LOIS DISTRIBUTIVES. APPLICATIONS AUX AUTOMATES STOCHASTIQUES

ELISABETH BURRONI

RÉSUMÉ. Les automates déterministes sont les algèbres d'une monade associée à un monoïde libre. Pour prolonger aux automates non déterministes et stochastiques un tel formalisme monadique, il est commode d'avoir recours à une notion plus riche que celle de monade, mais tout aussi basique : celle de loi distributive entre deux monades. La notion d'algèbre sur une monade se généralise alors par celle d'algèbre sur une loi distributive (voir [Bu' 74]). Les automates non deterministes et les automates stochastiques sont des algèbres sur des lois distributives dont la première monade est, comme dans le cas déterministe, la monade associée à un monoïde. Nous avons introduit le cas non déterministe à l'aide de la monade des parties dans [Bu' 74]. La monade des probabilités (voir [Tull 71] et [Giry 81]) nous permet alors de définir de manière analogue une nouvelle loi distributive et de présenter les automates stochastiques comme les algèbres sur cette loi distributive.

Cet article se situant au confluent des catégories, des automates et des probabilités, nous avons, pour la commodité du lecteur, fait les rappels utiles relatifs à chacune de ces disciplines (App.I, App.II et App.III). On trouvera également en appendice (App.IV et App.V) une construction détaillée de la monade des probabilités, et de la loi distributive qui la relie à la monade associée à un monoïde.

1. Automates déterministes

Soit $M = (M, e, m)$ un monoïde (e est l'unité de M et m, sa loi, sera notée multiplicativement : $(a,b) \mapsto ab$, pour tout $(a,b) \in M \times M$). On lui associe une monade $T_M = (T_M, \eta, \mu)$ sur la catégorie **Ens** des ensembles, définie de la manière suivante : l'endofoncteur T_M : **Ens** $\longrightarrow$ **Ens** associe, à tout ensemble X, l'ensemble $M \times X$ et, à toute application $f : X \longrightarrow Y$, l'application $M \times f : M \times X \longrightarrow M \times Y : (a,x) \mapsto (a, f(x))$; de plus, pour tout ensemble X, $\eta X : X \longrightarrow M \times X$ et $\mu X : M \times M \times X \longrightarrow M \times X$ sont les applications respectivement définies par $\eta X(x) = (e,x)$ et $\mu X(a,b,x) = (ab,x)$ (voir App.I.4, où, plus généralement, on associe une monade à tout monoïde d'une catégorie monoïdale).

Une T_M-*algèbre* est un couple (X, θ), où X est un ensemble et $\theta : M \times X \longrightarrow X$ une application vérifiant les axiomes (A_1) et (A_2) des algèbres (voir App.I.3), i.e vérifiant, pour tout $a, b \in M$ et tout $x \in X$,

$$\theta(e, x) = x \quad \text{et} \quad \theta(ab, x) = \theta(a, \theta(b, x))$$

Une T_M-algèbre s'appelle plus couramment un *M-ensemble*, mais, pour des raisons d'homogénéïté de terminologie avec la suite, nous lui donnerons également le nom de *M-automate déterministe* (de *loi de transition* θ). En se reportant à App.II.1, on constate trivialement le fait suivant :

Proposition 1 : *Dans le cas où $M = \Sigma^*$ (c'est-à-dire, le cas où M est le monoïde libre engendré par un ensemble fini Σ), une T_M-algèbre est la donnée d'un automate déterministe sur l'alphabet Σ.*

Pour décrire, en restant dans un cadre «monadique», les automates non déterministes et stochastiques, on a besoin d'une notion plus riche que celle de monade, celle de «loi distributive» entre deux monades (voir App.I.5).

2. Algèbres sur une loi distributive

Les automates non déterministes et stochastiques seront présentés plus loin comme deux cas particuliers d'une notion d'algèbre sur une loi distributive. Pour une introduction aux lois distributives, voir App.I.5 ; dans cette section nous introduisons ce nouveau type d'algèbre. Les données dans cette section sont les suivantes : $T = (T, \eta, \mu)$ et $T' = (T', \eta', \mu')$ sont deux monades sur une même catégorie $\mathbf{C}$, et $\tau : T/T'$ une loi distributive de T sur T', i.e c'est la donnée d'un triplet $\tau = (\tau, T, T')$, où $\tau : TT' \longrightarrow T'T$ est une transformation naturelle rendant commutatifs les diagrammes (B_i) rappelés dans App.I.5. Tous les résultats donnés ci-dessous ont été démontrés dans [Bu' 74].

2.1. On appelle *τ-algèbre* (appelée D-algèbre dans [Bu' 74]), la donnée d'un couple (X, θ), où X est un objet de $\mathbf{C}$ et $\theta : TX \longrightarrow T'X$ un morphisme de $\mathbf{C}$ rendant commutatifs les diagrammes suivants dans $\mathbf{C}$:

(A'_1)
$$\begin{tikzcd}
TX \arrow[rr, "\theta"] & & T'X \\
& X \arrow[ul, "\eta X"] \arrow[ur, "\eta' X"'] &
\end{tikzcd}$$

(A'_2)
$$\begin{tikzcd}
TTX \arrow[r, "T\theta"] \arrow[dr, "\mu X"'] & TT'X \arrow[r, "\tau X"] & T'TX \arrow[r, "T'\theta"] & T'T'X \arrow[dl, "\mu' X"] \\
& TX \arrow[r, "\theta"'] & T'X &
\end{tikzcd}$$

On retrouve les T-algèbres, comme cas particulier des τ-algèbres, en prenant $T' = id_{\mathbf{C}}$, $\eta' = \mu' = id_{T'}$ et $\tau = id_T$.

2.2. La donnée d'une τ-algèbre sur X est équivalente à la donnée d'une T-algèbre sur $T'X$ qui est τ-compatible avec la T'-algèbre libre $\mu'X$ (voir App.I.6). Plus précisément, si (X,θ) est une τ-algèbre, $(T'X, \mu'X \circ T'\theta \circ \tau X)$ est une T-algèbre qui est τ-compatible avec $(T'X, \mu'X)$; et inversement, si $(T'X, \theta)$ est une T-algèbre τ-compatible avec $(T'X, \mu'X)$, alors $(X, \theta \circ T\eta' X)$ est une τ-algèbre.

2.3. La donnée de la loi distributive $\tau : T/T'$ équivaut à la donnée d'un prolongement de la monade T sur $\mathbf{C}$ en une monade $\widetilde{T} = (\widetilde{T}, \widetilde{\eta}, \widetilde{\mu})$ sur la catégorie de Kleisli $Kl(T')$ (voir App.I.7), ce qui signifie que les diagrammes suivants commutent si, pour le second, on remplace le couple $(\Downarrow, \Uparrow)$ par chacun des couples $(\eta, \widetilde{\eta})$ et $(\mu, \widetilde{\mu})$ (où $F_{T'}$ est défini dans App.I.7) :

$$\begin{tikzcd} Kl(T') \arrow[r, "\widetilde{T}"] & Kl(T') \\ \mathbf{C} \arrow[u, "F_{T'}"] \arrow[r, "T"'] & \mathbf{C} \arrow[u, "F_{T'}"'] \end{tikzcd} \qquad \begin{tikzcd} Kl(T') \arrow[r, shift left] \arrow[r, shift right] \arrow[r, phantom, "\Uparrow"] & Kl(T') \\ \mathbf{C} \arrow[u, "F_{T'}"] \arrow[r, shift left] \arrow[r, shift right] \arrow[r, phantom, "\Downarrow"] & \mathbf{C} \arrow[u, "F_{T'}"'] \end{tikzcd}$$

Pour dire cela brièvement, ces diagrammes signifient que, moyennant l'identification de la catégorie $\mathbf{C}$ à son image $F_{T'}(\mathbf{C})$ dans $Kl(T')$ (voir App.I.7), la restriction de la monade $\widetilde{T}$ à $\mathbf{C}$ est la monade T. De plus, pour tout morphisme de Kleisli $u : X \dashrightarrow Y$ (se reporter à App.I.7 pour cette notation), $\widetilde{T}u$ est le morphisme composé $\tau Y \circ Tu : TX \dashrightarrow TY$.

$$\begin{tikzcd} & T'Y \\ X \arrow[ur, "u"] \arrow[r, dashed, "u"'] & Y \end{tikzcd} \qquad \begin{tikzcd} & TT'Y \arrow[r, "\tau Y"] & T'TY \\ TX \arrow[ur, "Tu"] \arrow[urr, "\widetilde{T}u"] \arrow[rr, dashed, "\widetilde{T}u"'] & & TY \end{tikzcd}$$

2.4. Si $\theta : TX \longrightarrow T'X$ est un morphisme de $\mathbf{C}$, alors (X,θ) est une τ-algèbre dans $\mathbf{C}$ ssi c'est une $\widetilde{T}$-algèbre dans $Kl(T')$; en effet, remarquant que θ est en fait un morphisme de Kleisli $TX \dashrightarrow X$, ça n'est qu'une question de traduction dans $Kl(T')$ des axiomes (A'_i) de τ-algèbres donnés dans 2.1 : $\theta \circ \eta X = \eta' X$ et $\theta \circ \mu X = \mu' X \circ T'\theta \circ \tau X \circ T\theta$ s'écrivent (voir App.I.7) $\theta \circ \eta X = \overline{id}_X$ et $\theta \circ \mu X = \mu' X \circ T'\theta \circ \widetilde{T}\theta = \theta \bar{\circ} \widetilde{T}\theta$, ou encore, moyennant le fait que la monade T est la restriction de la monade $\widetilde{T}$ à $\mathbf{C}$, elle-même considérée comme une sous-catégorie de $Kl(T')$, $\theta \bar{\circ} \widetilde{\eta} X = \overline{id}_X$ et $\theta \bar{\circ} \widetilde{\mu} X = \theta \bar{\circ} \widetilde{T}\theta$, qui ne sont autres que les axiomes (A_i) de $\widetilde{T}$-algèbres rappelés dans App.I.3.

2.5. Ceci nous conduit à définir la catégorie $Alg(\tau)$ des τ-algèbres comme étant égale à la catégorie $Alg(\widetilde{T})$ des $\widetilde{T}$-algèbres ; ses objets sont donc les τ-algèbres et ses morphismes les $u : (X,\theta) \dashrightarrow (X',\theta')$, i.e les $u : X \dashrightarrow X'$ vérifiant $u \bar{\circ} \theta = \theta' \bar{\circ} \widetilde{T}u$ (ce qui s'écrit $T'f \circ \theta = \theta' \circ Tf$ dans le cas particulier où $f : X \longrightarrow X'$ est un morphisme de $\mathbf{C}$).

On a un foncteur $F : Alg(T) \longrightarrow Alg(\tau)$, défini par $F(X,\theta) = (X, \eta' X \circ \theta)$ et $F(f) = \eta' X' \circ f$ (où $f : (X,\theta) \longrightarrow (X',\theta')$) qui permet d'identifier $Alg(T)$ à une sous-catégorie de $Alg(\tau)$; il admet un adjoint à droite $U : Alg(\tau) \longrightarrow Alg(T)$, défini par $U(X,\theta) = (T'X, \mu'X \circ T'\theta \circ \tau X)$ (voir 2.2) et $U(u) = \mu'X' \circ T'u$ (où $u : (X,\theta) \dashrightarrow (X',\theta')$). Ce couple d'adjoints (F,U) (c'est en fait un relèvement du couple d'adjoints $(F_{T'}, U_{T'})$ qui induit la monade T' sur $\mathbf{C}$ (voir App.I.7)) induit lui-même la monade $\widetilde{T'}$ sur $Alg(T)$, de sorte que l'on a un *foncteur de comparaison* plein et fidèle $Alg(\tau) \xrightarrow{\phi_\tau} Alg(\widetilde{T'})$, défini par $\phi_\tau(X,\theta) = ((T'X, \mu'X \circ T'\theta \circ \tau X), \mu'X)$ (voir 2.2 et App.I.6) et $\phi_\tau(u) = \mu'X' \circ T'u$ (où u est comme ci-dessus), qui vérifie les relations $\phi_\tau \circ F = F^{\widetilde{T'}}$ et $U^{\widetilde{T'}} \circ \phi_\tau = U$; ainsi la catégorie $Alg(\tau)$ est équivalente à la sous-catégorie pleine $\phi_\tau(Alg(\tau))$ de $Alg(\widetilde{T'}) \simeq Alg(\widehat{T})$ dont les objets sont les $((T'X,\theta),\mu'X)$ (voir 2.2 et App.I.6). Ce foncteur est un relèvement du foncteur de comparaison $Kl(T') \xrightarrow{\phi} Alg(T')$ rappelé dans App.I.7.

3. Automates non déterministes

3.1. Introduction : Les automates non déterministes constituent une première variante du cas déterministe. Soit Σ un ensemble fini ; dans un *automate non déterministe* sur l'alphabet Σ, l'action d'un mot sur un état a pour résultat, non pas un nouvel état (comme dans le cas déterministe), mais un ensemble d'états "possibles" (voir App.II.2) ; la *loi de transition* de cet automate est donc une application de la forme $\Sigma^* \times X \longrightarrow \mathcal{P}X$, où $\mathcal{P}X$ est l'ensemble des parties de X. Précisément, on peut redéfinir les automates non déterministes sur un alphabet Σ comme les algèbres d'une loi distributive entre la monade associée au monoïde Σ^* et la "monade des parties". Nous généralisons en prenant un monoïde M quelconque au lieu de Σ^* (voir [Bu' 74]), et nous appellerons *M-automate non déterministe*, une τ-algèbre, où $\tau : T_M/T_\mathcal{P}$ est la loi distributive que nous allons maintenant définir.

3.2. Précisément, considérons, comme dans le paragraphe 1, la monade T_M sur **Ens** associée à un monoïde $M = (M,e,m)$ donné ; considérons aussi la monade des parties $T_\mathcal{P} = (\mathcal{P}, \eta', \mu')$ sur **Ens**, où l'endofoncteur $\mathcal{P} : \mathbf{Ens} \longrightarrow \mathbf{Ens}$ associe à tout ensemble X l'ensemble $\mathcal{P}X$ de ses parties, et à toute application $f : X \longrightarrow Y$ l'application $\mathcal{P}X \longrightarrow \mathcal{P}Y$ qui, à toute partie A de X, associe son image $f(A)$ dans Y ; de plus, $\eta'X : X \longrightarrow \mathcal{P}X$ et $\mu'X : \mathcal{P}\mathcal{P}X \longrightarrow \mathcal{P}X$ sont respectivement les applications singleton et réunion. On montre facilement que l'application $\tau X : M \times \mathcal{P}X \longrightarrow \mathcal{P}(M \times X)$, définie, pour tout $(a,A) \in M \times \mathcal{P}X$, par $\tau X(a,A) = \{a\} \times A$, est la transformation naturelle d'une loi distributive $\tau : T_M/T_\mathcal{P}$. Une τ-algèbre est donc ici la donnée d'un ensemble X muni d'une application $\theta : M \times X \longrightarrow \mathcal{P}X$ vérifiant les axiomes (A'_i) ci-dessus, i.e tels que, pour tout $a,b \in M$ et tout $x \in X$, on ait :

$$\theta(e,x) = \{x\} \quad \text{et} \quad \theta(ab,x) = \cup_{y\in\theta(b,x)}\theta(a,y)$$

Se reportant à App.II.2, on obtient la proposition suivante :

Proposition 2 : *Dans le cas où $M = \Sigma^*$ (et où Σ est un ensemble fini), une τ-algèbre est la donnée d'un automate non déterministe sur l'alphabet Σ.*

3.3. La donnée d'une τ-algèbre $M \times X \xrightarrow{\theta} \mathcal{P}X$ est équivalente (voir 2.2 ci-dessus) à la donnée d'une T_M-algèbre $M \times \mathcal{P}X \xrightarrow{\widetilde{\theta}} \mathcal{P}X$ qui est τ-compatible avec la multiplication $\mu' X$, i.e telle que, pour tout $a \in M$, l'opérateur $\widetilde{\theta}(a, -)$: $\mathcal{P}X \longrightarrow \mathcal{P}X$ vérifie, pour tout $\mathfrak{A} \in \mathcal{P}\mathcal{P}X$, $\widetilde{\theta}(a, \cup_{A\in\mathfrak{A}} A) = \cup_{A\in\mathfrak{A}} \widetilde{\theta}(a, A)$. Plus précisément, la T_M-algèbre $\widetilde{\theta}$ associée à la τ-algèbre θ est définie par $\widetilde{\theta}(a, A) = \bigcup_{x\in A} \theta(a, x)$ (ainsi l'automate non déterministe sur les états se transforme en un automate déterministe sur les parties d'états, avec une condition de commutation aux réunions).

3.4. La catégorie de Kleisli de la monade $T_{\mathcal{P}}$ des parties est (plus précisément : est équivalente à) la catégorie **Rel** des relations entre ensembles ; en effet, un morphisme $u : X \dashrightarrow Y$ dans cette catégorie de Kleisli (voir la convention sur la notation des morphismes de Kleisli et leur composition dans App.I.7), i.e une application $u : X \longrightarrow \mathcal{P}Y$ dans **Ens** équivaut à la donnée de la relation $R_u \subset X \times Y$ définie par $(x, y) \in R_u$ ssi $y \in u(x)$; on dira donc que $u : X \dashrightarrow Y$ est une relation. Si maintenant $v : Y \dashrightarrow Z$ est une autre relation, alors, d'après App.I.7, $(v \bar{\circ} u)(x) = (\mu' Z \circ \mathcal{P}v \circ u)(x) = \bigcup_{y\in u(x)} v(y)$; par suite $(x, z) \in R_{v\bar{\circ}u}$ ssi il existe $y \in Y$ vérifiant $(x, y) \in R_u$ et $(y, z) \in R_v$; on obtient donc la proposition suivante :

Proposition 3 : *la composition $\bar{\circ}$ dans la catégorie de Kleisli* **Rel** *est la composition habituelle des relations.*

D'après 2.3, on a un prolongement $\widetilde{T_M}$: **Rel** $\longrightarrow$ **Rel** qui, à toute relation $u : X \dashrightarrow Y$, associe la relation $\widetilde{T_M}u = \tau Y \circ (M \times u) : M \times X \dashrightarrow M \times Y$: $(a, x) \mapsto \{a\} \times u(x)$. En particulier, pour $\theta : M \times X \dashrightarrow X$, la relation $\widetilde{T_M}\theta = \tau X \circ (M \times \theta) : M \times M \times X \dashrightarrow M \times X$ est définie par $(\widetilde{T_M}\theta)(a, b, x) = \{a\} \times \theta(b, x)$.

De plus, se référant à 2.4, une τ-algèbre $\theta : M \times X \longrightarrow \mathcal{P}X$ dans **Ens** est aussi une $\widetilde{T_M}$-algèbre $\theta : M \times X \dashrightarrow X$ dans **Rel**.

4. Automates stochastiques

4.1. Introduction : Soit Σ un ensemble fini ; un *automate stochastique* sur l'alphabet Σ, lorsqu'il lit un mot, donne une probabilité sur les états (voir App.III.5) ; sa loi de transition (appelée *probabilité de transition*) est donc une application de la forme $\Sigma^* \times X \longrightarrow \mathbb{P}X$, où $\mathbb{P}X$ est l'ensemble des probabilités sur X (on doit donc supposer que X est un espace mesurable). On obtient alors les automates stochastiques sur un alphabet Σ comme les algèbres d'une loi distributive entre la monade associée au monoïde Σ^* et la "monade des probabilités", ces monades étant définies, non pas sur **Ens**, mais sur la catégorie des espaces mesurables. Nous généralisons en prenant un monoïde M quelconque au lieu de Σ^*.

4.2. Considérons la catégorie **Mes** des applications mesurables entre espaces mesurables. C'est une catégorie monoïdale (voir App.I.1) d'unité 1 et de loi monoïdale le produit d'espaces mesurables. Dans cette catégorie monoïdale, fixons un monoïde $M = (M, e, m)$ (M est un espace mesurable, son unité et sa multiplication sont des applications mesurables ; voir App.I.2) et considérons la monade $T_{_M} = (T_{_M}, \eta, \mu)$ qui lui est associée sur **Mes** (voir App.I.4 ; cette monade est similaire à celle que nous avions définie sur **Ens** dans le paragraphe 1, sauf qu'ici, les applications ηX et μX sont mesurables pour tout espace mesurable X).

4.3. Décrivons la monade des probabilités $T_{\mathbb{P}} = (\mathbb{P}, \eta'', \mu'')$ sur **Mes** (voir [Tull 71], [Giry 81], et App.IV où nous donnons une preuve détaillée du fait que c'est une monade).

Son endofoncteur $\mathbb{P}$: **Mes** $\longrightarrow$ **Mes** associe à tout espace mesurable X l'ensemble $\mathbb{P}X$ des probabilités sur X, que l'on munit de la plus petite tribu rendant mesurables les évaluations $e_A : \mathbb{P}X \longrightarrow [0,1] : p \mapsto p(A)$, lorsque A parcourt la tribu donnée sur X ($[0,1]$ étant muni de sa tribu borélienne). De plus, si $f : X \longrightarrow Y$ est une application mesurable, $\mathbb{P}f : \mathbb{P}X \longrightarrow \mathbb{P}Y$ est l'application définie, pour tout $p \in \mathbb{P}X$, par $\mathbb{P}f(p) = f_*(p)$, la p-loi de f sur Y (voir App.III.2).

L'unité et la multiplication de cette monade sont définies comme suit :

- $\eta''X : X \longrightarrow \mathbb{P}X$ est l'application qui associe la probabililité de Dirac δ_x à l'élément $x \in X$;
- $\mu''X : \mathbb{P}\mathbb{P}X \longrightarrow \mathbb{P}X$ est l'application qui, à toute probabilité $\hat{p}$ sur $\mathbb{P}X$, associe la probabilité $\mu''X(\hat{p})$ sur X ainsi définie : pour toute partie mesurable A de X, $(\mu''X(\hat{p}))(A) = \int_{\mathbb{P}X} e_A d\hat{p} = \int_{\mathbb{P}X} e_A(p)\hat{p}(dp) = \int_{\mathbb{P}X} p(A)\hat{p}(dp)$ (c'est donc la $\hat{p}$-moyenne des $p(A)$ lorsque p parcourt $\mathbb{P}X$; voir App.III.2 pour la notation des intégrales).

Bien entendu, les applications $\mathbb{P}f = f_*$, $\eta''X$ et $\mu''X$ sont mesurables (car leurs composés avec chaque évaluation e_A sont mesurables ...voir App.III.1 et App.IV).

4.4. On définit une loi distributive $\tau : T_{_M}/T_{\mathbb{P}}$ de la manière suivante : pour tout espace mesurable X, l'application $\tau X : M \times \mathbb{P}X \longrightarrow \mathbb{P}(M \times X)$ est définie, pour tout $(a, p) \in M \times \mathbb{P}X$, par $\tau X(a, p) = \delta_a \otimes p$ (produit des probabilités δ_a et p) ; on vérifie (voir App.V pour une preuve détaillée) que c'est une application mesurable pour tout espace mesurable X et que cela définit une transformation naturelle $\tau : T_{_M}\mathbb{P} \longrightarrow \mathbb{P}T_{_M}$ qui vérifie les axiomes (B_i) des lois distributives (voir App.I.5). On obtient ainsi une loi distributive $\tau : T_{_M}/T_{\mathbb{P}}$.

4.5. Proposition 4 : *Soit $\tau : T_{_M}/T_{\mathbb{P}}$ la loi distributive décrite dans 4.4 ci-dessus. Une τ-algèbre est la donnée d'un M-automate stochastique, i.e (voir App.III.5), c'est la donnée d'un espace mesurable X et d'une application mesurable $\theta : M \times X \longrightarrow \mathbb{P}X$ tels que, pour tout $a, b \in M$, tout $x \in X$ et toute partie* *mesurable A de X, on ait :*

$$\theta(e,x)=\delta_x \quad \text{et} \quad (\theta(ab,x))(A) \overset{\star}{=} \int_X (\theta(a,y))(A)(\theta(b,x))(dy)$$

Dans le cas où $M=\Sigma^*$ (et où Σ est un ensemble fini), une τ-algèbre est la donnée d'un automate stochastique sur l'alphabet Σ.

Preuve : Il s'agit donc de comparer les axiomes (A'_i) de τ-algèbre (voir 2.1) avec les axiomes ci-dessus. L'axiome (A'_1) donne immédiatement $\theta(e,x)=\delta_x$ pour tout $x\in X$. Pour l'axiome (A'_2), soit $a,b\in M$, $x\in X$ et A une partie mesurable de X ; on obtient alors d'une part $((\theta\circ\mu X)(a,b,x))(A)=(\theta(ab,x))(A)$ et, d'autre part $((\mu''X\circ\mathbb{P}\theta\circ\tau X\circ(M\times\theta))(a,b,x))(A)=(\mu''X(\theta_*(\delta_a\otimes\theta(b,x))))(A)=\int_{\mathbb{P}X} e_A d(\theta_*(\delta_a\otimes\theta(b,x)))=\iint_{M\times X} e_A\circ\theta d(\delta_a\otimes\theta(b,x))=\iint_{M\times X}(\theta(c,y))(A)(\delta_a\otimes\theta(b,x))(dcdy)=\int_X(\theta(a,y))(A)(\theta(b,x))(dy)$. D'où l'égalité $\overset{\star}{=}$ cherchée. □

4.6. La donnée d'une τ-algèbre $M\times X\xrightarrow{\theta}\mathbb{P}X$ est équivalente (d'après 2.2 ci-dessus) à la donnée d'une T_M-algèbre $M\times\mathbb{P}X\xrightarrow{\widetilde{\theta}}\mathbb{P}X$, τ-compatible avec la multiplication $\mu''X$ (ainsi l'espace des probabilités d'un automate stochastique est un automate déterministe τ-compatible avec la multiplication de la monade $T_{\mathbb{P}}$ des probabilités), i.e telle que, pour tout $a\in M$ et tout $\hat{p}\in\mathbb{P}\mathbb{P}X$, on ait l'égalité : $\widetilde{\theta}(a,\mu''X(\hat{p}))=\mu''X(\widetilde{\theta}_*(\delta_a\otimes\hat{p}))$. Plus précisément, la T_M-algèbre $\widetilde{\theta}$ associée à la τ-algèbre θ est définie par $(\widetilde{\theta}(a,p))(A)=(\mu''X(\theta_*(\delta_a\otimes p)))(A)=\int_{\mathbb{P}X}e_A d(\theta_*(\delta_a\otimes p))=\iint_{M\times X}e_A\circ\theta d(\delta_a\otimes p)=\iint_{M\times X}(\theta(b,x))(A)(\delta_a\otimes p)(dbdx)=\int_X\theta(a,x)(A)p(dx)$ pour toute partie mesurable A de X.

4.7. La catégorie de Kleisli de la monade $T_{\mathbb{P}}$ des probabilités est la catégorie, notée **Stoch**, dont les objets sont les espaces mesurables et les morphismes les probabilités de transition, puisqu'un morphisme de Kleisli $u:X\dashrightarrow Y$ est une application mesurable $u:X\longrightarrow\mathbb{P}Y$ (voir App.I.7 et App.III.4). De plus, si $u:X\dashrightarrow Y$ et $v:Y\dashrightarrow Z$ sont deux probabilités de transition, leur composé $v\bar{\circ}u$ dans la catégorie de Kleisli **Stoch** est ainsi défini (voir App.I.7) : si $x\in X$ et si $\hat{p}$ est la probabilité image $v_*(u(x))$ sur $\mathbb{P}Z$ (c'est donc un élément de $\mathbb{P}\mathbb{P}Z$), alors $(v\bar{\circ}u)(x)=\mu''Z(\hat{p})$ est la probabilité sur Z qui, à toute partie mesurable C de Z, associe le réel $(\mu''Z(\hat{p}))(C)=\int_{\mathbb{P}Z}e_C d\hat{p}=\int_Y e_C\circ vd(u(x))=\int_Y e_C(v(y))(u(x))(dy)=\int_Y(v(y))(C)(u(x))(dy)$. On a donc obtenu l'égalité :

$$((v\bar{\circ}u)(x))(C)=\int_Y(v(y))(C)(u(x))(dy)$$

pour tout x de X et toute partie mesurable C de Z ; d'où la proposition suivante :

Proposition 5 [Giry 81] : *la composition $\bar{\circ}$ dans la catégorie de Kleisli* **Stoch** *est exactement la composition habituelle des probabilités de transition rappelée dans* App.III.4 : *$v\bar{\circ}u=vu$.*

D'après 2.3 ci-dessus, on a un prolongement $\widetilde{T_M}$: **Stoch** $\longrightarrow$ **Stoch** qui, à toute probabilité de transition $u:X\dashrightarrow Y$ associe la probabilité de transition $\widetilde{T_M}u=\tau Y\circ(M\times u):M\times X\dashrightarrow M\times Y:(a,x)\mapsto\delta_a\otimes u(x)$. En particulier, pour $\theta : M \times X \dashrightarrow X$, la probabilité de transition $\widetilde{T_M}\theta = \tau X \circ (M \times \theta)$: $M \times M \times X \dashrightarrow M \times X$ est définie par $(\widetilde{T_M}\theta)(a,b,x) = \delta_a \otimes \theta(b,x)$.

De plus, se référant à 2.4, une τ-algèbre $\theta : M \times X \longrightarrow \mathbb{P}X$ dans **Mes** est aussi une $\widetilde{T_M}$-algèbre $\theta : M \times X \dashrightarrow X$ dans **Stoch**.

5. Processus et chaînes de Markov homogènes

On conserve ici la loi distributive $\tau : T_M/T_{\mathbb{P}}$ du paragraphe 4.

5.1. Se référant aux App.III.6 et App.III.7, on obtient les corollaires suivants de la proposition 4 :

Corollaire 1 : *Dans le cas où le monoïde M est $\mathbf{R}_+$ (avec l'addition usuelle), que l'on munit de la tribu borélienne, une τ-algèbre étant un $\mathbf{R}_+$-automate stochastique, c'est donc un processus Markovien homogène.*

Corollaire 2 : *Dans le cas où le monoïde M est $\mathbf{N}$ (toujours pour l'addition usuelle), muni de la tribu discrète, une τ-algèbre étant un $\mathbf{N}$-automate stochastique (le monoïde $\mathbf{N}$ est librement engendré par $\Sigma = 1$), c'est donc une chaine de Markov homogène.*

5.2. Dans la catégorie de Kleisli **Stoch**, on a un *objet des entiers naturels* (i.e un NNO ... voir App.I.8) : c'est simplement l'ensemble $\mathbf{N}$ des entiers naturels muni de sa tribu discrète et des deux applications $z : 1 \longrightarrow \mathbf{N} : 0 \mapsto 0$ et $s : \mathbf{N} \longrightarrow \mathbf{N} : n \mapsto n+1$ (c'est un NNO dans **Ens**, donc dans **Mes** (car $\mathbf{N}$ est discret), et donc aussi dans **Stoch**, puisque le foncteur $F_{T_{\mathbb{P}}}$: **Mes** $\longrightarrow$ **Stoch** possède un adjoint à droite $U_{T_{\mathbb{P}}}$ (voir App.I.7, App.I.8 et [Bu 86])). Plus concrètement, si $1 \overset{\nu}{\dashrightarrow} X$ et $X \overset{\pi}{\dashrightarrow} X$ sont deux probabilités de transition, il existe une unique probabilité de transition $\varphi : \mathbf{N} \dashrightarrow X$ définie par la récurrence $\varphi(0) = \nu$ et $\varphi(n+1) = (\pi \bar{\circ} \varphi)(n)$, i.e rendant les diagrammes ci-dessous commutatifs dans la catégorie **Stoch** :

$$\begin{tikzcd} 1 \arrow[r, "z"] \arrow[dr, "\nu"'] & \mathbf{N} \arrow[r, "s"] \arrow[d, "\varphi"] & \mathbf{N} \arrow[d, "\varphi"] \\ & X \arrow[r, "\pi"'] & X \end{tikzcd}$$

Ceci traduit de manière catégorique le fait, rappelé dans App.III.6, que la donnée d'une probabilité ν sur X et d'une probabilité de transition $\pi : X \dashrightarrow X$ détermine entièrement les lois d'une chaîne de Markov homogène (elles s'écrivent, rappelons-le, $\pi_n \bar{\circ} \nu = \pi_n \nu$, où $\pi_n = \overbrace{\pi \bar{\circ} \ldots \bar{\circ} \pi}^{n \text{ fois}} = \overbrace{\pi \ldots \pi}^{n \text{ fois}}$). En effet, par définition de φ, ces lois s'écrivent ici $\pi_0 \bar{\circ} \nu = \nu = \varphi(0)$ et, pour tout $n \in \mathbf{N}^*$, $\pi_n \bar{\circ} \nu = \pi_n \bar{\circ} \varphi(0) = \pi_{n-1} \bar{\circ} (\pi \bar{\circ} \varphi(0)) = \pi_{n-1} \bar{\circ} \varphi(1) = \cdots = \pi \bar{\circ} \varphi(n-1) = \varphi(n)$.

APPENDICES

App.I. Rappels en théorie des catégories

App.I.1. Une *catégorie monoïdale* est définie par un 6-uplet $(\mathbf{V}, \otimes, I, \alpha, \lambda, \rho)$, où $\mathbf{V}$ est une catégorie, I un objet de $\mathbf{V}$, $\otimes : \mathbf{V} \times \mathbf{V} \longrightarrow \mathbf{V}$ un foncteur (appelés, respectivement, unité et loi monoïdale de $\mathbf{V}$) et α, λ et ρ sont des isomorphismes naturels de la forme (où X, Y, Z sont des objets de $\mathbf{V}$) :

$$\alpha_{X,Y,Z} : X \otimes (Y \otimes Z) \simeq (X \otimes Y) \otimes Z, \quad \lambda_X : I \otimes X \simeq X \quad \rho_X : X \otimes I \simeq X,$$

vérifiant les axiomes suivants :

$$\begin{tikzcd} X \otimes (I \otimes Y) \arrow[rr, "\alpha"] \arrow[dr, "X\otimes\lambda"'] & & (X \otimes I) \otimes Y \arrow[dl, "\rho\otimes Y"] \\ & X \otimes Y & \end{tikzcd}$$

$$\begin{tikzcd} X \otimes (Y \otimes (Z \otimes T)) \arrow[r, "\alpha"] \arrow[d, "X\otimes\alpha"'] & (X \otimes Y) \otimes (Z \otimes T) \arrow[r, "\alpha"] & ((X \otimes Y) \otimes Z) \otimes T \\ X \otimes ((Y \otimes Z) \otimes T) \arrow[rr, "\alpha"'] & & (X \otimes (Y \otimes Z)) \otimes T \arrow[u, "\alpha\otimes T"'] \end{tikzcd}$$

($\lambda_I = \rho_I : I \otimes I \longrightarrow I$ est une conséquence). On sait que toute catégorie monoïdale est équivalente à une catégorie monoïdale stricte, i.e. une catégorie monoïdale dans laquelle les isomorphismes α, λ et ρ sont des identités (la loi monoïdale est associative et l'unité est neutre), voir [McL 72]. Cela permet de supposer que la catégorie monoïdale dans laquelle on travaille est stricte et donc qu'on peut la noter par un triplet $(\mathbf{V}, \otimes, I)$.

Exemples : La catégorie des ensembles, munie des produits cartésiens et de l'ensemble 1, est une catégorie monoïdale $(\mathbf{Ens}, \times, 1)$.

Pour toute catégorie $\mathbf{C}$, on a une catégorie monoïdale $(\mathbf{C}^{\mathbf{C}},\ , \mathrm{id}_C)$ (notation multiplicative de la loi monoïdale, c'est-à-dire par une absence de symblole !). Les objets de $\mathbf{C}^{\mathbf{C}}$ sont les endofoncteurs de $\mathbf{C}$ et les morphismes, dont la composition est notée par «$\circ$», sont les transformations naturelles.

App.I.2. Un *monoïde* (ou, une *monade*, au sens premier donné par Benabou [Ben 67]) dans une catégorie monoïdale $(\mathbf{V}, \otimes, I)$ est défini par un triplet (M, e, m), où M est un objet de $\mathbf{V}$, $e : I \longrightarrow M$ et $m : M \otimes M \longrightarrow M$ (resp. unité et multiplication du monoïde) sont des morphismes de $\mathbf{V}$ vérifiant $m \circ (e \otimes M) = M = m \circ (M \otimes e)$ et $m \circ (m \otimes M) = m \circ (M \otimes m)$.

$$\begin{tikzcd} I \arrow[r, "e"] & M \arrow[r, shift left=2, "e\otimes M"] \arrow[r, shift right=2, "M\otimes e"'] & M \otimes M \arrow[l, "m"'] & M \otimes M \otimes M \arrow[l, shift right=1, "m\otimes M"'] \arrow[l, shift left=1, "M\otimes m"] \end{tikzcd}$$

App.I.3. Une *monade* sur une catégorie $\mathbf{C}$ est un monoïde dans la catégorie monoïdale $\mathbf{C}^{\mathbf{C}}$ mentionnée plus haut ; c'est donc la donnée d'un triplet $T = (T, \eta, \mu)$, où $T : \mathbf{C} \longrightarrow \mathbf{C}$ est un foncteur, $\eta : id_{\mathbf{C}} \longrightarrow T$ et $\mu : TT \longrightarrow T$ (l'unité et la multiplication de la monade) des transformations naturelles vérifiant $\mu \circ \eta T = T = \mu \circ T\eta$ et $\mu \circ \mu T = \mu \circ T\mu$.

$$\begin{tikzcd}
id_{\mathbf{C}} \arrow[r, "\eta"] & T \arrow[r, shift left=2, "\eta T"] \arrow[r, shift right=2, "T\eta"'] & TT \arrow[l, "\mu" description] & TTT \arrow[l, shift right=1, "\mu T"'] \arrow[l, shift left=1, "T\mu"]
\end{tikzcd}$$

Etant donné une monade $T = (T, \eta, \mu)$ sur une catégorie $\mathbf{C}$, une T-*algèbre* est la donnée d'un couple (X, θ), où X est un objet de $\mathbf{C}$ et $\theta : TX \longrightarrow X$ un morphisme de $\mathbf{C}$ rendant commutatifs les diagrammes suivants dans $\mathbf{C}$:

$$(A_1) \quad \begin{tikzcd}
TX \arrow[rr, "\theta"] & & X \\
& X \arrow[ul, "\eta X"] \arrow[ur, "id_X"'] &
\end{tikzcd}
\qquad\qquad
(A_2) \quad \begin{tikzcd}
TTX \arrow[r, "T\theta"] \arrow[d, "\mu X"'] & TX \arrow[d, "\theta"] \\
TX \arrow[r, "\theta"'] & X
\end{tikzcd}$$

Toutes les structures libres définissent des monades : par exemple, la monade «monoïdes libres», la monade «groupes abéliens libres», etc ... ont pour algèbres respectives les monoïdes, les groupes abéliens, etc ...

On définit la catégorie $Alg(T)$ des T-algèbres comme étant la catégorie dont les objets sont les T-algèbres et les morphismes les $f : (X, \theta) \longrightarrow (X', \theta')$, où $f : X \longrightarrow X'$ est un morphisme de $\mathbf{C}$ vérifiant $f \circ \theta = \theta' \circ Tf$. On a un foncteur d'oubli $U^T : Alg(T) \longrightarrow \mathbf{C}$ défini par $U^T(X, \theta) = X$ et $U^T(f) = f : X \longrightarrow X'$ si $f : (X, \theta) \longrightarrow (X', \theta')$; ce foncteur admet un adjoint $F^T : \mathbf{C} \longrightarrow Alg(T)$ qui, à tout objet X de $\mathbf{C}$, associe la T-*algèbre libre* $(TX, \mu X)$ associée a X (ou engendrée par X) et, à tout morphisme $f : X \longrightarrow Y$ de $\mathbf{C}$, le morphisme $Tf : (TX, \mu X) \longrightarrow (TX', \mu X')$ de $Alg(T)$.

Le couple d'adjoints (F^T, U^T), décrit ci-dessus, induit la monade T sur $\mathbf{C}$.

App.I.4. Fixons-nous un monoïde (M, e, m) dans une catégorie monoïdale $(\mathbf{V}, \otimes, I)$; on lui associe naturellement une monade $T_M = (T_M, \eta, \mu)$ sur $\mathbf{V}$, où l'endofoncteur T_M associe l'objet $M \otimes X$ (resp. le morphisme $M \otimes f : M \otimes X \longrightarrow M \otimes Y$) à tout objet X (resp. à tout morphisme $f : X \longrightarrow Y$) de $\mathbf{V}$; de plus, pour tout objet X de $\mathbf{V}$, les morphismes $\eta X : X \longrightarrow M \otimes X$ et $\mu X : M \otimes M \otimes X \longrightarrow M \otimes X$ sont respectivement égaux à $e \otimes X$ et $m \otimes X$.

App.I.5. Soit $T = (T, \eta, \mu)$ et $T' = (T', \eta', \mu')$ deux monades sur une même catégorie $\mathbf{C}$. Une *loi distributive* τ *de* T *sur* T' (ce que l'on notera $\tau : T/T'$ (voir [Beck 69]) est la donnée d'un triplet $\tau = (\tau, T, T')$, où $\tau : TT' \longrightarrow T'T$ est une transformation naturelle rendant commutatifs les diagrammes suivants dans $\mathbf{C}^{\mathbf{C}}$:

$$(B_1) \quad \begin{tikzcd}
TT' \arrow[rr, "\tau"] & & T'T \\
& T \arrow[ul, "T\eta'"] \arrow[ur, "\eta' T"'] &
\end{tikzcd}
\qquad\qquad
(B_2) \quad \begin{tikzcd}
TT' \arrow[rr, "\tau"] & & T'T \\
& T' \arrow[ul, "\eta T'"] \arrow[ur, "T'\eta"'] &
\end{tikzcd}$$

$$(B_3)\quad \begin{tikzcd} TT'T' \arrow[rr, "\tau T'"] \arrow[dr, "T\mu'"'] & & T'TT' \arrow[rr, "T'\tau"] & & T'T'T \arrow[dl, "\mu' T"] \\ & TT' \arrow[rr, "\tau"'] & & T'T & \end{tikzcd}$$

$$(B_4)\quad \begin{tikzcd} TTT' \arrow[rr, "T\tau"] \arrow[dr, "\mu T'"'] & & TT'T \arrow[rr, "\tau T"] & & T'TT \arrow[dl, "T'\mu"] \\ & TT' \arrow[rr, "\tau"'] & & T'T & \end{tikzcd}$$

On démontre qu'une loi distributive est une monade dans la 2-catégorie des monades (c'est une des manifestations du caractère ubiquitaire des monades !)

Exemple originel : La loi distributive "modèle" relie la monade "monoïdes libres" à la monade "groupes abéliens libres". En effet, le schéma d'origine des lois distributives est donné par la distributivité, dans un anneau, de la multiplication par rapport à l'addition, qui fait passer d'un mot de combinaisons linéaires à une combinaison linéaire de mots ; autrement dit, cette distributivité est décrite par une application de la forme $\tau X : (\mathbf{Z}^{(X)})^* \longrightarrow \mathbf{Z}^{(X^*)}$, soumise à des conditions du type des (B_i) ci-dessus (où X^* et $\mathbf{Z}^{(X)}$ sont respectivement le monoïde libre et le groupe abélien libre engendrés par l'ensemble X).

App.I.6. Les équivalences de J.Beck.

a) Le fait que le groupe abélien libre engendré par un monoïde M soit encore un monoïde (qui est l'anneau libre des polynômes sur X quand $M = X^*$) a conduit J.Beck à établir que la donnée d'une loi distributive $\tau : T/T'$ est équivalente :

- d'une part à la donnée d'un relèvement de la monade T' sur $\mathbf{C}$ en une monade $\widetilde{T'} = (\widetilde{T'}, \widetilde{\eta'}, \widetilde{\mu'})$ sur la catégorie $Alg(T)$ des T-algèbres, ce qui signifie que les diagrammes suivants commutent si, pour le second, on remplace le couple $(\Downarrow, \Uparrow)$ par chacun des couples $(\eta', \widetilde{\eta'})$ et $(\mu', \widetilde{\mu'})$:

$$\begin{tikzcd} Alg(T) \arrow[r, "\widetilde{T'}"] \arrow[d, "U^T"'] & Alg(T) \arrow[d, "U^T"] \\ \mathbf{C} \arrow[r, "T'"'] & \mathbf{C} \end{tikzcd} \qquad \begin{tikzcd} Alg(T) \arrow[r, shift left] \arrow[r, shift right] \arrow[r, phantom, "\Uparrow"] \arrow[d, "U^T"'] & Alg(T) \arrow[d, "U^T"] \\ \mathbf{C} \arrow[r, shift left] \arrow[r, shift right] \arrow[r, phantom, "\Downarrow"] & \mathbf{C} \end{tikzcd}$$

En bref, si $\tau : T/T'$ est une loi distributive, le premier diagramme signifie que l'image, par le foncteur $\widetilde{T'}$, d'une T-algèbre sur X est une T-algèbre sur $T'X$ (plus précisément, $\widetilde{T'}(X, \theta) = (T'X, T'\theta \circ \tau X)$), et le second diagramme que $\eta' = \eta$ et $\mu' = \mu$ (plus précisément, $\widetilde{\eta'}(X, \theta) = \eta' X$ et $\widetilde{\mu'}(X, \theta) = \mu' X$).

- d'autre part, à la donnée d'une monade composée $\widehat{T} = (\widehat{T}, \widehat{\eta}, \widehat{\mu})$ sur $\mathbf{C}$, où $\widehat{T} = T'T$ (plus précisément, $\widehat{\eta} = \eta'\eta$ et $\widehat{\mu} = (\mu'\mu) \circ (T'\tau T)$).

b) S'inspirant du fait que la donnée d'une structure d'anneau est équivalente à la donnée de deux opérations $\times$ et $+$ telles que la première soit distributive par rapport à la seconde, J.Beck a aussi prouvé que, s'il existe une loi distributive $\tau : T/T'$, alors la donnée d'une $\widehat{T}$-algèbre $(X, \widehat{\theta})$ dans $\mathbf{C}$ est équivalente à la donnée d'une $\widetilde{T'}$-algèbre $((X,\theta),\theta')$ dans $Alg(T)$ (avec $\widehat{\theta} = \theta' \circ T'\theta$; $\theta = \widehat{\theta} \circ \eta' TX$ et $\theta' = \widehat{\theta} \circ T'\eta X$), i.e à la donnée d'une T-algèbre (X,θ) et d'une T'-algèbre (X,θ') qui sont τ-compatibles, i.e telles que le diagramme suivant commute dans $\mathbf{C}$:

$$\begin{tikzcd} & TT'X \arrow[r, "\tau X"] \arrow[dl, "T\theta'"'] & T'TX \arrow[dr, "T'\theta"] & \\ TX \arrow[dr, "\theta"'] & & & T'X \arrow[dl, "\theta'"] \\ & X \arrow[r, "id"'] & X & \end{tikzcd}$$

Ce que l'on peut exprimer en disant que les catégories d'algèbres $Alg(\widehat{T})$ et $Alg(\widetilde{T'})$ sont équivalentes.

App.I.7. Catégorie de Kleisli d'une monade.

Soit $T = (T, \eta, \mu)$ une monade sur une catégorie $\mathbf{C}$. La *catégorie de Kleisli de* T, notée $Kl(T)$, est la catégorie suivante : ses objets sont les mêmes que ceux de $\mathbf{C}$; un morphisme, noté $u : X \dashrightarrow Y$, dans $Kl(T)$ est un morphisme de $\mathbf{C}$ de la forme $u : X \longrightarrow TY$. Le composé $v\bar{\circ}u : X \dashrightarrow Z$ dans $Kl(T)$, de deux morphismes $u : X \dashrightarrow Y$ et $v : Y \dashrightarrow Z$, est défini par $v\bar{\circ}u = \mu Z \circ Tv \circ u$.

$$\begin{tikzcd} & & TTZ \arrow[d, "\mu Z"] \\ & TY \arrow[ur, "Tv"] & TZ \\ X \arrow[ur, "u"] \arrow[r, dashed, "u"'] & Y \arrow[ur, "v"] \arrow[r, dashed, "v"'] & Z \end{tikzcd}$$

L'identité, dans $Kl(T)$, associée à un objet X, est notée $\overline{id}_X$ et est définie par $\overline{id}_X = \eta X : X \dashrightarrow X$.

Entre les catégories $\mathbf{C}$ et $Kl(T)$, on a un foncteur $U_T : Kl(T) \longrightarrow \mathbf{C}$ défini, pour tout objet X de $Kl(T)$, par $U_T(X) = TX$ et, pour tout $u : X \dashrightarrow Y$, par $U_T(u) = \mu Y \circ Tu : TX \longrightarrow TY$. Ce foncteur a un adjoint à gauche $F_T : \mathbf{C} \longrightarrow Kl(T)$ défini, pour tout objet X, par $F_T(X) = X$ et, pour tout morphisme $f : X \longrightarrow Y$ de $\mathbf{C}$, par $F_T(f) = \eta Y \circ f : X \dashrightarrow Y$ (le foncteur F_T permet souvent d'identifier $\mathbf{C}$ à la sous-catégorie $F_T(\mathbf{C})$ de $Kl(T)$).

Le couple d'adjoints (F_T, U_T), décrit ci-dessus, induit la monade T sur $\mathbf{C}$.

On a un foncteur plein et fidèle (dit *foncteur de comparaison*) $Kl(T) \xrightarrow{\phi} Alg(T)$ qui vérifie les relations (voir I.3 ci-dessus pour les foncteurs U^T et F^T) : $\phi \circ F_T = F^T$ et $U^T \circ \phi = U_T$ (il est défini par $\phi(X) = (TX, \mu X)$ et $\phi(u) = \mu Y \circ Tu$ (où $u : X \dashrightarrow Y$)) ; ainsi, la catégorie $Kl(T)$ est èquivalente à la sous-catégorie pleine $\phi(Kl(T))$ de $Alg(T)$ dont les objets sont les T-algèbres libres $(TX, \mu X)$.

En fait, $Alg(T)$ est une complétion de $\phi(Kl(T)) \simeq Kl(T)$ par conoyaux ; plus précisément, si (X, θ) est une T-algèbre, le morphisme $(TX, \mu X) \xrightarrow{\theta} (X, \theta)$ est un conoyau dans $Alg(T)$ du couple de morphismes de $\phi(Kl(T))$:

$$(TTX, \mu TX) \underset{\mu X}{\overset{T\theta}{\rightrightarrows}} (TX, \mu X)$$

Autrement dit, la T-algèbre (X, θ) est un quotient de la T-algèbre libre $(TX, \mu X)$).

App.I.8. Un *objet des entiers naturels* (ou NNO), dans une catégorie $\mathbf{C}$ avec objet final 1, est la donnée d'un objet N, muni de deux morphismes $1 \xrightarrow{z} N \xrightarrow{s} N$ assujettis à la propriété universelle suivante : pour tout objet X de $\mathbf{C}$, muni de deux morphismes $1 \xrightarrow{f} X \xrightarrow{g} X$, il existe un unique morphisme $\varphi : N \longrightarrow X$ rendant les diagrammes suivants commutatifs dans C :

$$\begin{tikzcd} 1 \arrow[r, "z"] \arrow[dr, "f"'] & N \arrow[r, "s"] \arrow[d, "\varphi"] & N \arrow[d, "\varphi"] \\ & X \arrow[r, "g"'] & X \end{tikzcd}$$

Exemple : L'ensemble $\mathbf{N}$ des entiers naturels, muni des deux applications : $1 \xrightarrow{z} \mathbf{N} \xrightarrow{s} \mathbf{N}$, définies par $z(0) = 0$ et, pour tout $n \in \mathbf{N}$, par $s(n) = n + 1$, est évidemment un NNO dans la catégorie **Ens**.

Proposition 6 (A.Burroni [Bu 86]) : *Tout foncteur admettant un adjoint à droite commute aux NNO (en fait, il commute aux extensions de Kan indexées par des graphes).*

App.II. Rappels en théorie des automates déterministes et non déterministes

On fixe un ensemble fini Σ qu'on appelle *alphabet* et dont les éléments sont appelés *symboles*. On note Σ^* le monoïde libre engendré par Σ. Les éléments de Σ^* s'appellent des *mots*. Un mot $a \in \Sigma^*$ est une suite finie $a = \alpha_1 \alpha_2 \dots \alpha_n$ de symboles $(\alpha_1, \alpha_2, \dots, \alpha_n \in \Sigma)$. L'entier n est la longueur de a. En particulier, pour $n = 0$, on a le *mot vide*, qu'on notera e, et qui est l'unité du monoïde Σ^*. Pour les deux définitions qui suivent, voir [Star 72].

App.II.1. Un *automate déterministe* sur l'alphabet Σ est la donnée d'un couple (X, δ), où X est un ensemble et $\delta : \Sigma \times X \longrightarrow X$ une application appelée *fonction de transition*. Elle s'étend (par induction sur la longueur des mots) en une application $\theta : \Sigma^* \times X \longrightarrow X$ (appelée *loi de transition* de l'automate déterministe) vérifiant, pour tout $a, b \in \Sigma^*$ et tout $x \in X$, les deux axiomes suivants :

$$\theta(e, x) = x \quad \text{et} \quad \theta(ab, x) = \theta(a, \theta(b, x))$$

Les éléments de X sont appelés les *états* de l'automate. L'égalité $\theta(a, x) = y$ s'interprète en disant que, lisant a, l'automate passe de l'état x à l'état y. Les deux axiomes ci-dessus reviennent à dire, d'une part que le mot vide ne fait pas changer d'état l'automate et, d'autre part que les mots agissent *séquentiellement* (i.e symbole après symbole) sur les états.

App.II.2. Un *automate non déterministe* sur l'alphabet Σ est encore la donnée d'un couple (X, δ), mais où la fonction de transition est cette fois de la forme $\delta : \Sigma \times X \longrightarrow \mathcal{P}X$ (où $\mathcal{P}X$ est l'ensemble des parties de X). Comme précédemment, l'application δ s'étend par induction en une application $\theta : \Sigma^* \times X \longrightarrow \mathcal{P}X$ (encore appelée *loi de transition* de l'automate non déterministe) vérifiant, pour tout $a, b \in \Sigma^*$ et tout $x \in X$, les deux axiomes suivants :

$$\theta(e, x) = \{x\} \quad \text{et} \quad \theta(ab, x) = \bigcup_{y \in \theta(b,x)} \theta(a, y)$$

Ici encore, les éléments de X sont appelés les *états* de l'automate. L'égalité $\theta(a, x) = A$ s'interprète en disant que A est l'ensemble des états "possibles" dans lesquels l'automate peut se mettre, partant de l'état x, quand il lit a ; les deux axiomes ci-dessus traduisent encore le fait que le mot vide n'a aucune action sur l'automate et que les mots agissent séquentiellement.

App.III. Rappels en théorie des probabilités et en théorie des automates stochastiques

Si X, Y, Z, etc ... sont des espaces mesurables, on désignera par $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, *etc* ... leurs tribus respectives. On considère l'ensemble $\mathbf{R}$ (ainsi que toutes ses parties) muni de la tribu borélienne.

App.III.1. Une probabilité sur l'espace mesurable X est une mesure positive p vérifiant $p(X) = 1$. Pour tout espace mesurable X, on désigne par $\mathbb{P}X$ l'ensemble des probabilités sur X ; dans tout cet article, l'ensemble $\mathbb{P}X$ sera considéré comme un espace mesurable en le munissant de la plus petite tribu rendant mesurables les évaluations $e_A : \mathbb{P}X \longrightarrow [0, 1] : p \mapsto p(A)$, où A parcourt la tribu $\mathcal{X}$ sur X. Ce qui équivaut à dire que, pour tout espace mesurable Y, une application $u : Y \longrightarrow \mathbb{P}X$ est mesurable ssi tous les composés $e_A \circ u$ sont mesurables, lorsque A parcout la tribu $\mathcal{X}$ sur X.

App.III.2. Soit $f : X \longrightarrow Y$ une application mesurable (appelée aussi *variable aléatoire* en théorie des probabilités) et p une probabilité sur X. On définit une probabilité image $f_*(p)$ sur Y en posant $(f_*(p))(B) = p(\overset{-1}{f}(B))$ pour toute partie mesurable B de Y ; on appelle habituellement *p-loi* de f cette probabilité image $f_*(p)$, et on note $p(f \in B)$ l'expression $p(\overset{-1}{f}(B))$ interprétée comme étant la *probabilité que f soit dans B*. Dans le cas où l'on a $p(f \in B) = 1$, on dit que f est *p-presque sûrement dans B*.

Dans le cas où $Y = \mathbf{R}$ et où f est mesurable positive, l'intégrale de f par rapport à p a un sens (elle peut être infinie) et se note $\int_X f dp = \int_X f(x)p(dx)$; si f est mesurable bornée, elle est p-intégrable (i.e son intégrale par rapport à p existe et est finie).

App.III.3. Soit X et Y deux espaces mesurables. On appelle *probabilité de transition* de X vers Y ([Nev 64]) une application $\pi : X \times \mathcal{Y} \longrightarrow [0,1]$ telle que, d'une part, pour tout $x \in X$, l'application $\pi(x, -) : \mathcal{Y} \longrightarrow [0,1]$ est une probabilité sur Y, et, d'autre part, pour tout $B \in \mathcal{Y}$, l'application $\pi(-, B) : X \longrightarrow [0,1]$ est mesurable.

Soit donc π une telle probabilité de transition de X vers Y, $f : Y \longrightarrow \mathbf{R}$ une application mesurable (positive ou bornée) et $x \in X$; on note $\int_Y f(y)\pi(x, dy)$ l'intégrale de f par rapport à la probabilité $\pi(x, -)$ sur Y.

Si π' est une autre probabilité de transition de Y vers Z, on définit une nouvelle probabilité de transition de X vers Z, appelée produit (ou composition) des probabilités de transition π et π', en posant, pour tout $x \in X$ et tout $C \in \mathcal{Z}$:

$$(\pi'\pi)(x, C) \overset{(1)}{=} \int_Y \pi'(y, C)\pi(x, dy)$$

Enfin, les probabilités de transition opèrent sur les probabilités de la manière suivante : si ν est une probabilité sur X, π une probabilité de transition de X vers Y, on définit une probabilité $\pi\nu$ sur Y en posant, pour tout $B \in \mathcal{Y}$:

$$(\pi\nu)(B) \overset{(2)}{=} \int_X \pi(x, B)\nu(dx)$$

En fait, la formule (2) est un cas particulier de la formule (1) en faisant $X = 1$ dans (1).

App.III.4. Considérons maintenant une application mesurable $u : X \longrightarrow \mathbb{P}Y$; une telle donnée équivaut au fait que, pour tout $x \in X$, $u(x)$ est une probabilité sur Y et que, par définition même de la tribu sur $\mathbb{P}Y$ (voir III.1 ci-dessus), pour toute partie mesurable B de Y, l'application $e_B \circ u : X \longrightarrow [0,1] : x \mapsto (u(x))(B)$ est mesurable ; ainsi, la donnée d'une application mesurable $u : X \longrightarrow \mathbb{P}Y$ est équivalente à la donnée de la probabilité de transition π de X vers Y définie par $\pi(x, B) = (u(x))(B)$ pour tout $x \in X$ et tout $B \in \mathcal{Y}$ ([Giry 81]).

Définition 1 : On appellera dorénavant *probabilité de transition* une application mesurable de la forme $u : X \longrightarrow \mathbb{P}Y$; en particulier, toute probabilité ν sur X s'identifie (en posant $\nu(0) = \nu$) à une probabilité de transition $\nu : 1 \longrightarrow \mathbb{P}X$.

Les formules (1) et (2) données dans III.3 ci-dessus se traduisent donc ici par :

$$((vu)(x))(C) \overset{(1')}{=} \int_Y (v(y))(C)(u(x))(dy) \quad \text{et} \quad (u\nu)(B) \overset{(2')}{=} \int_X (u(x))(B)\nu(dx)$$

pour tout $x \in X$, tout $C \in \mathcal{Z}$, toute probabilité ν sur X, et toutes probabilités de transition $u : X \longrightarrow \mathbb{P}Y$ et $v : Y \longrightarrow \mathbb{P}Z$.

App.III.5. Soit Σ un ensemble fini. Un *automate stochastique* sur l'alphabet Σ (voir [Star 72]) est la donnée d'un couple (X, θ), où X est un ensemble et $\theta : \Sigma^* \times X \longrightarrow \mathbb{P}X$ une application mesurable vérifiant les deux axiomes suivants :

$$\theta(e, x) = \delta_x \quad \text{et} \quad \theta(ab, x)(A) \overset{(1'')}{=} \int_X (\theta(a, y))(A)(\theta(b, x))(dy)$$

(où δ_x est la probabilité de Dirac au point x) pour tout $a, b \in \Sigma^*$, tout $x \in X$ et tout $A \in \mathcal{X}$. Les éléments de X sont appelés les *états* de l'automate et θ sa *probabilité de transition* (ce qui est bien naturel, d'après III.4 ci-dessus). L'égalité $\theta(a, x) = p$ s'interprète en disant que, pour toute partie mesurable A de X, $p(A)$ est la probabilité que l'automate, partant de l'état x et lisant a, aille dans A ; les deux axiomes ci-dessus signifient, d'une part que, partant d'un état x et lisant le mot vide, l'automate reste presque-sûrement dans toute partie mesurable de X contenant x et, d'autre part que cet automate fonctionne encore de façon séquentielle.

La similitude entre les formules $(1')$ et $(1'')$ n'a rien de fortuit : en effet, remarquant que, pour tout $a \in \Sigma^*$, l'application partielle $\pi_a = \theta(a, -)$ est aussi une probabilité de transition $X \longrightarrow \mathbb{P}X$, la formule $(1'')$ peut encore s'écrire, pour tout $a, b \in \Sigma^*$, tout $x \in X$ et tout $A \in \mathcal{X}$:

$$(\pi_{ab}(x))(A) = \int_X (\pi_a(y))(A)(\pi_b(x))(dy) = ((\pi_a\pi_b)(x))(A)$$

On obtient ainsi la *relation de Chapman-Kolmogoroff* (généralisée) des automates stochastiques sur un alphabet Σ (qui est en fait équivalente à la formule $(1'')$, i.e au second axiome des automates stochastiques) :

$$\pi_{ab} = \pi_a\pi_b \quad \text{pour tout} \quad a, b \in \Sigma^*$$

Définition 2 : On appelle *M-automate stochastique* (où M est un monoïde quelconque) la donnée d'un couple (X, θ), où X est un espace mesurable et $\theta : M \times X \longrightarrow \mathbb{P}X$ une probabilité de transition vérifiant les mêmes propriétés formelles que ci-dessus, i.e, avec les notations adoptées précédemment, vérifiant, pour tout $a, b \in M$ et tout $x \in X$, les axiomes :

$$\pi_e(x) = \delta_x \quad \text{et} \quad \pi_{ab} = \pi_a\pi_b$$

Utilisant la formule $(2')$ de III.4, on a une probabilité $\pi_a\nu$ sur X pour toute probabilité ν sur X qui est définie par $(\pi_a\nu)(A) = \int_X (\pi_a(x))(A)\nu(dx)$ pour toute partie mesurable A de X ; en particulier $\pi_a\delta_x = \pi_a(x)$ et $(\pi_e\nu)(A) = \int_X (\pi_e(x))(A)\nu(dx) = \int_X \delta_x(A)\nu(dx) = \int_X 1_A(x)\nu(dx) = \nu(A)$; ainsi $\pi_e\nu = \nu$.

App.III.6. Une *chaîne de Markov homogène* est un automate stochastique sur l'alphabet $\Sigma = 1$ (on identifie donc Σ^* à $\mathbf{N}$, avec son addition usuelle, que l'on munit de sa tribu discrète). Soit $\theta : \mathbf{N} \times X \longrightarrow \mathbb{P}X$ une telle chaîne de Markov, et notons encore $\pi_n(x) = \theta(n, x)$ pour tout $n \in \mathbf{N}$ et tout $x \in X$ ($(\pi_n(x))(A)$ étant la probabilité de passer de x dans A en n pas). La mesurabilité de θ étant équivalente à celle de tous les $\pi_n : X \longrightarrow \mathbb{P}X$, cette chaîne de Markov est équivalente à la donnée de la suite (π_n) de probabilités de transition vérifiant, pour tout $n, m \in \mathbf{N}$ et tout $x \in X$, les axiomes :

$$\pi_0(x) = \delta_x \quad \text{et} \quad \pi_{n+m} = \pi_n \pi_m$$

On dit que $\pi_n(x) = \pi_n \delta_x$ est *la loi de la chaîne au temps n, lorsqu'elle part de* x, et, plus généralement, si ν est une probabilité sur X, que $\pi_n \nu$ est *la loi de la chaîne au temps n, pour la loi initiale ν.*

En fait, grâce à la relation de Chapman-Kolmogoroff $\pi_{n+m} = \pi_n \pi_m$, les π_n avec $n \in \mathbf{N}^*$ sont entièrement déterminés par la probabilité de transition $\pi_1 : X \longrightarrow \mathbb{P}X$, de sorte que les lois d'une chaîne de Markov homogène sont entièrement déterminées par la donnée d'une probabilité de transition $\pi : X \longrightarrow \mathbb{P}X$ (appelée encore *probabilité de transition* de la chaîne de markov ; alors $\pi_n = \overbrace{\pi \ldots \pi}^{n \text{ fois}}$) et d'une probabilité ν sur X (on obtient donc toutes les lois $\pi_n \nu$ de la chaîne, et en particulier sa loi initiale $\pi_0 \nu$ qui n'est autre que ν, d'après III.5).

Les probabilistes ont coutume d'appeler plutôt chaîne de Markov homogène la donnée d'une suite de variables aléatoires dont les lois (voir III.2 ci-dessus) sont du type des $\pi_n \nu$ ci-dessus. Il est bien connu des probabilistes que l'on peut toujours construire, a posteriori, une telle chaîne de Markov homogène dont la probabilité de transition $\pi : X \longrightarrow \mathbb{P}X$ et la loi initiale ν sont données a priori : on considère l'espace $X^{\mathbf{N}}$, muni de la tribu produit ; on peut alors construire une unique probabilité P_ν (l'indice ν étant là pour "symboliser" le fait que l'on part avec la loi ν) sur cet espace produit, telle que la suite $(\mathbf{q}_n)$ des projections canoniques $X^{\mathbf{N}} \longrightarrow X$ réponde à la question. En effet, la tribu produit étant la plus petite tribu rendant mesurables les $\mathbf{q}_n$, P_ν est entièrement déterminée par les $P_\nu(\mathbf{q}_n \in A)$ pour tout $n \in \mathbf{N}$ et toute partie mesurable A de X (i.e par les P_ν-lois des $\mathbf{q}_n$ pour tout $n \in \mathbf{N}$; voir III.2 ci-dessus) : on pose d'abord $P_\nu(\mathbf{q}_n \in A) = (\pi_n \nu)(A) = \int_X (\pi_n(x))(A) \nu(dx)$, où $\pi_n = \overbrace{\pi \ldots \pi}^{n \text{ fois}}$; on prolonge ensuite P_ν aux pavés mesurables (i.e aux intersections finies de $\overset{-1}{\mathbf{q}_{n_i}}(A_i)$) en posant $P_\nu(\mathbf{q}_0 \in A_0$ et $\mathbf{q}_{n_1} \in A_1$ et $\ldots$ $\mathbf{q}_{n_k} \in A_k)$ $= \int_{A_0} \nu(dx_0) \int_{A_1} \pi_{n_1}(x_0, dx_1) \int_{A_2} \pi_{n_2 - n_1}(x_1, dx_2) \cdots \int_{A_k} \pi_{n_k - n_{k-1}}(x_{k-1}, dx_k)$; un théorème de prolongement nous permet alors de prolonger P_ν à tous les mesurables de $X^{\mathbf{N}}$. Cette définition de P_ν est ad hoc pour que $\pi_n \nu$ soit la P_ν-loi de $\mathbf{q}_n$ pour tout $n \in \mathbf{N}$, i.e la loi de la chaîne $(\mathbf{q}_n)$ au temps n, pour la loi initiale ν. En particulier, $\pi_0 \nu = \nu$ (voir III.5 ci-dessus) est la loi initiale de la chaîne $(\mathbf{q}_n)$; et, en notant P_x la probabilité P_{δ_x}, $P_x(\mathbf{q}_n \in A) = (\pi_n \delta_x)(A) = (\pi_n(x))(A)$, i.e $\pi_n(x)$ est la loi de la chaîne $(\mathbf{q}_n)$ au temps n, quand elle part de x.

App.III.7. Un *Processus de Markov homogène* (ou *processus Markovien homogène*) est un $\mathbf{R}_+$-automate stochastique (le monoïde $\mathbf{R}_+$ n'est ni libre ni finiment engendré ; on le munit de l'addition usuelle et de la tribu borélienne) ; c'est donc la donnée d'un espace mesurable X et d'une probabilité de transition $\theta : \mathbf{R}_+ \times X \longrightarrow \mathbb{P}X$ qui vérifie les axiomes (en notant $\pi_t(x)$ la probabilité $\theta(t, x)$

pour tout $t \in \mathbf{R}_+$ et tout $x \in X$) :

$$\pi_0(x) = \delta_x \quad \text{et} \quad \pi_{s+t} = \pi_s \pi_t$$

pour tout $s, t \in \mathbf{R}_+$ et tout $x \in X$ (attention, ici, a priori, la mesurabilité de tous les $\pi_t : X \longrightarrow \mathbb{P}X$ n'est pas une condition suffisante pour que $\theta : \mathbf{R}_+ \times X \longrightarrow \mathbb{P}X$ soit mesurable ; d'ailleurs, ici, on obtient plutôt les *processus Markoviens homogènes mesurables*, la mesurabilité de θ étant utile dans certains problèmes spécifiques d'intégration). $(\pi_t(x))(A)$ est la probabilité d'aller, en un temps t, de x dans A ; $\pi_t(x) = \pi_t \delta_x$ est la *loi du processus au temps t quand il part de x* ; plus généralement, si ν est une probabilité sur X, $\pi_t \nu$ est la *loi du processus au temps t, pour la loi initiale* ν.

Le problème de la construction, a posteriori, d'une famille de variables aléatoires dont les lois sont du type des $\pi_t \nu$ ci-dessus, i.e données, a priori, par une probabilité ν sur X et une famille (π_t) de probabilités de transition $X \longrightarrow \mathbb{P}X$, indexée par $\mathbf{R}_+$, et vérifiant la relation de Chapman-Kolmogoroff $\pi_{t+s} = \pi_t \pi_s$) est ici plus délicat. En effet, se référant à III.6 ci-dessus, on considère l'espace produit $X^{\mathbf{R}_+}$ (muni de sa tribu produit, i.e de la plus petite tribu rendant mesurables les projections canoniques $\mathbf{q}_t : X^{\mathbf{R}_+} \longrightarrow X$). Bien que l'on puisse toujours poser $P_\nu(\mathbf{q}_t \in A) = (\pi_t \nu)(A)$ pour tout $t \in \mathbf{R}_+$ et pour toute partie mesurable A de X (et donc définir P_ν sur les pavés mesurables de l'espace produit $X^{\mathbf{R}_+}$), on ne peut pas toujours prolonger P_ν en une probabilité sur tous les mesurables de $X^{\mathbf{R}_+}$; c'est cependant possible avec de bonnes hypothèses sur X (par exemple, si X est un espace de Borel, i.e isomorphe à un borélien de $[0, 1]$). Dans ce cas, la famille $(\mathbf{q}_t)$ forme un processus Markovien homogène dont les P_ν-lois sont, par définition de P_ν, les $\pi_t \nu$.

App.IV. La monade des probabilités

La monade $T_{\mathbb{P}} = (\mathbb{P}, \eta'', \mu'')$ des probabilités sur la catégorie **Mes** des applications mesurables entre espaces mesurables a été publiée pour la première fois par M.Giry [Giry 81][1] ; elle a donné une preuve complète mais succinte du fait que c'est une monade. Nous en donnons ici une preuve détaillée pour rendre ce travail plus accessible aux catégoriciens et aux probabilistes. Nous l'avons décrite rapidement dans la partie 4.3 de cet article. On reprend les notations fixées dans App.III.

App.IV.1. Rappelons que, pour définir une tribu sur l'ensemble $\mathbb{P}X$ des probabilités sur un espace mesurable X donné, nous avons utilisé les évaluations $e_A : \mathbb{P}X \longrightarrow [0, 1] : p \mapsto p(A)$ (où A est un partie mesurable de X (i.e $A \in \mathcal{X}$)) qui sont mesurables par définition même de cette tribu. En fait, pour toute application mesurable bornée $f : X \longrightarrow \mathbf{R}$, on considère l'évaluation $e_f : \mathbb{P}X \longrightarrow \mathbf{R}$

[1] Ollivier de la Tullaye [Tull 71] avait fait cette construction antérieurement dans un manuscrit récemment retrouvé.

définie par $e_f(p) = \int_X f dp$. Notons, pour tout espace mesurable X, $\mathcal{M}es(X)$ (resp. $\mathcal{M}es_b(X)$) l'espace vectoriel des applications mesurables (resp. mesurables bornées) de X dans $\mathbf{R}$, où $\mathbf{R}$ (tout comme $[0,1]$ ci-dessus) est muni de sa tribu borélienne.

Lemme 1 : *L'application* $\mathcal{X} \longrightarrow \mathcal{M}es_b(\mathbb{P}X) : A \longrightarrow e_A$ *se prolonge en une application* $Eval : \mathcal{M}es_b(X) \longrightarrow \mathcal{M}es(\mathbb{P}X) : f \longrightarrow e_f$ *qui est linéaire et qui commute aux* $\sup_n^{\uparrow}$ *(i.e aux* sup *dénombrables et croissants)* de $\mathcal{M}es_b(X)$.

<u>*Preuve*</u> : Par définition de l'application e_f, on a $e_{1_A} = e_A$ pour tout $A \in \mathcal{X}$. En utilisant la linéarité de l'intégrale et le théorème de convergence monotone, on obtient, selon que $f \in \mathcal{M}es_b(X)$ s'écrit $\sum_i \alpha_i f_i$ (où les α_i sont des réels et les f_i des applications mesurables bornées sur X ; il s'agit bien sûr de sommes finies ici) ou $f = \sup_n^{\uparrow} f_n$ (où les f_n forment une suite croissante d'applications mesurables bornées sur X) :

- d'une part $e_f(p) = \int_X \sum_i \alpha_i f_i dp = \sum_i \alpha_i \int_X f_i dp = \sum_i \alpha_i e_{f_i}(p)$ pour tout $p \in \mathbb{P}X$, i.e $e_{\sum_i \alpha_i f_i} = \sum_i \alpha_i e_{f_i}$;
- d'autre part $e_f(p) = \int_X \sup_n^{\uparrow} f_n dp = \sup_n^{\uparrow} \int_X f_n dp = \sup_n^{\uparrow} e_{f_n}(p)$ pour tout $p \in \mathbb{P}X$, i.e $e_{\sup_n^{\uparrow} f_n} = \sup_n^{\uparrow} e_{f_n}$.

Il en résulte immédiatement que l'application $Eval$ proposée est linéaire et commute aux $\sup_n^{\uparrow}$. On en déduit aussi que l'application e_f est mesurable lorsque, d'abord f est étagée sur X (en se limitant à des f_i de la forme 1_{A_i}, on obtient $e_{\sum_i \alpha_i 1_{A_i}} = \sum_i \alpha_i e_{A_i}$, de sorte que $e_{\sum_i \alpha_i 1_{A_i}}$ est mesurable) ; puis lorsque f est une fonction mesurable bornée positive sur X (elle s'écrit alors $f = \sup_n^{\uparrow} f_n$, où les f_n forment une suite croissante d'applications étagées sur X, de sorte que $e_{\sup_n^{\uparrow} f_n} = \sup_n^{\uparrow} e_{f_n}$ est mesurable) ; puis enfin lorsque f est simplement mesurable bornée (on écrit $f = f^+ - f^-$ avec f^+ et f^- mesurables bornées positives, de sorte que $e_f = e_{f^+} - e_{f^-}$ est mesurable). $\square$

App.IV.2. L'endofoncteur $\mathbb{P} : \mathbf{Mes} \longrightarrow \mathbf{Mes}$ est défini ainsi : à tout espace mesurable X, il associe l'ensemble $\mathbb{P}X$ des probabilités sur X, muni de la plus petite tribu rendant mesurables les évaluations e_A, lorsque A parcourt la tribu $\mathcal{X}$ sur X ; à toute application mesurable $f : X \longrightarrow Y$, il associe l'application $f_* : \mathbb{P}X \longrightarrow \mathbb{P}Y : p \mapsto f_*(p)$, (voir App.III.2). La fonctorialité de $\mathbb{P}$ résulte de l'égalité $(g \circ f)_* = g_* \circ f_*$. Pour vérifier que $\mathbb{P}$ est bien un endofoncteur sur **Mes**, Il faut juste s'assurer que, pour toute application mesurable f, l'application f_* est mesurable, i.e, par définition même de la tribu sur $\mathbb{P}Y$, que, pour tout $B \in \mathcal{Y}$, l'application $e_B \circ f_*$ est mesurable. Mais ceci est immédiat, puisque, pour toute probabilité p sur X, on a $(e_B \circ f_*)(p) = e_B(f_*(p)) = (f_*(p))(B) = p(\overset{-1}{f}(B)) = e_A(p)$, i.e $e_B \circ f_* = e_A$, où $A = \overset{-1}{f}(B)$; or e_A est mesurable puisque A l'est (car B et f sont supposés mesurables).

Remarquons en passant que, pour tout $A \in \mathcal{X}$, on a $\mathbb{P}1_A = (1_A)_* = e_A$; en effet, a priori $(1_A)_* : \mathbb{P}X \longrightarrow \mathbb{P}\mathbf{2} : p \mapsto (1_A)_*(p)$, mais en identifiant $\mathbb{P}\mathbf{2}$ à $[0,1]$ (toute probabilité $p \in \mathbb{P}\mathbf{2}$ s'identifiant à sa valeur $p(\{1\}) \in [0,1]$ qui la détermine entièrement), $(1_A)_*(p)$ s'identifie à $((1_A)_*(p))(\{1\}) = p(\overset{-1}{1_A}(\{1\})) = p(A) = e_A(p)$. On peut alors retrouver que $e_B \circ f_*$ est mesurable en écrivant $e_B \circ f_* = (1_B)_* \circ f_* = (1_B \circ f)_* = (1_A)_* = e_A$, où $A = \overset{-1}{f}(B)$.

App.IV.3. L'unité de la monade des probabilités est définie par $\eta''X : X \longrightarrow \mathbb{P}X : x \mapsto \delta_x$; La naturalité de η'' résulte du fait que, pour tout $x \in X$ et tout $B \in \mathcal{Y}$, on a $(f_*(\delta_x))(B) = \delta_x(\overset{-1}{f}(B)) = 1$ ou 0 (selon que $x \in \overset{-1}{f}(B)$ ou non) $= 1$ ou 0 (selon que $f(x) \in B$ ou non) $= \delta_{f(x)}(B)$, i.e $f_*(\delta_x) = \delta_{f(x)}$.

La mesurabilité de $\eta''X$ résulte de celle des $e_A \circ \eta''X$, lorsque A parcourt $\mathcal{X}$. Or, pour tout $x \in X$, $(e_A \circ \eta''X)(x) = e_A(\eta''X(x)) = e_A(\delta_x) = \delta_x(A) = 1_A(x)$, i.e $e_A \circ \eta''X = 1_A$ qui est mesurable, puisque A l'est.

$$\begin{tikzcd}
\mathbb{P}X \arrow[r, "f_*"] & \mathbb{P}Y \\
X \arrow[u, "\eta''X"] \arrow[r, "f"'] & Y \arrow[u, "\eta''Y"']
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbb{P}\mathbb{P}X \arrow[r, "(f_*)_*"] \arrow[d, "\mu''X"'] & \mathbb{P}\mathbb{P}Y \arrow[d, "\mu''Y"] \\
\mathbb{P}X \arrow[r, "f_*"'] & \mathbb{P}Y
\end{tikzcd}$$

App.IV.4. La multiplication de la monade des probabilités est ainsi définie : $\mu''X : \mathbb{P}\mathbb{P}X \longrightarrow \mathbb{P}X$ est l'application qui, à toute probabilité $\hat{p}$ sur $\mathbb{P}X$, associe la probabilité $\mu''X(\hat{p})$ sur X qui, à toute partie mesurable A de X, associe $(\mu''X(\hat{p}))(A) = \int_{\mathbb{P}X} e_A d\hat{p} = \int_{\mathbb{P}X} e_A(p)\hat{p}(dp) = \int_{\mathbb{P}X} p(A)\hat{p}(dp)$.

Vérifions que, pour tout $\hat{p} \in \mathbb{P}\mathbb{P}X$, on a bien $\mu''X(\hat{p}) \in \mathbb{P}X$. Soit donc $\hat{p}$ une probabilité sur $\mathbb{P}X$; d'abord $\mu''X(\hat{p})$ est positive (car e_A l'est) et $(\mu''X(\hat{p}))(X) = 1$ (car $e_X(p) = p(X) = 1$ pour tout $p \in \mathbb{P}X$). Soit maintenant (A_n) une suite de parties mesurables de X, 2 à 2 disjointes et telles que $\bigcup A_n$ soit aussi mesurable ; alors, utilisant le lemme 1 ci-dessus et le fait que, pour tout $n \in \mathbf{N}$, on a $1_{\bigcup_{k=0}^n A_k} = \sum_{k=0}^n 1_{A_k}$, on peut écrire $e_{\bigcup A_n} = e_{\sup_n^\uparrow \bigcup_{k=0}^n A_k} = \sup_n^\uparrow e_{\bigcup_{k=0}^n A_k} = \sup_n^\uparrow e_{\sum_{k=0}^n 1_{A_k}} = \sup_n^\uparrow \sum_{k=0}^n e_{A_k}$, de sorte que $(\mu''X(\hat{p}))(\bigcup A_n) = \int_{\mathbb{P}X} e_{\bigcup A_n} d\hat{p} = \int_{\mathbb{P}X} \sup_n^\uparrow \sum_{k=0}^n e_{A_k} d\hat{p} = \sup_n^\uparrow \sum_{k=0}^n \int_{\mathbb{P}X} e_{A_k} d\hat{p} = \sum_{k=0}^\infty (\mu''X(\hat{p}))(A_k)$, ce qui prouve la σ-additivité de $\mu''X(\hat{p})$.

Pour prouver la naturalité de μ'', on doit vérifier l'égalité $f_* \circ \mu''X = \mu''Y \circ (f_*)_*$ pour toute application mesurable $f : X \longrightarrow Y$. Soit donc $\hat{p} \in \mathbb{P}\mathbb{P}X$ et $B \in \mathcal{Y}$; alors $((f_* \circ \mu''X)(\hat{p}))(B) = (\mu''X(\hat{p}))(\overset{-1}{f}(B)) = \int_{\mathbb{P}X} e_A d\hat{p}$ où $A = \overset{-1}{f}(B)$, et $((\mu''Y \circ (f_*)_*)(\hat{p}))(B) = \int_{\mathbb{P}Y} e_B d(f_*)_*(\hat{p}) = \int_{\mathbb{P}X} e_B \circ f_* d\hat{p}$; il suffit donc d'utiliser l'égalité $e_B \circ f_* = e_A$ établie ci-dessus dans IV.2.

Reste donc à prouver la mesurabilité de $\mu''X$ pour tout espace mesurable X ; par définition de la tribu sur $\mathbb{P}X$, cela revient à la mesurabilité des $e_A \circ \mu''X$, lorsque A parcourt la tribu $\mathcal{X}$ donnée sur X. Or, pour tout $\hat{p} \in \mathbb{P}\mathbb{P}X$, on a $(e_A \circ \mu''X)(\hat{p}) = e_A(\mu''X(\hat{p})) = (\mu''X(\hat{p}))(A) = \int_{\mathbb{P}X} e_A d\hat{p} = e_{e_A}(\hat{p})$ (voir IV.1 ci-dessus pour la définition des e_f), de sorte que $e_A \circ \mu''X = e_{e_A}$; ceci prouve la mesurabilité de $e_A \circ \mu''X$, puisque e_A est mesurable (par définition de la tribu sur $\mathbb{P}X$) et bornée (voir le lemme 1).

Avant de quitter ce paragraphe, établissons le lemme suivant dont on se servira dans IV.5 :

Lemme 2 : *Pour tout* $f \in \mathcal{M}es_b(X)$ *et tout* $\hat{p} \in \mathbb{P}\mathbb{P}X$*, l'application* e_f *est* $\hat{p}$*-intégrable et l'on a l'égalité* :

$$\int_X f d(\mu'' X(\hat{p})) \stackrel{*}{=} \int_{\mathbb{P}X} e_f d\hat{p}$$

<u>*Preuve*</u> : Par le lemme 1, on sait que l'application $e_f : \mathbb{P}X \longrightarrow \mathbf{R}$ est mesurable. De plus, par définition même de la probabilité $\mu'' X(\hat{p})$, l'égalité $\stackrel{*}{=}$ est vraie lorsque $f = 1_A$, où $A \in \mathcal{X}$ (puisque $e_{1_A} = e_A$). Il en résulte immédiatement, par linéarité de l'intégrale, que l'égalité $\stackrel{*}{=}$ est vraie lorsque f est étagée. Elle est donc encore vraie lorsque f est mesurable bornée positive, puisqu'elle s'écrit alors $f = \sup_n^{\uparrow} f_n$, où les f_n sont étagées (on utilise le théorème de convergence monotone). Le passage au cas où f est seulement mesurable bornée n'est qu'une formalité en écrivant $f = f^+ - f^-$ (on a bien sûr exploité les propriétés de l'application $Eval$ du lemme 1). D'où le résultat. □

App.IV.5. Prouvons maintenant que $T_{\mathbb{P}} = (\mathbb{P}, \eta'', \mu'')$ est bien une monade.

- L'égalité $\mu'' X \circ \eta'' \mathbb{P}X = id_{\mathbb{P}X}$ résulte du fait que, pour tout $p \in \mathbb{P}X$, on a $\mu'' X(\eta'' \mathbb{P}X(p)) = p$; en effet, pour tout $A \in \mathcal{X}$, on a $(\mu'' X(\eta'' \mathbb{P}X(p)))(A) = (\mu'' X(\delta_p))(A) = \int_{\mathbb{P}X} e_A d\delta_p = e_A(p) = p(A)$.

- L'égalité $\mu'' X \circ \mathbb{P}\eta'' X = id_{\mathbb{P}X}$ résulte du fait que, pour tout $p \in \mathbb{P}X$, on a $\mu'' X((\eta'' X)_*(p)) = p$; en effet, pour tout $A \in \mathcal{X}$, on a $(\mu'' X((\eta'' X)_*(p)))(A) = \int_{\mathbb{P}X} e_A d(\eta'' X)_*(p) = \int_X e_A \circ \eta'' X dp = p(A)$, puisque $e_A \circ \eta'' X = 1_A$ (voir IV.3 ci-dessus).

- L'égalité $\mu'' X \circ \mu'' \mathbb{P}X = \mu'' X \circ \mathbb{P}\mu'' X$ résulte du fait que, pour tout $\hat{\hat{p}} \in \mathbb{P}\mathbb{P}\mathbb{P}X$, on a $\mu'' X(\mu'' \mathbb{P}X(\hat{\hat{p}})) = \mu'' X((\mu'' X)_*(\hat{\hat{p}}))$; en effet, pour tout $A \in \mathcal{X}$, on a :

- d'une part $(\mu'' X(\mu'' \mathbb{P}X(\hat{\hat{p}})))(A) = \int_{\mathbb{P}X} e_A d(\mu'' \mathbb{P}X(\hat{\hat{p}})) \stackrel{*}{=} \int_{\mathbb{P}\mathbb{P}X} e_{e_A} d\hat{\hat{p}} = \int_{\mathbb{P}\mathbb{P}X} e_{e_A}(\hat{p})\hat{\hat{p}}(d\hat{p}) = \int_{\mathbb{P}\mathbb{P}X}(\int_{\mathbb{P}X} e_A d\hat{p})\hat{\hat{p}}(d\hat{p})$ (la dernière égalité résultant de la définition des e_f, et l'égalité $\stackrel{*}{=}$ du lemme 2 ci-dessus) ;

- d'autre part $(\mu'' X((\mu'' X)_*(\hat{\hat{p}})))(A) = \int_{\mathbb{P}X} e_A d(\mu'' X)_*(\hat{\hat{p}}) = \int_{\mathbb{P}\mathbb{P}X} e_A \circ \mu'' X d\hat{\hat{p}} = \int_{\mathbb{P}\mathbb{P}X} e_A(\mu'' X(\hat{p}))\hat{\hat{p}}(d\hat{p}) = \int_{\mathbb{P}\mathbb{P}X}(\mu'' X(\hat{p}))(A)\hat{\hat{p}}(d\hat{p}) = \int_{\mathbb{P}\mathbb{P}X}(\int_{\mathbb{P}X} e_A d\hat{p})\hat{\hat{p}}(d\hat{p})$.

App.V. La loi distributive $\tau : T_M / T_{\mathbb{P}}$

Elle est définie par les $\tau X : M \times \mathbb{P}X \longrightarrow \mathbb{P}(M \times X) : (a, p) \mapsto \delta_a \otimes p$, pour tout espace mesurable X.

App.V.1. Pour la naturalité de τ, on vérifie la commutativité du diagramme suivant pour toute application mesurable $f : X \longrightarrow Y$:

$$\begin{tikzcd} M \times \mathbb{P}Y \arrow[r, "\tau Y"] & \mathbb{P}(M \times Y) \\ M \times \mathbb{P}X \arrow[u, "M\times f_*"] \arrow[r, "\tau X"'] & \mathbb{P}(M \times X) \arrow[u, "(M\times f)_*"'] \end{tikzcd}$$

On doit donc vérifier que, pour tout $(a,p) \in M \times \mathbb{P}X$, on a $\delta_a \otimes f_*(p) = (M \times f)_*(\delta_a \otimes p)$. Or, pour tout pavé mesurable $A \times B$ de $M \times Y$, on a, d'une part $(\delta_a \otimes f_*(p))(A \times B) = \delta_a(A)(f_*(p))(B) = \delta_a(A)p(\overset{-1}{f}(B))$, et, d'autre part $((M \times f)_*(\delta_a \otimes p))(A \times B) = (\delta_a \otimes p)((M \overset{-1}{\times} f)(A \times B)) = (\delta_a \otimes p)(A \times \overset{-1}{f}(B)) = \delta_a(A)p(\overset{-1}{f}(B))$.

La mesurabilité de τX résulte de celle des $e_{A\times B} \circ \tau X$, pour tout pavé mesurable $A \times B$ de $M \times X$. Or, pour tout $(a,p) \in M \times \mathbb{P}X$, on a $(e_{A\times B} \circ \tau X)(a,p) = e_{A\times B}(\delta_a \otimes p) = (\delta_a \otimes p)(A \times B) = \delta_a(A)p(B) = 1_A(a)e_B(p)$; la mesurabilité de $e_{A\times B} \circ \tau X$ résulte donc directement de celle de A et de B.

App.V.2. Reste à vérifier les axiomes (B_i) (voir App.I.5).

(B_1) : L'égalité $\tau X \circ (M \times \eta'' X) = \eta''(M \times X)$ résulte du fait que, pour tout $(a,x) \in M \times X$, on a $\delta_a \otimes \delta_x = \delta_{(a,x)}$.

(B_2) : L'égalité $\tau X \circ \eta \mathbb{P}X = \mathbb{P}\eta X$ résulte du fait que, pour toute probabilité p sur X et tout pavé mesurable $A \times B$ de $M \times X$, on a $((\eta X)_*(p))(A \times B) = p(\overset{-1}{\eta X}(A \times B)) = 1_A(e)p(B) = \delta_e(A)p(B) = (\delta_e \otimes p)(A \times B)$, la seconde égalité résultant du fait que $(\overset{-1}{\eta X})(A \times B) = \{x \in X \mid (e,x) \in A \times B\} = B$ ou $\emptyset$, selon que $e \in A$ ou non.

(B_3) : L'égalité $\tau X \circ (M \times \mu'' X) = \mu''(M \times X) \circ \mathbb{P}\tau X \circ \tau \mathbb{P}X$ résulte du fait que, pour tout $(a,\hat{p}) \in M \times \mathbb{P}\,\mathbb{P}X$ et tout pavé mesurable $A \times B$ de $M \times X$, on a, en posant $p = \mu'' X(\hat{p})$:

- d'une part $(\tau X((M \times \mu'' X)(a,\hat{p})))(A \times B) = (\delta_a \otimes p)(A \times B) = \delta_a(A)p(B)$;
- d'autre part $(\mu''(M \times X)((\tau X)_*(\tau(\mathbb{P}X)(a,\hat{p}))))(A \times B) = (\mu''(M \times X)((\tau X)_*(\delta_a \otimes \hat{p})))(A \times B) = \int_{\mathbb{P}(M\times X)} e_{A\times B} d(\tau X)_*(\delta_a \otimes \hat{p}) = \iint_{M\times\mathbb{P}X} e_{A\times B} \circ \tau X \, d(\delta_a \otimes \hat{p}) = \iint_{M\times\mathbb{P}X} (e_{A\times B}(\delta_b \otimes q))(\delta_a \otimes \hat{p})(dbdq) = \iint_{M\times\mathbb{P}X} ((\delta_b \otimes q)(A \times B))(\delta_a \otimes \hat{p})(dbdq) = \iint_{M\times\mathbb{P}X} \delta_b(A)q(B)(\delta_a \otimes \hat{p})(dbdq) = \int_{\mathbb{P}X} \delta_a(A)q(B)\hat{p}(dq) = \delta_a(A) \int_{\mathbb{P}X} q(B)\hat{p}(dq) = \delta_a(A) \int_{\mathbb{P}X} e_B(q)\hat{p}(dq) = \delta_a(A)p(B)$.

(B_4) : L'égalité $\tau X \circ \mu \mathbb{P}X = \mathbb{P}\mu X \circ \tau(M \times X) \circ (M \times \tau X)$ résulte du fait que, pour tout $(a,b,p) \in M \times M \times \mathbb{P}X$ et tout pavé mesurable $A \times B$ de $M \times X$, on a $((\mu X)_*(\tau(M \times X)((M \times \tau X)(a,b,p))))(A \times B) = ((\mu X)_*(\tau(M \times X)$ $(a, \delta_b \otimes p)))(A \times B) = ((\mu X)_*(\delta_a \otimes (\delta_b \otimes p)))(A \times B) = (\delta_{(a,b)} \otimes p)(\overset{-1}{\mu X}$ $(A \times B)) \overset{**}{=} \delta_{ab}(A)p(B) = (\delta_{ab} \otimes p)(A \times B) = (\tau X(\mu \mathbb{P}X(a,b,p)))(A \times B)$, l'égalité

$\stackrel{**}{=}$ résultant du fait que $\overset{-1}{\mu}X(A \times B) = \{(c,d,x) \in M \times M \times X \mid cd \in A$ et $x \in B\} = \{(c,d) \in M \times M \mid cd \in A\} \times B$.

Remerciements : Je tiens à remercier Jacques Penon pour les nombreuses discussions stimulantes que j'ai eues avec lui au cours de la rédaction de ce texte, ainsi que Philippe Bougerol pour m'avoir accordé du temps en répondant avec précision à toutes mes questions sur des points délicats de la théorie des probabilités.

Bibliographie

[Beck 69] J.Beck : Distributive laws, in Seminar on triples and categorical homology theory.119-140. LNM 80. Springer. 1969.

[Ben 67] J.Benabou : Introduction to bicategories, in Report of the Midwest category seminar. 1-77. LNM 47, Springer 1967.

[Bu 86] A.Burroni : Récursivité graphique (1ère partie) : catégorie des fonctions récursives formelles. *Cahiers de topologie et géométrie différentielle.* 49-79. Vol XXVII-1, 1986.

[Bu' 74] E.Burroni : Algèbres non déterministes et D-catégories. *Cahiers de topologie et géométrie différentielle.* 417-475. Vol XIV-4, 1973.

[Giry 81] M.Giry : A categorical approach to probability theory, in Categorical aspects of topology and analysis. 68-85, LNM 915. 1981.

[McL 72] S.Mac Lane : Categories for the working mathematician. GMT 5, Springer. 1972.

[Nev 64] J.Neveu : Calcul des probabilités. Masson. 1964.

[Star 72] P.H.Starke : Abstract automata. North-Holland. 1972.

[Tull 71] O.de la Tullaye : L'intégration considérée comme l'algèbre d'un triple. Rapport de Stage de D.E.A. Manuscrit. 1971.

Institut de Mathématiques de Jussieu, Université Paris 7, case 7012, 2, place Jussieu, 75251 Paris cedex 05 France.

E-mail address: eburroni@math.jussieu.fr