Les Cahiers de l'Analyse des Données
Vol. XXI - 1996 - n°2 - pp.243-254.

DÉ COMPOSITION DES MESURES ET FONCTIONS SUR UN ENSEMBLE FINI PROBABILISÉ MUNI D'UNE CLASSIFICATION ARBORESCENTE

(I) ARBRES ET HIÉRARCHIES DE PARTIES

[MES. FONCT. ARB. (I)]

Kh. BOUCHKI

N.B. Le présent exposé est extrait d'un mémoire préparé par l'auteur pour obtenir le Doctorat en Statistique. Le mémoire contient des démonstrations complètes. Ici, on se borne à l'enchaînement des définitions et propositions. Ce premier article concerne la seule structure d'ordre; mesures et fonctions seront considérées ensuite.

1 Ensemble muni d'une structure d'arbre

1.1 Définitions et notations pour les ensembles finis ordonnés

Soit A un ensemble fini muni d'une relation d'ordre H_A, ou H, pour laquelle on note: H(a, b): *a* est inférieur à *b*; on convient, comme d'usage, que:

$$\forall\, a \in A \;:\; H(a, a)\,.$$

On note:

Som A : l'ensemble des éléments maximaux, ou sommets de A:

$$\text{Som}\, A = \{\, a \mid a \in A\,;\, \forall\, b \in A : H(a, b) \Rightarrow (a = b)\} \quad ;$$

Ter A : l'ensemble des éléments minimaux, ou terminaux de A:

$$\text{Ter}\, A \;=\; \{\, a \mid a \in A\,;\, \forall\, b \in A : H(b, a) \Rightarrow (a = b)\} \quad ;$$

Is A : l'ensemble des éléments isolés de A, i.e. à la fois minimaux et maximaux: Is A = Ter A $\cap$ Som A :

$$\text{Is}\, A \;= \{\, a \mid a \in A\,;\, \forall\, b \in A : H(b, a) \vee H(a, b) \Rightarrow (a = b)\} \quad ;$$

Nod A : l'ensemble des nœuds, ou éléments de A admettant au moins un successeur distinct de lui-même: Nod A = A - Ter A :

$$\text{Nod}A = \{\, a \mid a \in A\,;\, \exists\, b \in A : H(b, a) \wedge (a \neq b)\} \quad ;$$

Pour tout élément a de A, on note:

Pred(a, A), ou Pred a, si aucune confusion n'est à craindre: l'ensemble des prédécesseurs de *a*, y compris *a*, dans A:

$$\mathrm{Pred}(a, A) \;=\; \{\, b \mid b \in A \;;\; H(a, b) \,\} \;;$$

Su(a, A), ou Su a : l'ensemble des *b* successeurs de *a* dans A et distincts de *a*:

$$\mathrm{Su}(a, A) \;=\; \{\, b \mid b \in A \;;\; b \neq a \;;\; H(b, a) \,\} \;;$$

(On notera que *a* est son propre prédécesseur, mais non son successeur.)

Sui(a, A), ou Sui a : l'ensemble des successeurs immédiats de *a* dans A; c'est-à-dire, l'ensemble des successeurs *b* de *a* tels qu'il n'y ait pas d'autre élément de A entre *b* et *a* :

$$\mathrm{Sui}(a, A) \;=\; \{ b \mid b \in A \;;\; b \neq a \;;\; \{u \mid u \in A \;;\; H(b, u) \;;\; H(u, a)\} = \{a, b\} \,\} \;;$$

On peut encore dire que Sui a est l'ensemble des éléments maximaux de Su a; il est non vide si et seulement si *a* est un nœud : a ∈ Nod A.

1.2 Définition et axiomes de la structure d'arbre hiérarchique

L'ensemble A, muni de la structure d'ordre H est un *arbre hiérarchique*, (en bref, ici, *arbre*), si et seulement si est satisfait l'axiome Ha :

Ha : ∀ a ∈ A, l'ensemble Pred a, muni de l'ordre induit par H, est totalement ordonné;

autrement dit, A est un arbre si et seulement si:

$$\forall\, a \in A,\, \forall\, x, y \in \mathrm{Pred}\, a \;:\; H(x, y) \,\vee\, H(y, x) \;;$$

Un arbre est dit *connexe* si et seulement si il satisfait à l'axiome Hc:

Hc : Card(Som A) = 1 ;

autrement dit, A a un unique sommet; celui-ci est souvent noté: som A.

Un arbre est dit *binaire* si et seulement si il satisfait à l'axiome Hb:

Hb : ∀ a ∈ Nod A : Card (Sui a) = 2 ;

autrement dit, un nœud de A a exactement deux successeurs immédiats.

Un arbre est dit *fourchu* si et seulement si il satisfait à l'axiome Hf:

Hf : ∀ a ∈ Nod A : Card (Sui a) ≥ 2 ;

autrement dit, un nœud de A a au moins deux successeurs immédiats.

Nous donnerons, sans démonstration, un scholie; où est introduite une notation usuelle.

Commentaire rappel

Scholie : Soit A un arbre, pour une structure d'ordre H:

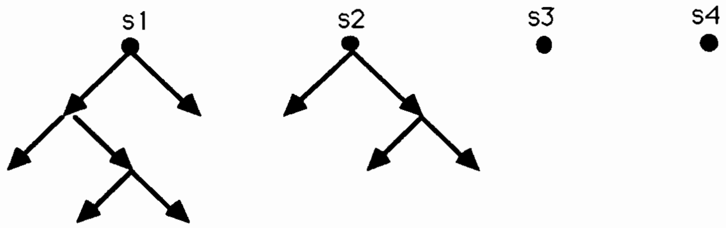
1) $\forall$ a $\in$ A - Som A, il existe un élément, unique, noté p(a), appelé le père de a, et tel que: a $\in$ Sui(p(a), A).

2) $\forall$ a $\in$ A, il existe un sommet unique, s $\in$ Som A, tel que H(a, s).

3) $\forall$ s $\in$ Som A, il existe t $\in$ Ter A, tel que H(t, s) .

1.3 Représentation graphique d'un arbre et langage associé

Soit A un ensemble fini, muni d'une structure d'arbre par la relation d'ordre H. Pour construire une représentation graphique de A, on peut représenter les sommets de A par des points convenablement espacés sur une ligne horizontale. Si un sommet *s* n'est pas point isolé (s $\in$ Som A; s$\notin$ Is A), on représente l'ensemble Sui s, des succeseurs immédiats de s, par des points liés à s par des flèches descendantes; et ainsi de suite, jusqu'à épuisement de l'ensemble A; les suites de flèches descendantes aboutissent toutes à un élément terminal. Pour la structure d'arbre telle qu'elle est considérée ici, l'orientation des flèches à droite ou à gauche est indifférente.



On dira encore, pour aîder à l'intuition, qu'on peut considérer le sommet unique d'un arbre connexe, comme une souche, ou un tronc; de chaque nœud partent autant de branches qu'il a de successeurs immédiats; un arbre non-connexe est une sorte de forêt. Mais il faut prendre garde que l'ordre H est défini de telle sorte que les branches sont inférieures au tronc; en ce sens, dans la notation mathématique, l'arbre est renversé.

On appelle *chemin descendant* de l'arbre A, une suite $\{a_1, a_2, \ldots, a_n\}$ d'éléments de A, telle que : $\forall$ i $\in$ $\{1, 2, \ldots, (n-1)\}$: $a_{i+1} \in$ Sui (a_i) ;

de même, on appelle *chemin ascendant* de l'arbre A, une suite $\{a_1, a_2, \ldots, a_n\}$ d'éléments de A, telle que : $\forall$ i $\in$ $\{1, 2, \ldots, (n-1)\}$: $a_i \in$ Sui (a_{i+1}) .

Lemme : Soit A un arbre *binaire*, connexe ou non. Notons:

S : nombre des sommets non isolés de A: S = Card (Som A - Is A) ;
N : nombre des nœuds de A qui ne sont pas sommet ;
T : nombre des éléments terminaux non isolés : T = Card (Ter A - Is A) ;

alors, on a la relation : N + 2.S = T ; en particulier, si l'arbre est *connexe*, S = 1; et l'on a: N = T - 2.

2 Relations entre arbres et au sein d'un arbre A

2.1 Sous-arbres

Une partie A' d'un arbre A, munie de la structure d'ordre induite par celle de A, satisfait nécessairement à l'axiome Ha (cf. §1.2): c'est donc un arbre. On dit que A' est un *sous-arbre* de A; et que A est un *sur-arbre* de A'. On dit que A et A' sont *coterminaux*, si Som(A) = Som(A') et Ter(A) = Ter(A').

On dit que A' est un sous-arbre *plein*, ou *dense* de A, si A' satisfait l'axiome Hd:

$$\mathbf{Hd} : \forall\, a, b \in A' : \quad A' \supset \{u \mid u \in A; H(a, u) ; H(u, b)\} \ ;$$

autrement dit: si A' contient a et b, il contient tout u, intermédiaire, dans A, entre a et b.

Un sous-arbre A' est dit entier, relativement à A, s'il satisfait l'axiome He:

$$\mathbf{He} : \forall\, a \in \mathrm{Nod}(A'),\ \forall\, b \in A : \ H(b, a) \ \Rightarrow \ \ldots$$

$$\exists\, u \in A' : \ (u \neq a) \ \wedge \ H(u, a) \ \wedge \ (\, H(b, u) \ \vee \ H(u, b)\,) \ \ ;$$

en bref: si b $\notin$ A', on trouve, dans A', dans la direction de *b*, un *u*, soit au-dessus, soit au-dessous de *b*.

2.2 Sections et relations entre sections pour un ensemble ordonné

Soit E un ensemble muni d'une relation d'ordre H. On appelle *section* de E une partie X, de E, telle que deux éléments distincts quelconques de X ne soient pas comparables pour la relation H :

$$\forall\, x, x' \in X \ : \ H(x, x') \ \Rightarrow x = x' \ .$$

Cette définition s'applique au cas où E est un arbre fini, A. En particulier, pour tout arbre A, Som A et Ter A sont des sections.

Une section X d'un arbre A est dite *totale* si :

$$\forall\, t \in \mathrm{Ter}\, A,\ \exists\, x \in X \ : \ H(t, x) \ ;$$

autrement dit si, $\forall$ t $\in$ Ter A, la suite Pred(t, A), des prédécesseurs de t dans A, rencontre X en un point x; lequel est unique, de par la définition même des sections.

Plus généralement, dans un ensemble E muni d'une relation d'ordre H, on dit qu'une section X *recouvre* une section Y si sont satisfaites les conditions:

$$\forall\, x \in X,\ \exists\, y \in Y \ : \ H(y, x) \ ;$$

$$\forall\, y \in Y,\ \exists\, x \in X \ : \ H(y, x) \ \wedge \ (\forall\, x' \in X : H(y, x') \Rightarrow x = x' \ ) \ ;$$

en bref, au-dessous de tout *x*, il y a au moins un *y*; au-dessus de tout *y* il y a un seul *x*.

On définit, sur l'ensemble Sec(E) des parties finies de E qui sont des sections, la relation HZ(Y, X) :

$$\forall X, Y \in \mathrm{Sec}(E) : \mathrm{HZ}(Y, X) \Leftrightarrow X \text{ recouvre } Y \;;$$

On démontre que HZ est une relation d'ordre sur Sec(E). Dans un arbre A, on a la relation HZ(Ter A, Som A). Il importe de noter ici que la relation HZ(Y, X) ne dépend pas de la structure de E tout entier, mais seulement de l'ordre induit sur $X \cup Y$; ce que précise l'énoncé qui suit.

Soit E, un ensemble muni d'une relation d'ordre H; X et Y, deux sections de E. La relation HZ(Y, X) est satisfaite si et seulement si $X \cup Y$, muni de la relation d'ordre induite par H, est un arbre pour lequel X et Y sont, respectivement, l'ensemble des sommets et l'ensemble des terminaux:

$$\mathrm{Som}(X \cup Y) = X \;;\; \mathrm{Ter}(X \cup Y) = Y \quad .$$

2.3 Relations entre sections d'un arbre

Pour un arbre A, on définit, entre sections, une relation HZZ, plus stricte que HZ. De façon précise: soit X un section de A, notons Ter A\X (ce qu'on lira: Ter de A sous X) l'ensemble des éléments terminaux de A qui sont successeurs d'un élément de X:

$$\mathrm{Ter}\, A\backslash X = \{t \mid t \in \mathrm{Ter}\, A \;;\; \exists\, x \in X \;:\; H(t, x)\} \quad .$$

Ceci posé, soit X et Y deux sections de A: on dira que X *repose* sur Y, si X recouvre Y et Y recouvre Ter A\X :

$$\mathrm{HZZ}(Y, X) \quad \Leftrightarrow \quad \mathrm{HZ}(Y, X) \;\wedge\; \mathrm{HZ}(\mathrm{Ter}\, A\backslash X, Y) \;.$$

On a l'implication :

$$\mathrm{HZZ}(Y, X) \quad \Rightarrow \quad \mathrm{Ter}\, A\backslash X = \mathrm{Ter}\, A\backslash Y \;.$$

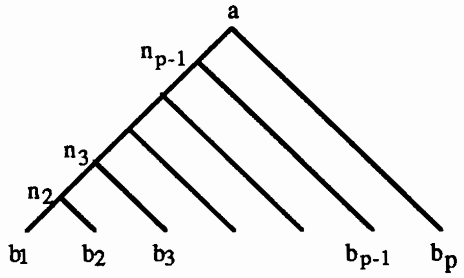
On démontre que la relation HZZ et une relation d'ordre sur Sec(A). À la différence de HZ, la relation HZZ, dépend essentiellemnt de A tout entier. On doit donc noter, si une confusion est à craindre, $\mathrm{HZZ}_A(Y, X)$.

Une partie A' de A est un sous-arbre *entier* de A (cf. *supra*, axiome He) si et seulement si on a: $\mathrm{HZZ}_A(\mathrm{Ter}\, A', \mathrm{Som}\, A')$. En particulier, si X et Y sont deux sections de A telles qu'on ait $\mathrm{HZZ}_A(Y, X)$, $X \cup Y$ est un sous-arbre entier de A (avec, cf. *supra*, $\mathrm{Som}(X \cup Y) = X$; $\mathrm{Ter}(X \cup Y) = Y$).

2. 4 Sur-arbres binaires associés à un arbre fourchu

Notre objet ultime est d'associer, à un ensemble muni d'une loi de probabilité et d'une classification hiérarchique, une base canonique orthonormée de fonctions sur cet ensemble. La base n'est complètement déterminée que si la hiérarchie est binaire: d'où le présent §.

La construction commence par le cas d'un sous-arbre ***élémentaire***. Soit $a \in$ Nod A: le sous-arbre élémentaire associé à a n'est autre que l'ensemble $\{a\} \cup$ Sui(a), muni de la structure induite par A. Pour cet arbre, a est l'unique sommet; et Sui(a) est l'ensemble des terminaux (de cardinal ≥ 2, puisque l'arbre est fourchu). Nous donnons sur une figure une construction de sur-arbre binaire ß(a), associé à $\{a\} \cup$ Sui(a); avec, pour ß(a), Som(ß(a) = $\{a\}$; Ter(ß(a) = Sui(a). La construction proposée ici dépend du choix d'un ordre total sur Sui(a); et, en général, tout parenthésage sur Sui(a) fournit une telle structure.



On pose Card(Sui(a)) = $p > 2$. On note Sui(a) = $\{b_1, b_2, \ldots, b_p\}$; alors, ß(a) s'écrit: $\{a\} \cup$ Sui(a) $\cup \{n_2, \ldots, n_{p-1}\}$; avec, sur ß(a), les relations d'ordre:

$$H(b_1, n_2) ; H(b_p, a) ; H(n_{p-1}, a) ;$$

$$H(b_2, n_2) ; H(b_3, n_3) ; \ldots; H(b_{p-1}, n_{p-1}) ;$$

$$H(n_2, n_3) ; \ldots; H(n_{p-2}, n_{p-1}) ;$$

et toutes celles qui s'en déduisent par transitivité. Il est clair que la structure de ß(a) dépend de l'ordre choisi sur Sui(a); un ordre et son inverse donnant, toutefois la même structure. Le cas où Card(Sui a) = 2 est trivial.

Plus généralement, pour un arbre fourchu A, on définit un sur-arbre ß(A) en appliquant la construction ci-dessus à chaque nœud de A; soit, en formule:

$$ß(A) = (\cup \{ß(a) \mid a \in \text{Nod } A\}) \cup \text{Is } A ;$$

formule qui n'est pas univoque, dans la mesure où elle ne l'est pas pour ß(a) si Sui(a) compte plus de deux éléments.

3 Opérations sur les arbres et sous-arbres : catégorie

Dans la suite, afin d'associer à un ensemble probabilisé muni d'une classification hiérarchique une décomposition des mesures et fonctions, on considérera un système d'opérateurs linéaires correspondant aux sous-arbres d'un arbre; avec des opérations que l'on peut, formellement, définir sur ceux-ci, ainsi qu'on le fera dans le présent §, en reprenant des notions de la théorie des catégories.

3.1 Somme directe d'arbres et de sous-arbres

Soit $\{(A_p, H_p) \mid p \in P\}$, une famille finie, indicée par p, d'ensembles A_p, deux à deux d'intersection vide, et dont chacun est muni d'une structure d'arbre H_p. Notons $A = \cup \{A_p \mid p \in P\}$, la réunion des A_p, et H, la structure d'ordre la moins fine sur A, induisant, sur chacun des A_p, la structure H_p ; Soit :

$$\forall x, x' \in A : H(x, x') \Leftrightarrow \exists p \in P : x, x' \in A_p \wedge H_p(x, x') ;$$

alors, H définit sur A une structure d'arbre, appelée *somme directe* des structures (A_p, H_p); et l'on note:

$$(A, H) = \oplus \{(A_p, H_p) \mid p \in P\} \ ; \quad \text{ou : } A = \oplus \{A_p \mid p \in P\} \ ;$$

[La somme directe peut se construire, en toute généralité, même si les A_p ne sont pas d'intersection vide: il faut alors prendre pour A une somme directe des ensembles A_p; i.e., en bref, distinguer les éléments qui se rencontrent dans plusieurs des A_p.].

Comme la réunion $\cup$, la somme directe $\oplus$ est associative et commutative.

Soit A un arbre, on associe, à tout $a \in A$, sa *branche*, Br(a, A) définie par: $Br(a, A) = \{x \mid x \in A ; H(x, a)\} = \{a\} \cup Su(a, A)$. Il est clair que Br(a, A) est un sous-arbre de A. Et on a la formule de somme directe:

$$A = \oplus \{Br(a, A) \mid a \in Som\, A\} \ ;$$

formule où Br(a, A) peut être appelée la *composantes connexe* du sommet *a*.

Plus généralement, soit $\{B_p \mid p \in P\}$ une famille finie de sous-arbres d'un même arbre (A, H). La réunion $B = \cup \{B_p \mid p \in P\}$, considérée avec la structure induite par H, s'dentifie à la somme directe $\oplus \{B_p \mid p \in P\}$, à condition que deux éléments {x, x'} appartenant à des B_p distincts entre eux soient non comparables pour la relation d'ordre H:

$$\forall x, x' \in B : H(x, x') \Rightarrow \exists p \in P : x, x' \in B_p \quad ;$$

on dit alors que les B_p sont deux à deux *étrangers* pour la relation H; ce qui équivaut encore à poser que les Ter A\Som$B_p$ sont deux à deux d'intersection vide (cf. §2.3 pour la notation Ter A\X, qu'on lit: Ter de A sous X).

3.2 Composition entre sous-arbres d'un arbre A

Soit A, un arbre; notons arb(A), l'ensemble des parties de A considérée chacune comme un sous-arbre, avec la structure d'arbre induite par A. Soit deux sous-arbres A1 et A2 tels que Ter A1 = Som A2; on dit que A1 et A2 sont *composables* (dans l'ordre donné); et l'on note A1 ∘ A2 (avec le même signe que pour une composition d'opérateurs), le sous-arbre A1 ∪ A2 de A.

L'on a:

$$\mathrm{Som}(A1 \circ A2) = \mathrm{Som}(A1) \quad ; \quad \mathrm{Ter}(A1 \circ A2) = \mathrm{Ter}(A2) \ .$$

Il est clair que la composition ainsi définie est associative; en ce sens que si A1 et A2 sont composables, ainsi que A2 et A3, on a:

$$((A1 \circ A2) \circ A3) = (A1 \circ (A2 \circ A3)) = A1 \circ A2 \circ A3 \ ;$$

la dernière écriture, sans parenthèse, étant justifiée par l'associativité.

On sait que, pour tout sous-arbre B de A, Som(B) et Ter(B) sont des sections de A. Toute section de A constitue un sous-arbre particulier. Pour la composition, on a:

$$B = \mathrm{Som}(B) \circ B = B \circ \mathrm{Ter}(B) = \mathrm{Som}(B) \circ B \circ \mathrm{Ter}(B) \ ;$$

il apparaît donc que, pour la composition, toute section X est une *unité à gauche* pour les sous-arbres B dont elle est l'ensemble des sommets ($X = \mathrm{Som}\, B$); et *unité à droite* pour les sous-arbres dont elle est l'ensemble des terminaux ($X = \mathrm{Ter}\, B$).

3.3 La catégorie des sous-arbres entiers d'un arbre donné

Dans la suite, on se servira plus particulièrement des sous-arbres entiers, d'un arbre A donné (associé à une classification hiérarchique). Notons arbe(A) l'ensemble de ces sous-arbres. On va définir sur arbe(A) une structure de catégorie. Le lecteur non habitué à cette notion pourra se borner à accepter telles quelles les définitions générales reprises ici pour arbe(A).

La classe des *objets* de arbe(A) est l'ensemble, Sec(A), des sections de A.

La classe des *flèches*, ou *morphismes* n'est autre que arbe(A); i.e. (§2.3):

$$\mathrm{arbe}(A) = \{B \mid A \supset B;\ \mathrm{HZZ}_A(\mathrm{Ter}\, B, \mathrm{Som}\, B)\} \ ;$$

pour $B \in \mathrm{arbe}(A)$, Som B et Ter B sont appelés respectivement *but* et *source* de B: $\mathrm{source}(B) = \mathrm{Ter}\, B \ ; \ \mathrm{but}(B) = \mathrm{Som}\, B$.

Deux morphismes B et C sont composables (cf. §3.2) si et seulement si la source de B (placé à gauche) coïncide avec le but de C (placé à droite):

$$\forall B, C \in \mathrm{arbe}(A) : \ \mathrm{Ter}\, B = \mathrm{Som}\, C \Rightarrow B \circ C = B \cup C \in \mathrm{arbe}(A) \ ;$$

que $B \circ C$ soit un sous-arbre entier de A résulte de la transitivité de HZZ.

Soit X et Y deux sections de A telles qu'on ait $\mathrm{HZZ}_A(X, Y)$; on note Hom(X, Y) l'ensemble des morphismes dont la source est X et le but, Y:

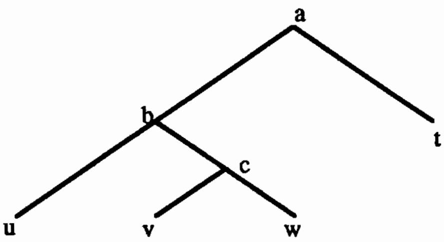
$$\mathrm{Hom}(X, Y) = \{B \mid A \supset B \ ; \ \mathrm{Ter}\, B = X \ ; \ \mathrm{Som}\, B = Y\} \ ;$$

en particulier: $\mathrm{HZZ}(X, Y) \Rightarrow X \circ Y = X \cup Y \in \mathrm{Hom}(X, Y)$.

On peut identifier les notions d'objet, de section et de morphisme unité.

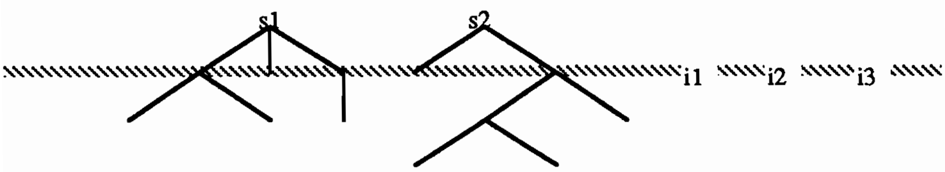
La catégorie arbe(A) est stable pour la somme directe: si les B_p sont des sous-arbres entiers et que $B = \oplus \{B_p \mid p \in P\}$ est défini, B est un sous-arbre entier. En effet, Ter A\SomB n'est autre que la réunion des Ter A\Som$B_p$; et puisque chacun des Ter B_p recouvre Ter A\Som$B_p$, la réunion de ceux-là, Ter B, recouvre la réunion de ceux-ci, qui n'est autre que Ter A\SomB

3.4 Expression d'un arbre en fonction de ses sous-arbres élémentaires



$$A = A(a) \circ (A(b) \oplus \{t\}) \circ (\{u\} \oplus A(c) \oplus \{t\}) \quad ;$$
$$A = A(a) \circ ((A(b) \circ (\{u\} \oplus A(c))) \oplus \{t\}) \quad ;$$

Partons de l'exemple ci-dessus. L'arbre A est connexe, de sommet *a* ; il a trois nœuds {a, b, c} et trois terminaux {t, u, v, w}. Pour tout nœud *n*, on note A(n) le sous-arbre élémentaire $\{n\} \cup$ Sui(n). Sont proposées deux expressions de l'arbre A, où figurent des sous-arbres élémentaires, des arbres réduits à un élément, et les signes de composition et de somme directe.



Notre propos est de montrer qu'une expression de cette forme existe pour tout arbre. Afin de procéder par récurrence sur le cardinal de l'arbre, on considère un énoncé plus général: étant donné un arbre A, tout sous-arbre plein, B, de A s'exprime par composition et somme directe à partir des sous-arbres élémentaires de A et d'éléments isolés. La proposition est évidente si B = Is B, ne comporte que des éléments isolés. Sinon, avec les notations ci-après, on écrit B sous la forme: $B1 \circ B2$; avec B1 décomposé en somme directe, et CardB2 < CardB :

$$V = \mathrm{Som}\,B - \mathrm{Is}\,B \quad ; \quad T1 = \cup\{\mathrm{Sui}(s) \mid s \in V\} \; ;$$
$$B1 = \mathrm{Som}\,B \cup T1 \quad ; \quad B2 = B - V \; ;$$
$$B1 = (\oplus\{A(s) \mid s \in V\}) \oplus (\oplus\{\{i\} \mid i \in \mathrm{Is}\}) \quad ; \quad B = B1 \circ B2 \quad ;$$

Sur la figure, B1 comprend ce qui est sur la ligne hachurée au au-dessus de celle-ci; B2, ce qui est sur la ligne et au-dessous.

4 Classifications hiérarchiques sur un ensemble fini

4.1 Rappel de la définition d'une classification hiérarchique

Soit I un ensemble fini ; on note P(I) l'ensemble de ses parties non vides:

$$P(I) = \mathfrak{P}(I) - \{\emptyset\} = \{X \mid I \supset X ; X \neq \emptyset\} .$$

L'inclusion définit sur P(I) une relation d'ordre H : $H(X, Y) \Leftrightarrow Y \supset X$.

Nous appellerons *classification hiérarchique* ou *classification* dans I, une partie A de P(I) qui, munie de la relation H, satisfait aux axiomes suivants

Axiome d'*intersection* : deux éléments de A, non comparables pour l'inclusion, sont d'intersection vide; ce qu'on peut écrire:

$$\mathbf{Hi} : \forall x, y \in A \ : \ x \cap y \in \{x, y, \emptyset\} \ ;$$

Axiome de *réunion* : tout nœud de A est réunion de ses successeurs :

$$\mathbf{Hu} : \forall n \in \text{Nod } A \ : \ n = \cup\{x \mid x \in \text{Su}(n)\} \ ;$$

plus précisément, il résulte de Hi que Sui(n), ensemble des successeurs immédiats de *n*, est une partition de *n* (en classes disjointes).

NB: Il importe de noter que l'on utilise communément, dans la pratique, des classifications *non-hiérarchiques*, ne satifaisant pas à Hi.

Si A satisfait aux axiomes Hi et Hu, A, muni de H, est un arbre; et A est fourchu; car si un nœud *n* n'avait qu'un seul successeur immédiat, celui-ci devrait coïncider avec *n*..

On appelle support de A, et on note Sup A, la réunion des éléments de A, considérés comme des parties de I :

$$\text{Sup } A = \cup\{x \mid x \in A\} = \cup\{x \mid x \in \text{Som } A\} = \cup\{x \mid x \in \text{Ter } A\} \ ;$$

plus généralement, une section X de A, telle que HZZ(X, Som A), définit, comme Som A et Ter A, une partition de Sup A. On dit que A est une hiérarchie *sur* Sup A et *dans* I. Dans la suite, on notera respectivement Clsf(I) et Prtt(I) l'ensemble des classifications et celui des partitions dans I.

Une partie A' de A est une classification dans I, si et seulement si A' est un sous-arbre entier de A. On dit alors que A' est une *sous-classification* de A. Si A' est une section de A, A' est une *partition* de son support Sup A'.

Une classification, A, dans I, est dite *fine* si tout $t \in$ Ter A est une partie de I réduite à un élément. Une classification A est dite *totale* si elle est fine et a pour unique sommet I ; alors, Som A = $\{I\}$; Ter A = $\{\{i\} \mid i \in I\}$.

Soit A un arbre fourchu: A définit sur Ter A, une classification arborescente fine $A^{Ter} = \{a^{Ter} \mid a \in A\}$; où : $a^{Ter} = \{t \mid t \in \mathrm{Ter}\, A; H(t, a)\}$.

Plus généralement, étant donné un ensemble I, et une surjection f d'une partie I', de I, sur Ter A, on a une classification A^I, (dans I, sur I') en posant:

$$A^I = \{a^I \mid a \in A\} \quad ; \quad \forall a \in A : a^I = \{i \mid i \in I' ; H(f(i), a)\} \ ;$$

4.2 Catégorie des classifications dans un ensemble

Reprenons maintenant les considérations du §3.3. Soit I un ensemble fini. On définit une catégorie dont l'ensemble des morphismes est l'ensemble, Clsf(I), des classifications dans I, muni des deux opérations $\{ \circ \ ; \ \oplus \}$.

Le sous-ensemble Prtt(I) des partitions, s'identifie aux objets (ou aux unités): la source de la classification A est Ter(A), son but est Som(A). Deux morphismes B et C sont composables dans l'ordre $B \circ C$ si et seulement si la source de B coïncide avec le but de C; on alors:

$$\forall B, C \in \mathrm{Clsf}(I) : \ \mathrm{Ter}\, B = \mathrm{Som}\, C \ \Rightarrow B \circ C = B \cup C \in \mathrm{Clsf}(I) \ ;$$

Soit X et Y deux objets (ou morphismes unité), i.e. deux partitions dans I. L'ensemble Hom(X, Y) est non-vide si et seulement si X et Y ont même support et que X est plus fine que Y. On a alors, en particulier:

$$(X \cup Y) \in \mathrm{Hom}(X, Y).$$

La somme directe de deux objets ou de deux morphismes B et C, (les objets n'étant que des cas particuliers des morphismes), est définie si et seulement si $\mathrm{Sup}(B) \cap \mathrm{Sup}(C) = \emptyset$; on a alors $B \oplus C = B \cup C$.

On peut écrire ici des formules analogues à celles proposées au §3.4 pour exprimer un arbre en fonction de ses sous-abres élémentaires. Soit X et Y deux partitions dans I, avec Sup(X) = Sup(Y) = s ; et X plus fine que Y. On a sur s, 4 partitions; qui sont, de la plus grossière à la plus fine:

$$\{s\} \ ; \quad Y \quad ; \quad X \quad ; \quad \{\{i\} \mid i \in s\} \ ;$$

à partir de ces partitions, on construit sur I une classification totale A, i.e. fine (distinguant tous les éléments) et de sommet I :

$$A = \{s, I\text{-}s, I\} \circ (B \oplus C) \ \in \ \mathrm{Hom}(\{\{i\} \mid i \in I\}, \{I\}) \ ;$$

où on a noté :

$$B = (Y \cup \{s\}) \circ (Y \cup X) \circ (X \cup \{\{i\} \mid i \in s\}) \in \mathrm{Hom}(\{\{i\} \mid i \in s\}, \{s\}) \ ;$$

$$C = (\{I\text{-}s\} \cup \{\{i\} \mid i \in I\text{-}s\}) \ \in \ \mathrm{Hom}(\{\{i\} \mid i \in I\text{-}s\}, \{I\text{-}s\}) \ ;$$

Toutefois, $\{I\} \cup Y \cup X \cup \{\{i\} \mid i \in I\}$ est également une classification fine et totale sur I; et c'est évidemment la plus simple de celles où rentrent X et Y.

4.3 Introduction de la partie vide dans une hiérarchie

Dans la suite, pour la décomposition des fonctions sur I, il sera commode de faire se correspondre la fonction constante et la partie vide, $\emptyset$. À cette fin, étant donné une classification, A, sur I, on notera $A°$ l'ensemble $A \cup \{\emptyset\}$, $A°$ étant muni d'une relation d'ordre, H, définie par l'inclusion; à ceci près que $\emptyset$ sera considéré comme supérieur à tous les *a* de A : $\forall\ a \in A : H(a, \emptyset)$; bien que, au contraire, $\emptyset$ soit inclus dans tout *a*. Ainsi, on a pour successeurs immédiats de $\emptyset$ dans $A°$ les sommets de A: $\mathrm{Sui}(\emptyset, A°) = \mathrm{Som}\ A$.

Soit $B°$ un sous-arbre de $A°$, avec $\emptyset \in B°$, som $B° = \emptyset$: si $B°$ n'est pas réduit à $\{\emptyset\}$, et que $B°$ est entier (cf. §2.1, axiome He), on a Sup B = Sup A = I. De même, on convient que $\{\emptyset\}$ est la partition la plus grossière sur I. On notera :

$$\mathrm{Prtt}°(I) = \mathrm{Prtt}(I) \cup \{\{\emptyset\}\} \quad ;$$

$$\mathrm{Clsf}°(I) = \mathrm{Clsf}(I) \cup \{B° \mid B \in \mathrm{Clsf}(I)\ ;\ \mathrm{Sup}\ B = I\} \cup \{\{\emptyset\}\} \quad ;$$

NB: avec cette convention, *on n'a pas* $B° \in \mathrm{Clsf}°(I)$, si $\mathrm{Sup}(B) \neq I$. La notation $B°$ sera cependant parfois employée même si Sup $B \neq I$.

En convenant de placer $\{\emptyset\}$ au-dessus de tout élément de Prtt(I) ayant pour support I, (bien que $\emptyset$ ne figure dans aucune partition,) on a adjoint à la catégorie des partitions dans I un nouvel objet $\{\emptyset\}$; ayant, de plus, attribué I comme support à $\{\emptyset\}$, on peut convenir que $\mathrm{Hom}(X, \{\emptyset\})$ est non vide si et seulement si Sup X = I, auquel cas $\mathrm{Hom}(X, \{\emptyset\})$ s'identifie à l'ensemble des classifications de source X, complétée chacune par le sommet $\emptyset$:

$$\mathrm{Hom}(X, \{\emptyset\}) = \{B° \mid B \in \mathrm{Clsf}(I)\ ;\ \mathrm{Sup}\ B = I\ ;\ \mathrm{Ter}\ B = X\} \ .$$

On a bien ainsi une catégorie ayant pour ensemble de ses objets Prtt°(I); et pour ensemble de ses morphismes Clsf°(I).