aversion pour le risque. En effet, par le théorème de dérivation des fonctions implicites, on trouve en tout point de E

$$\left. \frac{dx_2}{dx_1} \right|_{u=u_0} = - \frac{\Pi u'(x_1)}{(1 - \Pi) u'(x_2)} \quad (21)$$

$$\begin{aligned}
& \frac{d^2 x_2}{dx_1^2} \Big|_{u=u_0} \\
&= \frac{1}{((1-\Pi) u'(x_2))^3} \det \begin{pmatrix} 0 & \Pi u'(x_1) & (1-\Pi) u'(x_2) \\ \Pi u'(x_1) & \Pi u''(x_1) & 0 \\ (1-\Pi) u'(x_2) & 0 & (1-\Pi) u''(x_2) \end{pmatrix} \\
&= -\frac{\Pi}{1-\Pi} \frac{1}{u'(x_2)} \left[u''(x_1) + \frac{\Pi}{1-\Pi} \left(\frac{u'(x_1)}{u'(x_2)} \right)^2 u''(x_2) \right] \quad (22)
\end{aligned}$$

Dès lors, lorsque l'utilité marginale est strictement positive,

$$\frac{dx_2}{dx_1} \Big|_{u=u_0} < 0$$

et

$$\frac{d^2 x_2}{dx_1^2} \Big|_{u=u_0} > 0 \quad (\text{resp. } < 0)$$

dès que $u''(x) < 0$ pour tout x (resp. dès que $u''(x) > 0$ pour tout x).

Partant, si $u''(x) < 0$ pour tout x (cas d'aversion pour le risque), toute courbe d'indifférence sans risque est décroissante et tourne sa concavité vers le haut (c'est-à-dire est convexe); par contre, si $u''(x) > 0$ pour tout x (cas de goût pour le risque), toute courbe d'indifférence sans risque est décroissante et concave.

De plus, la courbure d'une courbe d'indifférence en un point de la bissectrice du premier quadrant dépend directement de l'indice absolu r_a et de la probabilité Π : de façon précise, en un tel point (x_1, x_2) , avec $x_1 = x_2$, on a

$$\frac{d^2 x_2}{dx_1^2} \Big|_{\substack{x_1=x_2 \\ u=u_0}} = -\frac{\Pi}{1-\Pi} \frac{u''(x_1)}{u'(x_1)} \left(1 + \frac{\Pi}{1-\Pi} \right) = r_a(x_1) \frac{\Pi}{(1-\Pi)^2}. \quad (23)$$

Dans la pratique, on fait le plus souvent appel à des fonctions d'utilité u , construites à partir de fonctions élémentaires classiques, telles que les expressions $u'(x)$ et $\frac{d}{dx} r_r(x)$ sont positives (ou nulles), tandis que $u''(x)$ et $\frac{d}{dx} r_a(x)$ sont négatives (ou nulles). En guise d'exemples, citons les fonctions

suivantes :

- i) $u(x) = a + b \log(x + c)$ avec $a, b > 0$, $c > 1$ et $x > 0$
- ii) $u(x) = a - \frac{1}{x+b}$ avec $a \geq 0$, $b > 0$ et $x \geq 0$
- iii) $u(x) = -e^{-ax}$ avec $a > 0$ et $x \in \mathbb{R}$
- iv) $u(x) = x^{1-a}$ avec $0 < a < 1$ et $x > 0$
- v) $u(x) = -x^{-(a-1)}$ avec $a > 1$ et $x > 0$.

Bibliographie

- [1] Bair J., Equations fonctionnelles et mathématique financière, *Mathématique et Pédagogie*, 117, 1998, pp. 49-58.
- [2] Bair J., Haesbroeck G., Hansen G., Philippart S., *Fonctions convexes : théorie et applications économiques*, Dossier du Gemme n° 1, Université de Liège, année académique 1995-1996.
- [3] Broquet C., Cobbaut R., Gillet R. et van den Berg A., *Gestion de portefeuille*, De Boeck Université, Bruxelles, 1997.
- [4] Janssen J., *Traité de mathématique appliquée pour l'assurance, l'économie, la finance*, vol. 2, Office International de Librairie, Bruxelles, 1993.
- [5] Laffont J.J., *Economie de l'incertain et de l'information*, vol. 2, Economica, Paris, 1991.
- [6] Philipps O.T., Ravindran A., Solberg J.J., *Operations research : principles and practice*, J. Wiley & Sons, New York, 1987.
- [7] Roger P., *Les outils de la modélisation financière*, P.U.F., Paris, 1991.
- [8] Takayama A., *Analytical Methods in Economics*, Harvester Wheatsheaf, New York, 1994.
- [9] Winston W.L., *Operations research : applications and algorithms*, Duxbury Press, Boston, 1987.

Adresse de l'auteur :

G. Haesbroeck

Université de Liège

Faculté d'Economie, de Gestion et de Sciences Sociales

Boulevard du Rectorat 7, bâtiment B31

4000 LIEGE

Belgique

Une courte nuit pour un grand jour : le 11 août 1999!

F. Clette, J.-R. Gabryl,

L'événement

A l'idée de courte nuit est directement associée celle de solstice d'été, qui se produira cette année le 21 juin. Que se passera-t-il donc le mercredi 11 août 1999, presque deux mois après le solstice ?

En plein milieu de la journée, va se produire une éclipse totale de soleil qui sera visible dans l'extrême Sud de la Belgique, mais aussi à travers toute l'Europe. Bien à l'avance, on en parle déjà dans des livres, les journaux et à la télévision.

Cet événement est sans doute le plus grandiose que la mécanique céleste puisse nous offrir en spectacle. Imaginez en effet que dans son mouvement orbital autour de notre planète, la lune va passer entre le soleil et la terre. L'alignement sera suffisamment précis pour que son ombre se projette sur la surface terrestre. Cette ombre parcourera le globe terrestre durant 3 heures, à une vitesse d'environ 3000 km/h, plus vite que les avions supersoniques. Elle est entourée d'une zone de pénombre plus vaste. Dans cette zone, le soleil n'est jamais complètement masqué (éclipse partielle), un phénomène curieux mais assez courant puisqu'on a pu en voir récemment, en mai 1994 et octobre 1996, en Belgique.

En revanche, pour les témoins privilégiés qui se trouveront sur la trajectoire de l'ombre (Fig. 1), le soleil sera complètement caché durant quelques minutes (2 minutes maximum en Belgique, 2 min 27 s au maximum en Roumanie centrale). Un spectacle étonnant et toujours très impressionnant à vivre soi-même, l'observateur étant "enveloppé" dans ce phénomène de dimension cosmique : obscurité crépusculaire, la couronne dans le ciel étoilé, rafraîchissement de la température, perturbation des animaux, des plantes et des activités humaines. Cette totalité ne se produit pas d'un coup, elle est précédée et suivie d'une phase partielle de 1 h 15 m environ au cours de laquelle la lune grignote lentement le soleil.

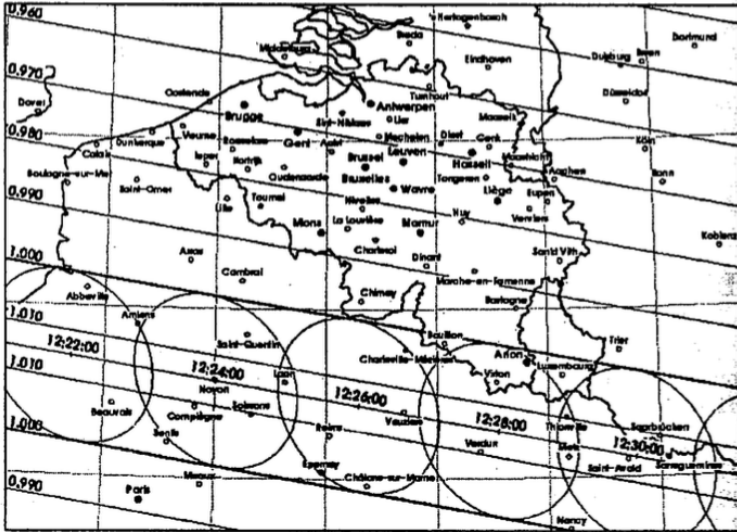


Fig. 1 : Trajectoire de l'ombre en Belgique

Un brin de mécanique céleste

Sachant que le cycle des phases lunaires (période synodique) est de 29, 53 jours et que la lune repasse donc souvent entre la terre et le soleil (Nouvelle lune), il est légitime de se demander pourquoi les éclipses sont un événement si rare (2 éclipses solaires sur trois ans sur tout le globe de la terre, mais intervalles moyens de plusieurs siècles en un lieu donné).

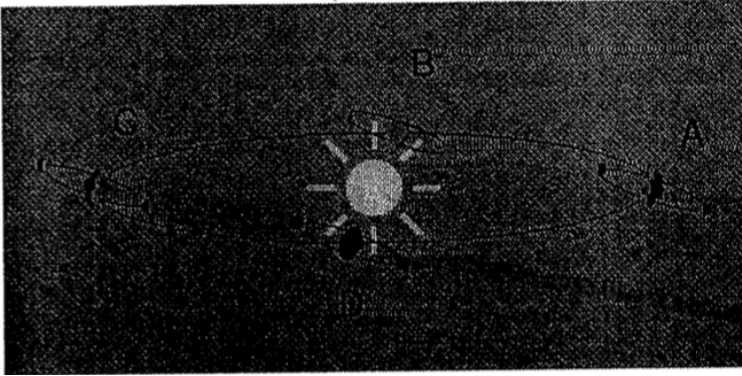


Fig. 2 : Mouvements de la terre et de la lune

L'explication tient en peu de mots : terre et lune ne circulent pas dans le même plan (Fig. 2). En effet, le plan dans lequel la lune se meut autour de la terre est incliné d'environ 5° par rapport à l'écliptique, qui est le plan de révolution de la terre autour du soleil. Le lieu commun de ces deux plans, une droite, s'appelle la ligne des noeuds. La condition pour qu'une éclipse se produise est que la ligne des noeuds pointe vers le soleil et que la lune se situe simultanément à l'un des noeuds (éclipse de soleil au noeud intercalé entre le soleil et la lune, éclipse de lune au noeud opposé). Sinon, l'ombre de la lune passe à côté de la terre, et l'ombre de la terre passe à côté de la lune (Fig. 3).

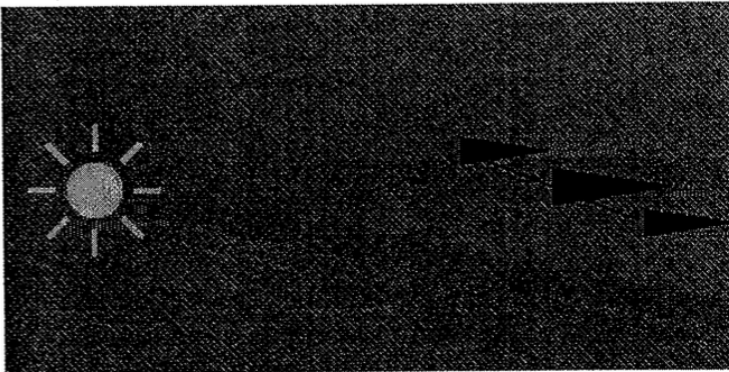


Fig. 3 : Projection des ombres

Si le plan orbital lunaire était parfaitement fixe dans l'espace, la ligne des noeuds resterait constamment parallèle à une direction donnée. On aurait donc un alignement entre la ligne des noeuds et le soleil tous les six mois, et un noeud donné dans cet alignement ferait face au soleil une fois par an, à la même date anniversaire (Fig. 2). Mais ce n'est pas le cas : la dynamique du système solaire entraîne ce plan dans un lent mouvement de rotation par rapport à l'écliptique ($0,05^\circ$ par jour). Du coup, la ligne des noeuds ne conserve pas une orientation constante. La précession qui l'anime va la réaligner sur le soleil en un peu moins de six mois, et il ne faudra plus que 346,62 jours pour qu'un même noeud se représente face au soleil. Toutefois, cela ne signifie pas encore que l'on aura une éclipse de soleil à chaque fois. En effet, il n'y a pas un nombre entier de lunaisons dans cette période de 346,62 jours !

Il existe par contre, par un hasard assez extraordinaire, un multiple commun entre la période de lunaison et la période de retour d'un noeud : 6585 jours (en arrondi), c'est-à-dire 223 lunaisons ou 19 "retours de noeud", soit encore 18 ans 11 jours et 8 heures. Ce cycle appelé Saros, avait déjà été remarqué au temps des Chaldéens (1000 ACN). On peut donc prendre rendez-vous dans 18 ans mais pas au même endroit. A cause des 8 heures supplémentaires, la terre aura tourné d'un tiers de tour et l'ombre de la lune se produira environ 120° de longitude plus à l'Ouest.

Les résultats d'observations d'éclipses

Les éclipses totales de soleil sont généralement le point de rendez-vous d'équipes scientifiques, principalement des physiciens solaires. En effet, durant la totalité, le disque aveuglant du soleil (photosphère, attention aux yeux !) disparaît exactement derrière la lune. Coïncidence extraordinaire mais vraiment fortuite (quelle chance !), les diamètres apparents du soleil et de la lune sont identiques. Le ciel est alors suffisamment obscurci pour découvrir l'atmosphère extérieure du soleil, la couronne (Fig. 4), plus d'un million de fois moins lumineux que la photosphère, et cela depuis quasi le limbe solaire jusqu'à plus de 4 rayons solaires au-dessus de celui-ci (Fig. 6).

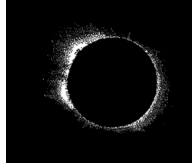


Fig. 4 : La couronne du 26/02/1998

Un spectacle grandiose mais aussi à exploiter scientifiquement pour en savoir plus sur cette atmosphère tellement chaude (un million de degrés) que les astrophysiciens, qui n'ont pas voulu y croire eux-mêmes, n'y ont rien compris pendant plus de 50 ans. En outre, la couronne est en expansion continue et emplit tout le système solaire d'un flux de particules électrisées (plasma) par lequel la terre, et sa magnétosphère, sont soumis à l'influence solaire de façon continue. Ce flux de matière est appelée le vent solaire (Fig. 5).

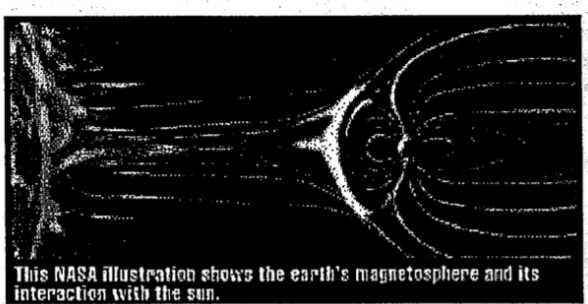


Fig. 5 : Interaction soleil-terre via le vent solaire

Ce vent solaire n'est pas uniforme, mais secoué de bourrasques, appelées éjections de masse coronales. Provoquées notamment par des éruptions solaires, de tels événements génèrent les aurores polaires, mais aussi des perturbations pour la technologie actuelle, par certains côtés si fragile : télécommunications, engins spatiaux, lignes haute-tension, etc.

Nous savons maintenant que la couronne varie en permanence sous le contrôle des champs magnétiques qui lui donnent sa structure très inhomogène : grands jets en fuseau étiré, intenses condensations coronales en arcade, cavités, feuillets, plumes polaires. Ces champs magnétiques à grande

échelle se déploient dans la couronne mais sont ancrés en profondeur sous la photosphère. Ils sont générés par un mécanisme de “dynamo”, en action à la base de la couche convective. A ce titre, l’observation des structures coronales, pourtant éloignées de la surface solaire, peut nous donner des indices sur ce qui se produit dans les profondeurs inaccessibles. C’est un des grands thèmes dans la quête des astrophysiciens depuis de longues années.

La structure magnétique coronale est en évolution permanente sous l’action de l’activité solaire, c’est-à-dire en gros, le rythme auquel du nouveau flux émerge à la surface, ce qui se traduit invariablement par un feu d’artifice d’éruptions accompagnées de jets de particules et d’émission de rayonnements durs (courte longueur d’onde : ultraviolet et rayon X), d’où le nom d’“activité”.

Or, l’activité solaire est modulée périodiquement selon un cycle long d’une durée moyenne de 11 ans (Fig. 6). Ce cycle, encore incompris malgré qu’il ait été découvert il y a plus de 150 ans, constitue une des grandes énigmes non résolues du soleil, importante pour comprendre l’impact à long terme de notre étoile sur la stabilité de l’environnement habitable par la vie sur terre. Comme on peut s’y attendre, l’apparence de la couronne change considérablement au cours de ce cycle.

En haute activité, condensations et jets émergent radialement dans toutes les directions (symétrie circulaire). En phase de faible activité, quelques jets émergent parallèlement à l’équateur solaire avec des régions polaires sombres parsemées de plumes polaires (forme aplatie) (Fig. 6). Les indices de l’activité solaire, couvrant presque 2 siècles, sont collectés, et des prévisions sont disséminées au niveau mondial depuis l’Observatoire Royal de Belgique (Sunspot Index Data Center).

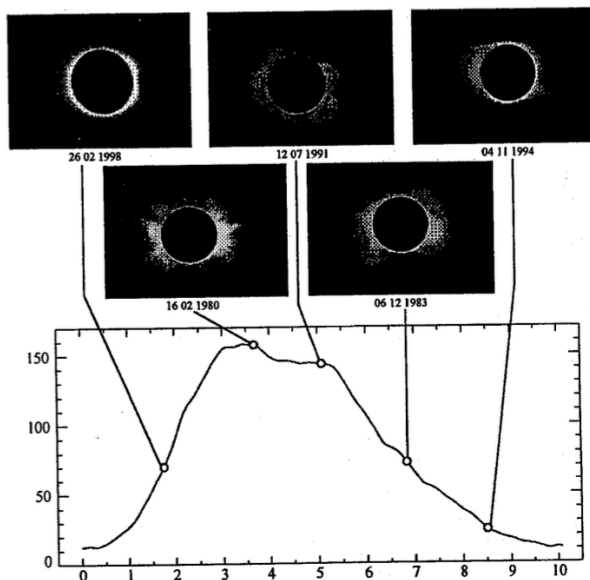


Fig. 6 : Morphologie de la couronne en fonction du cycle d'activité solaire

Se préparer à l'éclipse du 11 août prochain

Tout un chacun dispose en fait d'une place de premier rang pour admirer l'éclipse. L'œil nu constitue la meilleure approche, avec éventuellement en appoint des jumelles ou une petite lunette. La couronne, clou du spectacle, offre des dimensions généreuses (2-3 degrés, 4 fois le diamètre solaire). Aux instants où commence et finit la totalité, guettez l'apparition fulgurante et brève des grains de Baily (derniers rayons de soleil filtrant par les volées lunaires) ou de la chromosphère et des protubérances, seule source de couleur pendant l'éclipse (émission rouge vif de l'hydrogène).

Observez l'ambiance autour de vous : sons, température, éclairage, comportements. Le ciel assombri laissera voir quelques planètes et les étoiles les plus brillantes. Pour les reconnaître, on peut se référer à la position du soleil éclipsé. Partant de là, on peut toiser les distances angulaires à l'aide de la main tendue à bout de bras (poing fermé = 10° , mais étendue du pouce à l'auriculaire = 20°).

Vous l'avez compris, aucun filtre de protection n'est nécessaire durant les 2 minutes de la totalité. En revanche, tout au long des phases partielles et tant qu'un fragment de photosphère subsiste, le soleil est aussi aveuglant qu'à l'ordinaire. **Pas question de regarder le soleil sans protection adéquate**, on risque de perdre définitivement la vue par brûlure de la rétine. Histoire d'être sûr, utilisez des lunettes d'éclipses homologuées (label CE), utilisant des filtres noirs ou métallisés (atténuation nécessaire 100.000 fois!), qui seront en vente avant l'éclipse. Les recettes de "grand-mères" dangereuses ne manquent pas. N'essayez pas ! Les yeux vides pour la vie pour un jeu de 30 secondes avec un soleil en éclipse ? A votre avis, ça vaut le coup ?

Avec tout appareil d'optique, la radiation solaire est encore plus concentrée. Pour éviter tout danger, le plus simple et efficace (surtout en groupe) est d'utiliser l'image projetée sur un écran. Surtout ne pas approcher de l'oculaire : une surveillance est nécessaire pour les distraits. Record de la sécurité et de la simplicité : la chambre obscure. Il peut s'agir d'une boîte vide en carton percée au centre d'une des faces d'un trou d'aiguille ou d'un miroir masqué d'un carton percé pour projeter une image du soleil en croissant sur n'importe quel écran (un mur, le sol, etc).

Ce n'est pas tous les jours non plus que le soleil se déguise en croissant de lune. Alors prêts pour le 11 août ? Nous croisons les doigts et nous vous souhaitons un ciel tout dégagé avec le "grand oeil noir" au milieu.

Renseignements complémentaires

- Département de Physique Solaire, Observatoire Royal de Belgique, Av. Circulaire 3, 1180 Bruxelles.
- Journées d'informations au Planétarium du Heysel, Av. De Bouchout 10, 1090 Bruxelles, spécialement pour les enseignants. Dates proposées : 23/04 et 05/05 dès 14 h. Réservation à l'Observatoire.
- Web : <http://www.oma.be/KSB-ORB/INFO/menu.html>.
- Téléphone "Eclipse" : 02/373.67.77.
- Brochure "Dans l'ombre de la Lune", 66 pages, 250 F (6,20 euros) à verser sur le compte 000-0346177-81 de l'Observatoire Royal de Belgique, avec la mention "Brochure Eclipse en Français".
- Posters sur commande.
- Cassette vidéo sur commande.

Revue des revues

M. Fremal,

Math-Ecole (revue suisse), n° 174-octobre 1996.

Au sommaire :

- **François Jaquet**, *Editorial : Est-ce la fin des maths modernes ?*
- **Daniela Medici et Paola Vighi**, *Une histoire ... improbable*
Les auteurs ont introduit les probabilités à l'école primaire. Plusieurs situations, expérimentées dans les classes, sont proposées.
- **Pierre Favre**, *ICME - 8ème congrès international*
Cette rencontre s'est tenue à Séville en juillet 1996, elle a réuni plus de 4.000 personnes. L'auteur nous livre ses impressions.
- **Michel Chastellain**, *CABRidées : une machine à multiplier*
La moyenne géométrique, les triangles semblables et une macro construction "report de segment" permettent de construire, à l'aide de *Cabri-Géomètre*, un segment dont la mesure est égale au produit des mesures de deux segments donnés.
- **François Jaquet**, *Entre addition et multiplication*
L'auteur étudie les effets des énoncés de problèmes sur les représentations et stratégies des élèves. Le thème des parentés entre situations additives et multiplicatives est étudié.
- **Colomba Boggini et François Jaquet**, *"Le Puzzle"*
De nouveaux moyens romands d'enseignement "Mathématique 1P" sont mis en place en première année de l'école primaire. Plusieurs activités sont expérimentées dans les classes, celle intitulée "Le Puzzle" est décrite.
- **Antoinette Raccio**, *Participer au rallye mathématique romand avec une classe de 1ère-2ème-3ème années, possible ?*
L'auteur décrit une expérience qui s'est déroulée dans une classe regroupant quatorze élèves de trois degrés différents de l'école primaire. Plusieurs questions sont posées :
 - jusqu'à quel point les interactions entre les élèves peuvent-elles être profitables aux trois degrés dans un travail de groupe ?
 - quel peut être l'apport du travail de groupe, degrés mélangés, pour les élèves des trois degrés ?
 - comment s'organise et travaille un groupe comprenant des élèves de degrés différents ?

Deux problèmes du 3ème Rallye mathématique servent de support à l'expérience décrite dans l'article. En principe, ces exercices ne devaient être accessibles qu'aux élèves de troisième.

En conclusion, les interactions entre les trois degrés sont riches d'apports pour tous les élèves.

– **Michel Brêchet**, *L'aire et les erreurs*

Lors de l'acquisition des connaissances nécessaires à la maîtrise de l'aire du parallélogramme, des élèves "faibles" de 13-14 ans rencontrent des difficultés.

L'auteur décrit et explique certaines de celles-ci, il propose aussi quelques pistes pour donner du sens au concept d'aire.

– **Christophe Mironneau**, *Petite réflexion sur quelques problèmes de logique*

Quelques problèmes proposés dans le recueil de Schneider, E. et J.-B., intitulé *69 problèmes de logique pour apprendre à raisonner aux enfants de 9 à 13 ans*, ont été expérimentés. L'auteur nous fait part de ses réflexions.

– *Notes de lecture*

- Charnay R. et Mante M., Préparation à l'épreuve de mathématiques du concours de professeur des écoles - Tomes I et II, Paris, Hatier pédagogie, 1995.
- Charnay R., Pourquoi des mathématiques à l'école?, Paris, ESF éditeur, 1996.

Math-Ecole (revue suisse), n° 175-décembre 1996.

Au sommaire :

- **Luc-Olivier Pochon**, *Editorial* : <http://www.unine.ch/irdp/math-eco/>
Brève analyse du phénomène de l'informatique à l'école.
- **Marianne Nicole**, *Des mathématiques gustatives*
Compte rendu d'une expérience qui s'est déroulée avec des enfants présentant un trouble du langage oral et écrit.
- **Dominique Valentin**, *Vive les livres à compter!*
Ce sont souvent des activités ludiques qui permettent aux jeunes enfants de donner du sens aux nombres. L'équipe de didactique de l'INRP de Paris a parfois recours aux livres à compter pour entraîner les enfants au dénombrement. Le travail réalisé par deux institutrices de deuxième année d'école maternelle est présenté.
Les compétences développées se situent sur plusieurs plans : transversal, langagier et mathématique.
- **Michel Chastellain**, *CABRidées : Un support didactique pour une introduction à l'axonométrie*
Ce mode de représentation des solides est assez abstrait pour les élèves de 8ème et 9ème années scientifiques. L'apport pédagogique et didactique de Cabri-géomètre se révèle intéressant pour l'étude des représentations axonométriques.
- *5ème Rallye mathématique transalpin*
Ce concours s'adresse à des classes réparties par régions, selon quatre catégories (3ème, 4ème, 5ème et 6ème années de l'école primaire). Les problèmes de l'épreuve d'essai sont présentés.
- *Découpages de carrés en triangles semblables*
Le problème suivant est étudié : combien existe-t-il de découpages d'un carré en six triangles semblables? Quelques solutions sont proposées, mais combien y en a-t-il : 89, 93, 97, ...? Le concours est ouvert.
- **Isabelle Vogt**, *La maturité professionnelle : une nouvelle ouverture pour l'apprentissage*
- "Histoires" de chercher ...
 - Francis Perret (NAJAROS), mathémagicien, propose deux problèmes dont celui du carré supermagique 1996.
 - Un problème pour les élèves (dès 10-12 ans) : *On souhaite partager un quadrillage carré en deux parties isométriques en ne suivant*

que les traits du quadrillage ou ses diagonales. Combien y a-t-il de solutions différentes ?

– *La revue des revues*

– *Notes de lecture*

- Panoramath 96, Comité International des Jeux mathématiques. Coédition C.I.J.M.-A.P.M.E.P.-ACL, Paris, 1996.
- Jeux 4, de l'intérêt des problèmes de rallyes, publication de l'APMEP, n° 97, Paris, 1995
- Ces problèmes qui font les mathématiques : Enseigner la géométrie dans l'espace au collège et au lycée par Bernard Destainville, brochure n° 99 de l'APMEP
- Faire de la Géométrie en jouant avec Cabri-géomètre par Roger Cuppens, brochures n° 104 et 105 de l'APMEP
- Mathematica dans les lycées par Robert Cabessa, LEP, Lausanne, 1996.

M. FREMAL

Olympiades

C. Festraets,

Voici les solutions des problèmes posés le second jour à l'Olympiade Mathématique Internationale de 1998.

Problème 4. *Trouver tous les couples (a, b) d'entiers strictement positifs tels que $ab^2 + b + 7$ divise $a^2b + a + b$.*

Solution.

$ab^2 + b + 7$ divise $a^2b + a + b$, donc il existe un entier non nul k tel que

$$\begin{aligned} k &= \frac{a^2b + a + b}{ab^2 + b + 7} \\ &= \frac{1}{b} \left(\frac{a^2b^2 + ab + b^2}{ab^2 + b + 7} \right) \\ &= \frac{1}{b} \left(a + \frac{b^2 - 7a}{ab^2 + b + 7} \right). \end{aligned} \tag{1}$$

1) Si $b^2 - 7a = 0$, alors $a = \frac{b^2}{7}$ et $k = \frac{b}{7}$.

En posant $b = 7t$, avec $t \in \mathbb{N}^*$, on obtient les solutions $(a, b) = (7t^2, 7t)$.

2) Si $b^2 - 7a > 0$, il faut que

$$\begin{aligned} b^2 - 7a &\geq ab^2 + b + 7 \\ b^2(1 - a) &\geq 7a + b + 7, \end{aligned}$$

ce qui est impossible car $a \geq 1$.

3) Si $b^2 - 7a < 0$, il faut que

$$\begin{aligned} 7a - b^2 &\geq ab^2 + b + 7 \\ a(7 - b^2) &\geq b^2 + b + 7 \end{aligned}$$

comme le 2ème membre est positif, les seules valeurs possibles sont $b = 1$ ou $b = 2$.

Pour $b = 1$, l'égalité (1) devient

$$k = a - \frac{7a - 1}{a + 8} = a - 7 + \frac{57}{a + 8}.$$

$a + 8$ doit être un diviseur de 57 strictement supérieur à 8; donc deux possibilités,

$$\begin{aligned} a + 8 &= 19, \text{ d'où } a = 11 \\ a + 8 &= 57, \text{ d'où } a = 49, \end{aligned}$$

ce qui nous fournit deux solutions $(11, 1)$ et $(49, 1)$.

Pour $b = 2$, l'égalité (1) devient

$$k = \frac{1}{2} \left(a - \frac{7a - 4}{4a + 9} \right).$$

$7a - 4$ doit être divisible par $4a + 9$, autrement dit, le pgcd de $7a - 4$ et de $4a + 9$ doit être $4a + 9$.

Calculons ce pgcd :

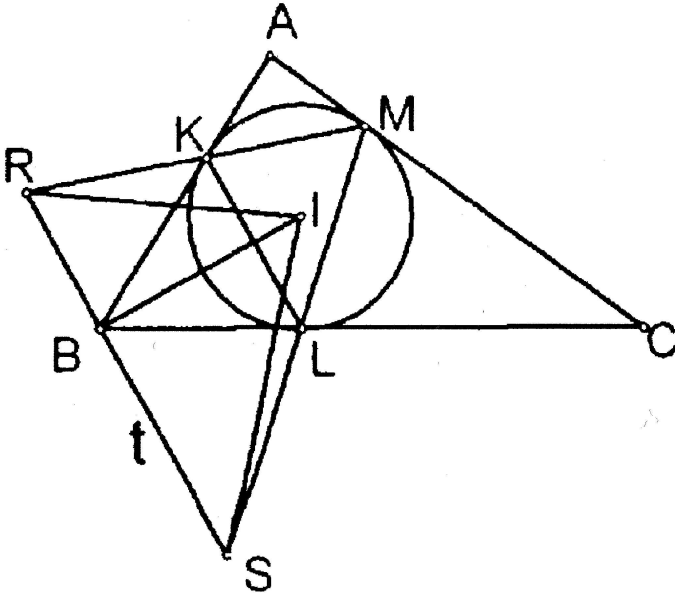
$$\begin{aligned} (7a - 4, 4a + 9) &= (-a - 22, 4a + 9) \\ &= (a + 22, 79) \\ &= 1 \text{ ou } 79. \end{aligned}$$

$4a + 9 = 1$ et $4a + 9 = 79$ n'ont pas de solution entière pour a .

Les seules solutions du problème sont donc $(a, b) = (7t^2, 7t)$ avec $t \in \mathbb{N}^*$, $(a, b) = (11, 1)$ et $(a, b) = (49, 1)$.

Problème 5. *Soit I le centre du cercle dans le triangle ABC . Ce cercle est tangent aux côtés BC, CA et AB du triangle, en les points L, M et K respectivement. La droite parallèle à KL passant par B coupe les droites MK et ML respectivement en R et S . Prouver que l'angle RIS est aigu.*

Solution.



$$\begin{aligned}
 \widehat{RKB} &= \widehat{MKA} \\
 &= \widehat{MLK} \text{ (interceptent le même arc)} \\
 &= \widehat{LSB} \text{ (} KL // LS \text{)} \\
 \widehat{RBK} &= \widehat{LKB} \text{ (} KL // RS \text{)} \\
 &= \widehat{KLB} \text{ (triangle isocèle } KBL \text{)} \\
 &= \widehat{SLB} \text{ (} KL // RS \text{)}.
 \end{aligned}$$

Les triangles KRB et SBL sont donc semblables et

$$\frac{BK}{BS} = \frac{RB}{BL}$$

et comme $BK = BL$, on a

$$BK^2 = BS \cdot RB. \tag{2}$$

$RS \parallel KL$ et $IB \perp KL$, donc $IB \perp RS$ et dans les triangles rectangles IBR , IBS , on a

$$\begin{aligned} IR^2 &= RB^2 + IB^2 \\ IS^2 &= BS^2 + IB^2 \\ IR^2 + IS^2 &= RB^2 + BS^2 + 2IB^2 \\ &= (RB + BS)^2 - 2RB \cdot BS + 2IB^2 \end{aligned}$$

en tenant compte de (2)

$$\begin{aligned} IR^2 + IS^2 &= RS^2 - 2BK^2 + 2IB^2 \\ &= RS^2 + 2(IB^2 - BK^2). \end{aligned}$$

Le triangle IKB est rectangle, d'où $IB^2 - BK^2 = IK^2$ et l'égalité précédente devient

$$IR^2 + IS^2 - RS^2 = 2IK^2 > 0.$$

Or,

$$IR^2 + IS^2 - RS^2 = 2IR \cdot IS \cdot \cos \widehat{RIS}$$

donc $\cos \widehat{RIS} > 0$ et l'angle \widehat{RIS} est aigu.

Problème 6. On considère toutes les applications f de l'ensemble \mathbb{N}^* de tous les entiers strictement positifs dans lui-même vérifiant

$$f(t^2 f(s)) = s(f(t))^2,$$

quels que soient s et t dans \mathbb{N}^* .

Déterminer la plus petite valeur possible de $f(1998)$.

Solution.

1) $\forall x \in \mathbb{N}^*$,

$$f(f(x)) = f(1^2 f(x)) = x f^2(1) = k^2 x. \quad (3)$$

2) $\forall x, y \in \mathbb{N}^*$,

$$f(f(x)) = k^2 x \quad \text{et} \quad f(f(y)) = k^2 y$$

si $f(x) = f(y)$, alors $x = y$, donc f est une injection de \mathbb{N}^* dans \mathbb{N}^* .

3) $\forall x \in \mathbb{N}^*$,

$$f(kx^2) = f(x^2 f(1)) = 1 \cdot (f(x))^2 = f^2(x). \quad (4)$$

4) $\forall x, y \in \mathbb{N}^*$,

$$\begin{aligned} [f(x) \cdot f(y)]^2 &= f^2(x) \cdot f^2(y) \\ &= f^2(x) \cdot f(ky^2) && \text{par (4)} \\ &= f(x^2 f(f(ky^2))) && \text{par définition de } f \\ &= f(x^2 k^2 ky^2) && \text{par (3)} \\ &= f(k(kxy)^2) \\ &= f^2(kxy) && \text{par (4)} \end{aligned}$$

d'où

$$f(x) \cdot f(y) = f(kxy) \quad (5)$$

en particulier, si $y = 1$

$$f(x) \cdot k = f(kx)$$

et (5) devient

$$f(x) \cdot f(y) = kf(xy). \quad (6)$$

Si on fait $x = y$, on obtient

$$f^2(x) = kf(x^2)$$

et par induction, pour tout $t \in \mathbb{N}^*$,

$$[f(x)]^t = k^{t-1} f(x^t). \quad (7)$$

5) Soit p un nombre premier, p^α la plus haute puissance de p divisant $f(x)$ et p^β la plus haute puissance de p divisant k .

De (7), il vient

$$\begin{aligned} \alpha t &\geq \beta(t-1) \\ \alpha &\geq \beta \frac{t-1}{t}. \end{aligned}$$

Cette inégalité est vraie quel que soit $t \in \mathbb{N}^*$, d'où

$$\begin{aligned} \alpha &\geq \beta \\ p^\alpha &\geq p^\beta \end{aligned}$$

et $\forall x \in \mathbb{N}^*$, k est un diviseur de $f(x)$.

6) Soit p un nombre premier ; $\frac{1}{k} f(p)$ est un entier positif.

Supposons que

$$\begin{aligned}\frac{1}{k} f(p) &= ab \\ f(f(p)) &= f(kab) \\ k^2 p &= f(a) \cdot f(b) \quad \text{par (3) et (5)} \\ p &= \frac{f(a)}{k} \cdot \frac{f(b)}{k}.\end{aligned}$$

$\frac{f(a)}{k}$ et $\frac{f(b)}{k}$ sont entiers, mais p est premier, donc l'un des deux facteurs est égal à 1.

Soit $\frac{f(a)}{k} = 1$, alors $f(a) = k = f(1)$, d'où $a = 1$ car f est injective.

$\frac{1}{k} f(p)$ ne peut être un produit de deux facteurs, donc, quel que soit p premier, $\frac{1}{k} f(p)$ est premier.

7) $f(1998) = f(2 \cdot 3^3 \cdot 37) = \frac{1}{k^5} f(2) \cdot [f(3)]^3 \cdot f(37)$.

Pour minimiser $f(1998)$, définissons f par

$$\begin{aligned}f(1) &= k = 1 \\ f(2) &= 3, & f(3) &= f(f(2)) = 2 \\ f(37) &= 5, & f(5) &= f(f(37)) = 37 \\ f(p) &= p, & & \text{pour tout } p \text{ premier distinct de } 2, 3, 5, 37.\end{aligned}$$

f est totalement définie, car $\forall n \in \mathbb{N}^*$,

$$n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \dots p_t^{\alpha_t} \quad (\text{les } p_i \text{ étant premiers})$$

$$\Rightarrow f(n) = [f(p_1)]^{\alpha_1} \cdot [f(p_2)]^{\alpha_2} \dots [f(p_t)]^{\alpha_t}$$

et on obtient

$$f(1998) = 3 \cdot 2^3 \cdot 5 = 120.$$

Des problèmes et des jeux

C. Festraets,

Quelle puissance problème n° 211 de M. et P. n° 119.

Soit $n = (1998^{1998})^{1998}$. Désignons par s la somme des chiffres de n , par t celle de s et par u celle de t . Que vaut u ?

Solution de J. JANSSEN de Lambermont.

$n = (1998^{1998})^{1998}$. Comme 1998 est un multiple de 9, le nombre n , puissance et donc multiple de 1998, le sera aussi de même que s , la somme de ses chiffres ; ce sera encore pareil pour t et pour u .

Majorons n en partant de $n' = (2000^{2000})^{2000}$

$$\begin{aligned} n &< n' = 2000^{4.000.000} = 2^{4.000.000} \cdot 10^{3 \times 4.000.000} \\ &< (2^3)^{1.333.334} \cdot 10^{12.000.000} < 10^{1.333.334} \times 10^{12.000.000} = 10^{13.333.334}, \end{aligned}$$

donc $n < 10^{13.333.334}$.

n ne peut donc comprendre plus de 13.333.335 chiffres et est un multiple de 9. La somme de ces chiffres, s (multiple de 9), doit donc être inférieure ou égale à

$$9 \times 13.333.334 = 120.000.006$$

et comme s n'est pas nul, on a

$$9 \leq s \leq 120.000.006.$$

t , la somme des chiffres de s , est aussi un multiple de 9 et tel que

$$9 \leq t \leq 8 \times 9 = 72.$$

Et comme u est encore un multiple de 9, on obtient

$$9 \leq u \leq 9 \text{ soit } \boxed{u = 9}.$$

N. BERCKMANS de Waterloo, P. DASSY de Liège, J. FINOULST de Diepenbeek et C. VAN HOOSTE de Marbaix-la-Tour ont envoyé des solutions similaires.

Soyons constructifs | problème n° 212 de M. et P. n° 119.

On donne dans le plan trois cercles concentriques et tels que le plus grand des trois rayons est inférieur à la somme des deux autres. Construire un triangle équilatéral dont chaque sommet appartient à un cercle différent.

Solution de C. VAN HOOSTE de Marbaix-la-Tour.

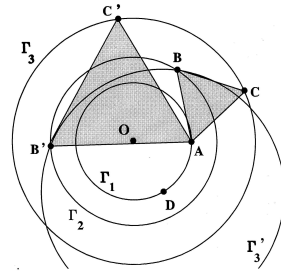
Soient $\Gamma_1, \Gamma_2, \Gamma_3$ les cercles de centre O , respectivement de rayons r_1, r_2 et r_3 tels que

$$r_1 < r_2 < r_3 \leq r_1 + r_2.$$

Si ABC est un triangle équilatéral tel que

$$A \in \Gamma_1, B \in \Gamma_2, C \in \Gamma_3,$$

alors B est l'image de C par la rotation de centre A et d'angle 60° (dans le sens trigonométrique).



Cette même rotation applique Γ_3 sur le cercle Γ'_3 de rayon r_3 et de centre D tel que le triangle OAD soit équilatéral puisque D doit être l'image de O par ladite rotation.

Le point B doit nécessairement être un des points d'intersection de Γ_2 et de Γ'_3 .

De cette courte analyse, nous déduisons le procédé de construction du triangle ABC :

- choisir un point A quelconque de Γ_1 ;
- construire D appartenant à Γ_1 tel que $|AD| = r_1$;
- tracer le cercle Γ'_3 de centre D et de rayon r_3 , cercle qui coupe Γ_2 en B ;
- construire C appartenant à Γ_3 tel que le triangle ABC soit équilatéral.

Comme la rotation de centre A et d'angle -60° applique D sur O (OAD étant un triangle équilatéral de côté r_1), le cercle Γ'_3 a pour image Γ_3 et B appartenant à Γ'_3 a pour image C appartenant à Γ_3 (l'angle ABC ayant une amplitude de 60°). Il s'ensuit que le procédé de construction décrit ci-dessus est correct pour autant que les cercles Γ_2 et Γ'_3 aient au moins

un point d'intersection, donc pour autant que la distance $|OA|$ entre leurs centres soit telle que

$$r_3 - r_2 \leq |OA| \leq r_3 + r_2.$$

Etant donné que $|OA| = r_1$, ces conditions sont remplies vu les hypothèses sur les rayons de cercles.

Nous noterons que, lorsque $r_3 < r_1 + r_2$, les cercles Γ_2 et Γ'_3 sont sécants et ont deux points d'intersection. Le problème admet alors deux solutions, les triangles ABC et $AB'C'$, le point A étant fixé sur Γ_1 . Quand $r_3 = r_1 + r_2$, les cercles Γ_2 et Γ'_3 sont tangents et n'ont qu'un seul point commun; le problème n'a alors qu'une seule solution.

Bonnes solutions de N. BERCKMANS de Waterloo, J. FINOULST de Diepenbeek et J. JANSSEN de Lambermont.

Plus ou moins problème n° 213 de M. et P. n° 119.

La suite (a_n) est définie par

$$\begin{aligned} a_1 &= a_3 = 1 \\ a_2 &= a_4 = -1 \\ a_n &= a_{n-1} \cdot a_{n-2} \cdot a_{n-4}, \quad n \in \{5, 6, 7, \dots\} \end{aligned}$$

Déterminer a_{1999} .

Solution de H.-J. SEIFFERT de Berlin.

Clairement, on a $a_n^2 = 1$ quel que soit $n \in \mathbb{N}^*$.

En utilisant plusieurs fois la relation de récurrence, on obtient, pour $n \in \mathbb{N}^*$,

$$\begin{aligned} a_{n+7} &= a_{n+6} \cdot a_{n+5} \cdot a_{n+3} = (a_{n+5} \cdot a_{n+4} \cdot a_{n+2}) \cdot a_{n+5} \cdot a_{n+3} \\ &= a_{n+5}^2 \cdot a_{n+4} \cdot a_{n+3} \cdot a_{n+2} = a_{n+4} \cdot a_{n+3} \cdot a_{n+2} \\ &= (a_{n+3} \cdot a_{n+2} \cdot a_n) \cdot a_{n+3} \cdot a_{n+2} = a_{n+3}^2 \cdot a_{n+2}^2 \cdot a_n \\ &= a_n. \end{aligned}$$

Puisque $a_1 = a_3 = a_7 = 1$ et $a_2 = a_4 = a_5 = a_6 = -1$, il s'ensuit que pour tout $n \in \mathbb{N}^*$,

$$a_n = \begin{cases} 1 & \text{si } n \equiv 0, 1 \text{ ou } 3 \pmod{7} \\ -1 & \text{si } n \equiv 2, 4, 5, 6 \pmod{7}. \end{cases}$$

En particulier, nous avons $a_{1999} = -1$ car $1999 \equiv 4 \pmod{7}$.

Bonnes solutions de P. DASSY de Liège, J. DILLIES de Le Bizet, J. FINOULST de Diepenbeek, J. JANSSEN de Lambermont, J. RONDOU de Heverlee, C. VAN HOOSTE de Marbaix-la-Tour et C. VILLERS de Hyon.

Les solutions des problèmes suivants doivent me parvenir au plus tard le 1er novembre 1999.

219. De quoi lire

Cette petite boutique est ouverte sept jours sur sept. Durant l'année 1998, on y a vendu au moins un livre chaque jour et au total 600 livres sur l'année complète. Démontrer qu'il y a une période de jours consécutifs pendant laquelle on a vendu exactement 129 livres.

220. Opérations élémentaires

Prouver que la somme de deux nombres premiers impairs consécutifs est le produit d'au moins trois facteurs premiers (non nécessairement distincts).

221. Encore un carré

Les diagonales AC et BD d'un carré se coupent en E . Sur le côté AB du carré et vers l'extérieur, on construit un triangle AFB rectangle en F et tel que $|AF| = 6$ et $|BF| = 8$. Trouver la longueur de EF .

AU SECOURS!

Dans cette rubrique, seront publiés de temps en temps des problèmes pour lesquels nous n'avons pas de solution. Votre aide pour les résoudre est la très bienvenue.

Si vous aussi, vous avez des problèmes sur lesquels vous séchez, envoyez-les, d'autres lecteurs pourront peut-être apporter leur concours.

AS1 C_1 est un cercle fixé, C_2 est un cercle fixé intérieur à C_1 . Par P , point variable sur C_1 , on mène les tangentes t' et t'' à C_2 . Ces droites coupent C_1 en A et B . Démontrer que l'enveloppe de la droite AB est un cercle.

(Problème communiqué par J. OOMS de Chimay)

AS2 Soient a_1, a_2, \dots, a_n des nombres positifs non nuls. Démontrer l'inégalité

$$\frac{a_1}{a_2 + a_3} + \frac{a_2}{a_3 + a_4} + \dots + \frac{a_{n-1}}{a_n + a_1} + \frac{a_n}{a_1 + a_2} > \frac{n}{4}.$$

(Problème proposé en 1969 à une olympiade en URSS)