PROBABILITES, RELATIVISATIONS DESCRIPTIONNELLES, ET REPRESENTATION DES COMPLEXITES ET DE LEURS MESURES *SANS AMPUTATION DU SENS*

Mioara Mugur-Schächter [*]

Résumé. Dans ce travail on construit un algorithme d'identification de la loi *factuelle* de probabilité qu'il convient d'affirmer dans une 'situation probabiliste' donnée quelconque. Sur cette base on établit où la théorie des communications de messages de Shannon comporte du 'sens', et en quoi ce sens consiste. On montre en outre que ce sens intervient d'une manière cruciale tout au cours de la syntaxe de Shannon. Enfin, on esquisse une représentation relativisée des 'complexités' et l'on indique des manières de mesurer ces complexités sans amputer les contenus sémantiques impliqués.

Abstract. In this work an algorithm is constructed that permits to identify the *factual* probability law to be asserted in any given factual 'probabilistic situation'. On this basis it is established where Shannon's theory of communication of messages, contains 'significance', and of what this significance consists. It is furthermore shown that this significance plays a crucial role throughout Shannon's syntax. Finally a relativized representation of 'complexities' is outlined and modalities are indicated for measuring these complexities without amputating their semantic contents.

I. INTRODUCTION

Depuis quelques décennies l'essence et les aspects innombrables de ce qu'on appelle 'complexité' ne cessent de frapper les esprits avec force et ampleur croissantes. On voudrait *comprendre* les complexités. On voudrait en estimer la 'valeur' sans leur substituer des réductions arbitraires et dérisoires. Le concept d'entropie informationnelle d'une loi de probabilité, introduit par Shannon [1] et affiné mathématiquement par Khinchin [2], paraissait offrir quelques espoirs d'avancer vers ce but, même s'il restait encore très obscur en quel sens la complexité d'entités souvent *individuelles* et toujours censées exister *indépendamment de toute **communication***, pourrait systématiquement être évaluée en termes probabilistes et informationnels.

Mais actuellement cette voie est en impasse. Depuis déjà les années 1970 les mathématiciens s'enferment dans un refus pur et simple de l'entière théorie de l'information de Shannon. Ce refus est fondé sur l'idée que le concept formel d'une mesure de probabilité au sens moderne de Kolmogorov, serait dépourvu de toute interprétation factuelle définissable clairement. Il en découlerait que le calcul moderne des probabilités est à confiner au statut d'un domaine des mathématiques pures, logiquement coupé des situations 'probabilistes' au sens factuel. Cette vue, promue par Kolmogorov lui-même et par Chaitin, a engendré une scission dans l'évolution des recherches concernant la complexité : D'une part est née une *théorie de la complexité 'algorithmique'*. Celle-ci – sans plus faire aucun usage d'un concept de probabilité – construit des mesures formelles de la complexité des suites de signes utilisées dans les programmes informatiques. Ces mesures déconstruisent à tel point les contenus sémantiques des entités factuelles qui constituent l'objet des programmations, que parler de 'complexité' dans un tel contexte participe de l'ironie et du détournement. D'autre part, l'étude des systèmes, de

---

[*] Centre pour la Synthèse d'une Epistémologie Formalisée (*CeSEF*)

[1] Shannon, E.C., "The mathematical Theory of Communication", *Bell Syst., Techn. Journ.,* 27, 379-423 ; 623-656 (1948).

[2] Khinchin, A.I., "Mathematical Foundations of Information Theory", *Dover Publications* (1957) (traduction de deux articles parus en russe en 1953).

l'organisation, des démarches constructives au sens de Simon et de Le Moigne, de la complexité telle qu'elle est conçue par Morin – où l'accent tombe sur la *structure du sens* – s'étoffe de plus en plus. Mais sans s'extraire d'un seul pouce de l'exclusivement qualitatif.

Dans l'exposé qui suit je montrerai d'abord que la question cruciale de la signification factuelle associable au concept abstrait de mesure de probabilité, admet une solution. Ensuite j'esquisserai les principes d'une représentation non-réductrice et susceptible d'estimations numériques, du concept de complexité d'une entité quelconque. Il s'agira d'une approche inachevée. Mais cette approche est enracinée dans la microphysique moderne et incorporée dans une épistémologie formalisée où tout développement est guidé [3].

II. LE PROBLEME DE LA SIGNIFICATION D'UNE MESURE DE PROBABILITE

II.1. Les probabilités modernes 'classiques'

La théorie moderne 'classique' des probabilités a été formulée par Kolmogorov [4]. Elle est fondée sur le concept d'un 'espace de probabilité' $[U, \tau, p(\tau)]$ où : ***(a)*** $U\equiv\{e_i\}$ (avec $i\in I$ et I un ensemble d'indices)] est un *univers (ensemble) d'événements élémentaires* e_i engendré par la répétition d'une *procédure (ou expérience)* $\mathscr{P}$ qui, malgré l'assertion de sa reproductibilité 'identique', produit des événements élémentaires qui *varient* à l'intérieur de l'univers U ; ***(b)*** τ est une *algèbre d'événements* définie sur U [5], un *événement e* de τ consistant en un sous-ensemble de U et étant posé s'être 'réalisé' à chaque fois que s'est réalisé un événement élémentaire e_i quelconque appartenant au sous-ensemble considéré ; ***(c)*** $p(\tau)$ est une *mesure de probabilité* définie sur l'algèbre τ [6].

Une paire $[\mathscr{P},U]$ contenant une procédure $\mathscr{P}$ 'identiquement' reproductible et l'univers correspondant U d'événements élémentaires, constituent ensemble un *phénomène aléatoire*. Sur un univers U donné d'événements élémentaires on peut définir tout un ensemble d'algèbres τ d'événements. Donc il est possible de former différentes associations $[[\mathscr{P},U], [U, \tau, p(\tau)]]$ correspondant toutes au même phénomène aléatoire $[\mathscr{P},U]$.

Par comparaison avec les représentations mathématiques précédentes du concept de probabilité (Bernoulli, von Mises, etc.) où seulement une 'loi' de probabilité était définie en termes mathématiques, le concept de Kolmogorov d'un *espace* de probabilité $[U, \tau, p(\tau)]$ à marqué un

---

[3] L'exposé qui suit fait usage d'éléments tirés d'un autre exposé beaucoup plus élaboré, qui paraîtra ailleurs. Mais on peut trouver en annexe une version française complète de la méthode épistémologique mentionnée.

[4] Kolmogorov, A.N., "Foundations of the Theory of Probabilities", *Chelsea Publishing Company*, 1950. (Traduction de la monographie originale allemande "Grundbegriffe der Wahrscheinlichkeitrechnung", Ergebnisse der Mathematik (1933).

[5] Une algèbre définie sur un ensemble S est un ensemble de sous-ensembles de S – contenant S lui-même ainsi que l'ensemble vide $\emptyset$ – et qui est tel que si les sous-ensembles A et B sont dans S alors S contient également $A\cup B$ et $A\text{-}B$.

[6] Une mesure de probabilité définie sur τ consiste en un ensemble de nombres réels $p(A)$ dont chacun est associé à un événement A de τ et qui satisfont aux conditions suivantes : : $0\leq p(A)\leq 1$, $p(U)=1$ (normation), $p(\emptyset)=0$, et $p(A\cup B)\leq p(A)+p(B)$ où l'égalité se réalise ssi A et B sont disjoints (n'ont aucun événement élémentaire e_i en commun $(A\cap B=\emptyset)$. En fait, ***l'inclusion du cas limite p(A)=0 crée confusion*** (cf. III.2.5).

énorme progrès : *via* ce concept, les représentations des situations factuelles que l'on dénomme 'probabilistes' dans le langage courant – qui n'étaient caractérisées que sur une base intuitive, même si elle faisait usage de nombres – s'insèrent désormais dans la syntaxe mathématique très élaborée de la théorie des mesures sur des ensembles.

II.2. Sur l'interprétation d'une mesure abstraite de probabilité

L'unique élément spécifiquement 'probabiliste' d'un espace de probabilité est la mesure de probabilité $p(\tau)$. Or à ce jour même l'application du concept formel de mesure de probabilité, à des situations qui au plan factuel sont incontestablement 'probabilistes', *n'a pas encore pu être fondée sur une interprétation explicitement construite d'une mesure de probabilité*. On ne sait même pas désigner avec précision le *sens* de l'affirmation qu'en telle ou telle circonstance concrète il *'existe'* une *'loi de probabilité'* au sens empirique. *A fortiori* on ne sait pas indiquer une procédure pour construire cette loi. La spécification d'un tel sens-et-procédure corrélés – *même en principe seulement* – suffirait déjà pour asseoir un concept *factuel* de probabilité qui puisse être regardé comme une interprétation d'une mesure de probabilité formelle. Mais, et ceci est très surprenant, une telle spécification manque totalement. Dans tel ou tel cas donné où l'on constate de manière factuelle des dispersions statistiques *cependant qu'un ensemble de conditions globales* ***stables*** *est spécifié*, on se borne à juste affirmer, sur la base d'équipartitions *a priori* des événements élémentaires, qu'il existerait la loi de probabilité déterminée par le nombre des cas favorables à tel ou tel événement rapporté au nombre des cas possibles. Mais en général les équipartitions *a priori* des événements élémentaires ne sont pas confirmées par les mesures effectives des fréquences relatives de ceux-ci.

Ce problème reste encore confidentiel. Chez la plupart des physiciens, chez les spécialistes de la communication, chez les mathématiciens qui ne font qu'utiliser la théorie des probabilités sans la placer au cœur de leurs recherches, chez l'homme de la rue, il s'est installé une 'confiance' profane selon laquelle toute question qu'on peut se poser posséderait sûrement une réponse quelque part dans quelques travaux spécialisés. Des croyances de cette sorte, dures comme du granite, se forment autour de toute question scientifique. Ces croyances sont le terrain fragile mais nécessaire sur lequel roule l'évolution des sciences. Mais ceux qui développent des recherches concernant les fondements de la théorie des probabilités, sont tout à fait conscients, eux, que le concept de mesure de probabilité pose un problème d'interprétation dont l'importance est vitale. Kolmogorov lui-même [7] écrit dès 1963 [8] :

> « I have already expressed the view …that the basis for the applicability of the results of the mathematical theory of probability to real random phenomena must depend in some form on the *frequency concept of probability*, the unavoidable nature of which has been established by von Mises in a spirited manner…..(But) The frequency concept (of probability (ma spécification)) which has been based on the notion of limiting *frequency* as the number of trials increases to infinity, does not contribute anything to substantiate the applicability of the results of probability theory to real practical problems where we have always to deal with a finite number of trials ».

[7] Kolmogorov, N. A ., in *Shankhya*, 1963.
[8] Segal, J., "Théorie de l'information : science, techniques et société", *Thèse* (manuscrit), p. 587, note 783.

Cette citation mérite toute l'attention. On ne peut pas être plus clair. Explicitons sa signification. Il s'est constitué une notion plus ou moins floue, mais assez agissante, selon laquelle le théorème des grands nombres fonderait de manière *déductive* ***stricte*** l'existence d'une loi de probabilité factuelle, tout en spécifiant aussi sa structure. Or il n'en est rien. Le théorème des grands nombres, on le sait, affirme ce qui suit (j'utilise les notations ensemblistes traditionnelles).

Etant donné un ensemble $\{e_j,\ j\equiv 1,2,....q\}$ d'événements e_j [9], (ou d'événements élémentaires, indifféremment), ***SI*** une loi factuelle de probabilité $\{p(e_j),\ j\equiv 1,2,\ ....q\}$ ***existe*** sur cet ensemble, ***ALORS***, pour tout événement e_j et toute paire (ε,δ) de deux réels *arbitrairement petits*, il existe un entier N_0 tel que lorsque le nombre N des réitérations 'identiques' du phénomène aléatoire qui agit devient égal à, ou plus grand que N_0, la (méta) ***PROBABILITE***

$$\mathbf{P}[\,|(n(e_j)/N-p(e_j))]\leq\varepsilon] \qquad (1)$$

pour que **[**la valeur absolue de la différence $(n(e_j)/N - p(e_j))$ entre d'une part la fréquence relative $n(e_j)/N$ *mesurée* pour l'événement e_j et d'autre part la probabilité factuelle $p(e_j)$ de cette événement, devienne plus *petite* ou égale à ε **]**, devient, elle, plus *grande* que, ou égale à, $(1\text{-}\delta)$. On peut exprimer tout cela d'une manière synthétique et rigoureuse à l'aide d'une seule ligne bien connue d'écriture entièrement symbolique :

$$\forall j,\ \forall(\varepsilon,\delta), \quad \exists N_0:\ \forall(N\geq N_0) \Rightarrow \mathbf{P}\,[\,|(n(e_j)/N - p(e_j)\,|\,]\leq\varepsilon\,] \geq (1-\delta) \qquad (2)$$

Cette même assertion est quelquefois exprimée d'une façon moins précise en disant que *si* une loi de probabilité $\{p(e_j),\ j\equiv 1,2,\ ....q\}$ *existe* sur l'ensemble d'événements $\{e_j,\ j\equiv 1,2,\ ...,\ q\}$, *alors* pour tout j, lorsque N 'tend vers l'infini' la valeur absolue de la différence entre la fréquence relative $(n(e_j)/N$ et la probabilité $p(e_j)$, *tend vers 0 'en probabilité'*, i.e. *presque* certainement. ***Presque*** certainement, *pas* certainement, parce que l'expression $\mathbf{P}[\,|(n(e_j)/N - p(e_j)\,|] \leq \varepsilon\,]$ désigne elle même seulement une (méta)*probabilité* (indiquée ici par $\mathbf{P}$), non une certitude, d'un (méta)événement $[\,|(n(e_j)/N - p(e_j)\,|] \leq \varepsilon\,]$ ('méta' en ce sens que sa définition fait intervenir les événements e_j, et qu'en ce sens elle leur est conceptuellement postérieure).

Donc dans le théorème des grands nombres l'*existence* de la loi de probabilité $\{p(e_j),\ j\equiv 1,2....q\}$ n'est pas déduite, elle est ***posée***. Ce qu'on *montre* est que la tendance de l'évolution lorsque N s'accroît, de chaque fréquence relative $n(e_j)/N$, vers la probabilité $p(e_j)$ assignée à l'événement e_j par la loi qu'on pose exister, est, *elle*, très probable, *au sens d'une autre* ***loi de probabilité***. Donc *en ce qui concerne la signification de l'existence d'une loi de probabilité, on ne trouve là qu'une régression infinie*.

Quant à la *forme* de la loi de probabilité factuelle $\{p(e_j),\ j\equiv 1,2....q\}$, le théorème des grands nombres n'en construit une définition factuelle – la fameuse 'définition fréquentielle' exprimée à

[9] Afin de ne pas noyer l'essence du problème dans d'autres questions, je suppose constamment des ensembles *finis* d'événements.

l'aide de la (méta)probabilité ***P*** – que sur la base de la postulation de l'*existence* de cette loi. Les fréquences relatives ne jouent en *(2)* un rôle de *détermination* des valeurs-limite $p(e_j)$, que si elles sont *soumises* à l'existence postulée de valeurs-limite $p(e_j)$.

C'est vers tout cela que pointe le texte de Kolmogorov cité plus haut. Et dès qu'on y pense vraiment il saute aux yeux qu'en effet il est gênant de relier un système formel comme la théorie mathématique des probabilités, à des situations probabilistes factuelles, en l'absence de tout modèle du concept de loi factuelle de probabilité.

> ***Ce qui manque est un modèle FACTUEL du concept de loi de probabilité****. La définition abstraite d'une mesure de probabilité doit être la* **formalisation** *d'un tel modèle factuel défini. Elle ne peut pas en être le* **générateur***. les 'situations probabilistes' appartiennent à la factualité physique.*

Déjà bien avant Kolmogorov tout un nombre d'auteurs ont manifesté des réserves induites par cette situation. Par exemple R. J. Solomonoff écrit [10] :

> « Probability theory tells how to derive a new probability distribution from old probability distributions…….. It does *not* tell how to get a probability distribution from data in the real world ».

Mais c'est Kolmogorov qui a développé dernièrement une attitude extrême à cet égard. Dans les années 1980 il a opposé un net refus au célèbre concept de Shannon d'entropie informationnelle (une forme entropique $H(S) \equiv \Sigma_i p_i log(1/p_i)$, mais qui est associée non pas à une statistique de fréquences relatives $\{n_i/N,\ i \equiv 1,2,...q\}$ comme dans la théorie de Boltzmann, mais à une mesure de probabilité $\{p_i,\ i \equiv 1,2,\ldots q\}$ supposée agir *factuellement* sur l'alphabet $\{a_i,\ i \equiv 1,2,...q\}$ de signes a_i à coder et à transmettre, émis par une source S d''information'). Kolmogorov est allé jusqu'à prôner *l'élimination du concept formel de probabilité de la base de toutes les représentations considérées jusqu'ici comme des applications de ce concept* [11] :

> « **1**. Information theory must precede probability theory, and not be based on it. By the very essence of this discipline, the foundations of information theory have a finite combinatorial character.
> **2**. The applications of probability theory can be put on a uniform basis. It is always a matter of consequences of hypotheses about the impossibility of reducing in one way or another the complexity of the descriptions of the objects in question. Naturally this approach to the matter does not prevent the development of probability theory as a branch of mathematics being a special case of general measure theory.
> **3**. The concepts of information theory as applied to infinite sequences give rise to very interesting investigations, which, without being indispensable as a basis of probability theory, can acquire a certain value in the investigation of the algorithmic side of mathematics as a whole ».

Le père des probabilités modernes voulait donc traiter désormais les questions d'information et de complexité sans recours au concept de probabilité. Il voulait les traiter à l'aide seulement d'analyses combinatoires et d' « hypothèses d'impossibilité de réduction d'une manière ou d'une

---

[10] Solomonoff, R. J., "An inductive inference machine", *IRE National Record*, 5, 1957 (cf. réf. 90 p. 594).

[11] Kolmogorov, A., "Combinatorial foundations of information theory and the calculus of probabilities", *Russia Mathematical Surveys*, 38, pp. 29-40, (1983) (cf. réf. 90).

autre, de la complexité des ***descriptions*** [12] des objets en question ». Quant aux probabilités, il voulait les confiner à une branche de la théorie mathématique des mesures, sans droit d'application directe. Il voulait enfermer dans une cage abstraite le concept de probabilité *qui a été dicté par l'expérience concrète* !!! C'est une proposition qui vient d'une autorité dans la matière et donc elle mérite d'être gardée présente à l'esprit. Mais c'est une proposition extrémiste. Parmi les mathématiciens cette proposition a pu d'ores et déjà infléchir la direction des recherches concernant la complexité. Pour un physicien cependant, il est tout simplement inconcevable qu'un concept formel comme celui de mesure de probabilité – dont l'origine première est factuelle – ne pointe pas en retour vers un sens factuel, constructible d'une manière explicite. *Dans l'entière microphysique, les descriptions probabilistes sont* **primordiales** en ce sens qu'elles y émergent dans la toute première phase descriptionnelle. En physique macroscopique les probabilités sont conçues comme l'effet de l'ignorance de 'détails' qui en principe sont connaissables et que la théorie primordiale *non*-probabiliste serait apte en principe à traiter de façon à obtenir des résultats certains. Mais en dessous des descriptions probabilistes des microétats il n'existe pas des données connaissables extérieures aux lois de probabilité qui constituent ces descriptions *directement* ; en microphysique on ne dispose d'aucune théorie non-probabiliste qui puisse en principe offrir des descriptions certaines. Si l'on voudra absolument disposer d'une telle théorie, il faudra un jour la construire ***sur la base*** *de la théorie primordiale probabiliste*. Il paraît absurde de concevoir que la signification factuelle du concept de loi physique de probabilité ne soit pas constructible, quand ce concept intervient ainsi avec ce statut primordial à la *base* de l'entière physique.

Alors comment expliquer la situation esquissée plus haut ? J'ai montré ailleurs [13] que la formalisation de Kolmogorov, malgré les remarquables enrichissements qu'elle a apportés à l'état précédent du concept de probabilité, n'incorpore *pas* les caractéristiques *générales* d'une 'situation probabiliste' quelconque. La production d'un exemple suffit pour le démontrer. Or l'exemple qu'on peut évoquer est indiscutable et énorme : les descriptions probabilistes des microétats ***dépassent*** la théorie classique des probabilités ; celle-ci est débordée par l'entière physique des microétats. En ces conditions, certaines données nécessaires pour identifier dans sa globalité la 'signification' de la notion de loi de probabilité, sont restées cachées. Mais elles peuvent être discernées via une analyse appropriée, plus exhaustive. Celle-ci se trouve développée à l'intérieur d'une *méthode générale de conceptualisation relativisée* (*MCR*) dont la source se trouve précisément dans l'étude des probabilités et de la logique impliquées dans le formalisme quantique (cf. les indications bibliographiques 15).

---

[12] ***Notons bien cette distinction entre description d'un objet et l'objet décrit : elle se montrera essentiellement consonante avec les conclusions qui émergeront ici***.

[13] Mugur Schächter, M., :

A. "Spacetime Quantum Probabilities I:..... ", *Founds. of Phys*., Vol. 21, (1991) ;

B. "Toward a Factually Induced Space-Time Quantum Logic", *Founds. of Phys*., Vol. 22, (1992) ;

C. "From Quantum Mechanics to Universal Structure of Conceptualization and Feedback on Quantum Mechanics", *Founds. of Phys*., Vol. 23, (1993) ;

D. "Quantum Probabilities, Operators of State Preparation, and the Principle of Superposition", *Int. J. of Theoretical Phys.,* Vol. 31, No.9, 1992.

III. ESQUISSE CONSTRUCTIVE D'UN MODELE CONCERNANT UNE LOI FACTUELLE DE PROBABILITE

III.1. Très bref préalable sur *MCR*

Une 'description', de par la définition du concept, comporte nécessairement une entité-objet de la description et une grille de qualification de celle-ci (une 'vue' à travers laquelle on qualifie, ou encore, un 'regard' qualifiant). La méthode de conceptualisation relativisée – *MCR* – introduit une expression canonique d'une description quelconque. Cette méthode a été exposée amplement dans chacune de ses phases [14]. Dans ce texte, pour auto-suffisance, je ne re-dirai que ce qui suit concernant *MCR* et le canon d'expression d'une description, qu'elle introduit. Toutefois il est clair qu'une véritable compréhension de ce qui suit ne peut être acquise qu'en examinant un exposé complet de la méthode, de préférence le dernier en date.

(1) *MCR* est fondée sur la relativisation systématique de toute description, à la triade $(G,œ_G,V)$ qui y est impliquée où : *'G'* dénote l'opération – physique, abstraite, mixte – par laquelle est introduite l'entité-objet de la description que l'on veut faire ; *'$œ_G$'* dénote l'entité-objet-de la description elle-même ; *'V'* dénote la 'vue', la grille de qualification introduite afin de qualifier.

(2) Toute description, par exigence de méthode, est à mettre sous la forme canonique $D/G,œ_G,V/$. On peut aussi dire – ce qui revient au même – que toute description est relative au *référentiel épistémique (G,V)* dans lequel elle est élaborée.

(3) Entre l'opération de génération *G* d'entité-objet et l'entité-objet qui en résulte, il est posé une relation de un-à-un : $G \leftrightarrow œ_G$. Ce postulat méthodologique est loin d'être une évidence. Mais dans le cas des descriptions de microétats il s'impose inéluctablement, et une analyse attentive montre qu'il s'impose également *en toute généralité* lorsqu'on veut édifier une méthode de décrire qui bannisse a priori toute insertion de fausses absolutisations.

(4) Toute vue *V* est dotée d'une ***structure*** prescrite strictement par la définition du concept : une vue *V* est un ensemble *fini* de *vues-aspect Vg* où *'g'* est un indice d'aspect ; un aspect *g* consiste en la donnée d'*une dimension sémantique de qualification* (couleur, poids, etc.) portant un ensemble *fini* de 'valeurs' *gk* de l'aspect *g* (pour couleur : rouge, bleu, etc.) portées toutes par la dimension sémantique posée. (Dans chaque ces particulier les indices généraux *'g'* et *'gk'* sont remplacés par des indices spécifiques de ce cas-là. La vue-aspect *Vg* est définie *si et seulement si* tous les objets (outils,

---

[14] Mugur-Schächter, M.,

A."Esquisse d'une représentation générale et formalisée des descriptions et le statut descriptionnel de la mécanique quantique", *Epistemological Letters*, Lausanne, cahier 36, (1984) ;

B. "Spacetime Quantum Probabilities II : Relativized Descriptions and Popperian Propensities", *Founds. of Phys.*, Vol. 22 (1992).

C. "Une méthode de conceptualisation relativisée... ", *Revue Int. de Systémique*, Vol. 9 (1995).

D. "Objectivity and Descriptional Relativities", *Foundations of Science* 7, 73-180 (2002) ;

E. "Quantum Mechanics versus a Method of Relativised Conceptualisation", in *Quantum Mechanics, Mathematics, Cognition and Action : Proposals for a Formalised Epistemology*, M. Mugur-Schächter and A. van der Merwe, eds., Kluwer Academic Publishers (2003).

Etc..

appareils) et toutes les modalités opérationnelles sont spécifiées, qui permettent d'affirmer que lors d'un 'examen' *via* Vg d'une entité-objet $œ_G$, a été obtenue telle ou telle valeur gk (une seule) de g. Une vue est un ***filtre*** – fini – de qualification : face aux aspects-et-valeurs-d'aspect qui n'y sont pas contenus, elle est *aveugle*, elle ne les perçoit pas [15]. Les qualifications d'espace et de temps s'accomplissent à l'aide de deux types très particuliers de *'vues-cadre' d'espace-temps V(ET)* (réductibles selon le cas à une vue d'espace ou de temps seulement).

(5) Etant donnée une paire *(G,Vg)*, ces deux opérateurs épistémiques peuvent *exister l'un relativement à l'autre*, ou non. Si un examen par Vg de l'entité-objet $œ_G$ introduite par G produit quelque valeur bien définie gk de g, il y a existence relative et la paire *(G,Vg)* constitue un référentiel épistémique. En ce cas, en appliquant la grille de qualification Vg, à l'effet $œ_G$ de l'opération de génération G, c'est à dire en effectuant une succession *[G.Vg]* des deux opérations épistémiques G et Vg, l'on obtient une description de l'entité-objet $œ_G$ via la grille de qualification consistant en la vue-aspect Vg, i.e. une description relativisée $D/G,œ_G,Vg/$. Dans le cas contraire il n'y a pas existence relative (ou mutuelle) de $œ_G$ et Vg – donc non plus de G et Vg – et alors l'appariement *(G,Vg)* est à éliminer parce qu'il ne peut pas conduire à l'élaboration d'*une description relativisée* $D/G,œ_G,Vg/$. Si la succession *[G.Vg]* conduit systématiquement à un et même résultat gk à chaque fois qu'elle est réalisée, la description correspondante $D/G,œ_G,Vg/$ est ***individuelle***. Dans le cas contraire elle est ***probabiliste***. Ces considérations s'étendent d'une façon évidente à une paire G,V où V est une vue à plusieurs vues-aspect Vg. En ce cas on parle de la possibilité, ou non, d'*une description relativisée* $D/G,œ_G,V/$. *Toute entité-objet de nature* ***physique*** *existe face à au moins une vue-aspect Vg distincte de toute vue-****cadre*** *d'espace-temps V(ET), mais elle n'existe relativement à aucune vue-cadre V(ET) considérée* ***seule****, séparément de tout autre aspect Vg différent de tout aspect-cadre ET : c'est* ***le principe-cadre d'espace-temps***. Le référentiel épistémique de la description relativisée d'une entité-objet physique incorpore donc nécessairement au moins une vue-aspect Vg distincte de toute vue-cadre *V(ET)* ; en outre, par convention, il incorpore également une vue-cadre d'espace-temps *V(ET)* face à laquelle existe au moins une *description* partielles $D/G,œ_G,Vg/$ de l'entité-objet $œ_G$ où la vue-aspect Vg – distincte de toute *V(ET)* – appartient à la vue V de $D/G,œ_G,V/$.

(6) Tout au long d'une *chaîne de conceptualisation*, les descriptions reliées dans cette chaîne s'organisent selon un ordre *hiérarchique* – la description d'ordre *1* (i.e ; la première description *de la chaîne*), la description d'ordre *2* qui est une méta-description face à celle d'ordre 1, description d'ordre *3* qui face à celle d'ordre *1* est une méta-méta-description et face à celle d'ordre 2 est une méta-description, etc..

(7) Le passage de l'une des descriptions d'une chaîne, à la suivante, est dicté par *le principe de séparation*, au sens qui suit. Par construction, toute description $D/G,œ_G,V/$ est une cellule *finie* de conceptualisation, accomplie dans le référentiel épistémique *(G,V)* donné, puisque l'effet de G est $œ_G$

[15] On mesure la différence avec le concept grammatical-logique de 'prédicat'.

exclusivement et que le nombre de qualifications disponibles *a priori* à l'aide de toute vue V, est fini. Lorsque les ressources qualificationnelles disponibles dans (G,V) sont épuisées, toute continuation du processus de conceptualisation exige le passage à un *nouveau* référentiel épistémique où une *autre* description, correspondante, différente de la description $D/G,œ_G,V/$ accomplie dans (G,V), est à élaborer ***séparément***.

(8) Selon *MCR* toute 'connaissance communicable' est *description*. ***Rien*** d'autre que des descriptions ne peut être connu d'une manière communicable, ni des 'faits' extérieurs aux psychismes qui ne sont pas décrits, ni des fais psychiques (émotions, sentiments, etc.) qui ne sont pas exprimés par quelque description plus ou moins explicite, verbale ou d'une autre nature. Notamment :

> *Lorsqu'on reconstruit à l'intérieur de MCR le concept de probabilité,* ***les événements élémentaires et les événements – au sens probabiliste – acquièrent la forme de DESCRIPTIONS relativisées*** : ***leur statut MCR n'est pas celui D'ENTITES-OBJET $œ_G$, mais de DESCRIPTIONS relativisées de telles entités-objet***.

Ceci – qui est à relier à la note 12 – se révélera hautement non-trivial et organisateur lorsqu'il s'agit, notamment, de représentations relativisées de 'complexités'.

Cela suffira pour indiquer maintenant schématiquement de quelle manière les relativisations descriptionnelles permettent d'associer un modèle au concept factuel de loi de probabilité. Ceux qui désireront placer le chapitre en cours, dans un contexte-*MCR* pleinement élaboré, pourront consulter dans la réf. 16E la reconstruction du concept général de probabilité conformément à toutes les exigences du noyau de *MCR*.

III.2. Esquisse d'un modèle concernant l'existence et la structure d'une loi factuelle de probabilité

L'investigation se construira par une suite d'exemples. De petite évidence en petite évidence il se constituera une nouveauté.

III.2.1. Préalable : un tableau morcelé. Soit un référentiel épistémique (G_T,V) où le générateur G_T sélectionne comme entité-objet l'image de la solution intégrée du puzzle d'un tableau T [16] d'un paysage, et V est une vue comportant une vue-aspect de couleur Vc munie d'un nombre fini de valeurs cr de couleur et une vue-aspect-*cadre* d'espace $V(E)$, et une vue-aspect $V\phi c$ de *forme de couleur* ; donc la vue globale introduite est $V\equiv Vc\cup V(E)\cup V\phi c$. La description correspondante au référentiel (G_T,V) est donc $D/G_T,T,Vc\cup V(E)\cup V\phi c/$.

Supposons que le puzzle de T contient 100 carrés et que chaque carré porte l'inscription d'une paire de coordonnées d'espace (x_k,y_h) où la coordonnée x_k est tirée d'un ensemble de 10 coordonnées

[16] Ne pas confondre ici *'T'* avec un indice de *temps* intervenant dans vue-cadre d'espace-temps ou de temps : dans ce qui suit, pour simplicité, l'analyse est essentiellement spatiale.

successives $\{x_k,\ k\equiv 1,2\ldots 10\}$ inscrites sur un axe d'espace ox, et la coordonnée y_h est tirée d'un ensemble de 10 coordonnées successives $\{y_h,\ h\equiv 1,2\ldots 10\}$ inscrites sur un axe d'espace oy de même origine $o(x_1,y_1)$ que l'axe ox, et perpendiculaire sur celui-ci. Sur la solution intégrée du puzzle les coordonnées (x_1,y_1) indiquent un coin du tableau T, et les coordonnées (x_{10},y_{10}) en marquent le coin diagonalement opposé. De cette façon le tableau T visible sur la solution du puzzle est couvert d'une grille plane de référence spatiale où les paires (x_k,y_h) définissent les 'valeurs' assignées à la vue-aspect-cadre d'espace $V(E)$. Considérons un référentiel épistémique 'local' (G_κ,V) où G_κ sélectionne *un* seul carré – dénotons-le $\kappa(x_k,y_h)$ – et V est le *même* que dans le référentiel global (G_T,V). Soit $D/G,\kappa(x_k,y_h),V/$ la description relativisée correspondant à (G_κ,V). Elle consiste en une 'forme de couleurs' qui couvre ce carré, construite avec la vue *III*.

Supposons que ***(a)*** les dimensions globales de T et la distance entre deux coordonnées x successives et deux coordonnés y successives, sont telles, qu'en général le carré $\kappa(x_k,y_h)$ est assez petit pour qu'il ne porte (en gros) qu'*une seule valeur de couleur* ; et que ***(b)*** une description partielle donnée $D/G,\kappa(x_k,y_h),Vc/$ de $\kappa(x_k,y_h)$ réalisée *sans* tenir compte de l'aspect de l'aspect $V\phi c$ de *forme*-de-couleur i.e. en regardant exclusivement la valeur de couleur sur $\kappa(x_k,y_h)$, se réalise (plus ou moins approximativement) sur *bien plus* d'un seul carré [17].

Distinguons maintenant par un indice $j\equiv 1,2,\ldots.q$ les descriptions partielles $D/G,\kappa(x_k,y_h),Vc/$ qui sont mutuellement différentes. Alors, en conséquence de *(a)* et *(b)*, on obtient un ensemble $\{Dj\}$, $j\equiv 1,2,\ldots.q$, de q descriptions partielles $D/G,\kappa(x_k,y_h),Vc/$ mutuellement distinctes où *q est beaucoup plus petit que 100*.

Mélangeons les carrés et versons-les dans une urne. A partir de ce point nous esquissons une suite de 'jeux' qui conduiront à l'interprétation proposée.

III.2.2. Jeu d'illustration du pouvoir de reconstruction de l'ordre spatial (ou d'espace-temps). Faisons les *100* tirages successifs possibles, en plaçant à chaque fois le carré $\kappa(x_k,y_h)$ tiré, à la place, sur la grille de référence spatiale, qui est désignée par les coordonnés (x_k,y_h). Alors, après exactement le *100ème* tirage, le tableau T se trouvera reconstitué *sans même avoir eu à regarde les indices j de valeurs-de-couleurs portés par les carrés*. On aura suivi exclusivement les indications de l'ordre spatial, fournies par les coordonnées (x_k,y_h). Rien d'infini ni rien d'aléatoire ne sera intervenu. Tout aura été *fini* et, en dépit du morcellement et du mélange des morceaux, tout aura aussi été ***certain***. La grille de référence spatiale possède un pouvoir d'organisation topologique qui est *indépendant de tout 'contenu', pourvu seulement que ce contenu – quel qu'il soit – porte des étiquetages d'emplacement spatial*. Cette remarque s'étend d'une manière évidente à un 'tableau évolutif' fragmenté et dont les fragments portent des étiquettes d'espace-*temps*.

[17] Notons que l'aspect *V(E)* reste actif mais que l'expression de son activité est absorbée dans les indices *k,h* de la notation $\kappa(x_k,y_h)$.

III.2.3. Puzzle à un seul exemplaire de T. Maintenant procédons autrement. Mélangeons les *100* carrés et versons-les dans l'urne comme au paragraphe *2*. Faisons ensuite les *100* tirages successifs possibles, mais *ignorons* les inscriptions (x_k,y_h) d'emplacement spatial et jouons cette fois au jeu de puzzle. Par contre, sur chaque carré, considérons cette fois la description globale correspondante, $D/G,\kappa(x_k,y_h),V/$, celle qui *inclut* l'aspect de *forme* (de couleur, dans l'espace physique). De nouveau, après le *100*^{ème} tirage le tableau *T* se trouvera reconstitué. Mais nous aurons eu à tâtonner pour placer chaque morceau, nous aurons fait des essais et des erreurs, jusqu'à ce que, guidés par les *structures* des formes-de-couleurs portées par chaque carré – que les indices *j* n'indiquent que partiellement – nous aurons identifié à chaque fois le 'bon' emplacement, avant de passer au tirage suivant. La structure de formes-de-couleurs portée par chaque carré aura été décisive *surtout par ses contenus proches des* ***bords*** *des carrés*, où elle détermine ***des cohérences de voisinage*** : une sorte d''attraction par continuité' entre les contenus sémantiques de formes-de-couleur observables en proximité des bords des morceaux. Cette fois le pouvoir indépendant d'organisation topologique des coordonnées spatiales toutes nues, aura été court-circuité et remplacé par ces différentes 'attractions par continuités de voisinage'. Et de nouveau rien d'infini ne sera intervenu, *ni d'aléatoire*, malgré l'existence d'essais et erreurs. Car ceux-ci, avec évidence, tiennent à des caractères de la situation définie qui sont d'une nature radicalement différente de celle d'une incertitude prévisionnelle au sens probabiliste. Cet exemple aussi, comme le précédent, peut se généraliser à un 'tableau *évolutif*' fragmenté, dont les étiquettes d'espace-*temps* des fragments sont court-circuitées. (Dans la recherche du coupable d'un crime, par exemple, on joue, en essence, à un puzzle généralisé d'espace-temps de cette sorte).

III.2.4. Puzzle à plusieurs exemplaires de T. Munissons-nous maintenant de *1000* exemplaires du même tableau *T* et procédons pour tous ces exemplaires au même mélange et mise dans l'urne que dans les cas précédents. Nous aurons donc dans l'urne *100.000* de carrés mélangés. Procédons maintenant comme au point *III.2.3*. Que se passera-t-il ? Au bout de *100.000* tirages nous aurons *certainement* reconstitué les *1000* exemplaires du tableau *T*, en général après beaucoup de tâtonnements, et pas nettement l'un *après* l'autre, mais d'une manière entremêlée qui ne sépare complètement tous les exemplaires qu'avant le tout dernier tirage. Aucun trait essentiellement nouveau n'apparaîtrait avec 10^n exemplaires où *n* est fini mais arbitrairement grand, quel que soit *n*. Et ce jeu aussi peut s'étendre à un ensemble de 'tableaux évolutifs'. De nouveau rien d'aléatoire ne sera intervenu, malgré l'existence d'essais et erreurs et malgré l'intrication accrue du processus. *Un jeu de puzzle, quelle que soit sa complication, de par sa nature ne comporte pas du hasard.*

L'attraction de continuité sémantique sur les bords des morceaux, comportée par l'aspect de forme d'espace tient le hasard à distance.

III.2.5. Un 'jeu de probabilité' avec un seul exemplaire de T. Comment naît alors du 'hasard probabiliste' ? Par un changement qui au premier abord paraîtra insignifiant, je ferai maintenant apparaître tout à coup tous les caractères introduits par une loi de probabilité : des suites arbitrairement longues d'événements élémentaires, des fréquences relatives correspondantes, de l'*aléatoire* : le changement d'apparence insignifiante que j'ai annoncé s'avérera avoir en fait été un véritable saut conceptuel.

Utilisons le même tableau *T*, en *un* seul exemplaire. Mais au lieu de procéder comme à l'un ou l'autre des points qui précèdent, jouons le 'jeu de probabilité' suivant. Mélangeons les *100* carrés et versons-les dans l'urne. Tirons ensuite un carré de l'urne. *Notons* l'étiquette *j* constatée, puis ***remettons le carré dans l'urne sans nous soucier, ni de la forme-de-couleur portée par le carré, ni de l'endroit spatial où, sur une grille de coordonnées d'espace*** (x_k,y_h)***, il conviendrait de placer ce carré si l'on voulait reconstruire le tableau : l'aspect*** $V\phi c$ ***de forme-de-couleur-dans-l'espace reste MUET, et donc aussi, a fortiori, le critère de continuité de voisinage d'espace des formes-de-couleur***. Mélangeons de nouveau les carrés dans l'urne et répétons la procédure autant de fois que nous voulons.

J'affirme que cette fois, en conséquence de la modification de procédure, nous sommes *en 'situation probabiliste'*. En effet, contrairement à ce qui se passait dans tous les cas qui précèdent, cette fois, ***avant*** *chaque tirage*, se trouve entièrement reconstitué un certain ensemble de conditions *invariantes* qui définissent une 'procédure *reproductible*' $\mathscr{P}$ et un phénomène aléatoire $(\mathscr{P},U)$ au sens courant de ces termes, avec $U \equiv \{Dj\}, j \equiv 1,2....q$. Qu'arrivera-t-il alors ? *Peut-on le prévoir ?*

Si le nombre *N* des tirages est beaucoup plus grand que *q*, on peut faire deux constatations *C1* et *C2* évidentes :

C1. Puisque l'entier contenu initial de l'urne est reconstitué après chaque tirage, toutes les valeurs $j \equiv 1,2,\ldots.q$ qui avaient été possibles lors du tirage précédent, sont possibles aussi lors du nouveau tirage qui suit. D'un tirage à l'autre aucune possibilité n'est 'consommée' irréversiblement, comme aux points précédents.

C2. Corrélativement, le contenu de l'urne ne s'épuise *jamais*. Rien ne met plus *fin* à la suite des résultats qu'on peut obtenir par des répétitions d'un tirage. Cette suite est de longueur arbitraire, donc indéfinie, donc elle peut s'accroître 'vers l'infini'.

Je passe maintenant à deux autres assertions qui ne sont pas ressenties comme des certitudes. La première est la réponse *R1* à la question *Q1* suivante : « Si l'on continue les tirages aussi longtemps que l'on veut, est-ce que *toutes* les *q* valeurs-de-couleur *j* finiront nécessairement par apparaître, ou pas ? ». La deuxième assertion est la réponse *R2* à une seconde question *Q2* qui suit : « Si l'on continue les tirages aussi longtemps qu'on veut, comment évoluera la fréquence relative *n(j)/N* de réalisation d'une valeur-de-couleur *j* donnée? ». Je crois pouvoir affirmer à l'avance *en tant qu'un* ***fait*** *psychologique* (peut-être installé par une longue période de réflexions sur les jeux) que, après une brève réflexion (et éventuellement un débat), il s'établira du consensus quasi-unanime concernant les réponses suivantes :

R1. Il est *presque certain* que, si l'on accroît suffisamment le nombre *N* des tirages, l'on fera apparaître toutes les *q* valeurs *j* de forme-de-couleurs.

R2. Si le nombre *N* des tirages s'accroît sans aucune limite imposée *a priori*, la fréquence relative $n(j)/N$ manifestera – plus tôt ou plus tard mais *presque certainement* et pour *tout j* – une certaine convergence. A savoir : la fréquence relative $n(j)/N$ tendra *vers le rapport* $n_T(j)/100$ *(T : tableau) qui réfère le nombre des carrés qui,* ***dans le tableau T****, portent la valeur-de-couleurs étiquetée j, au nombre total 100 des carrés*.

Mais *pourquoi y aurait-il une telle convergence* ? Et surtout *pourquoi précisément vers* ***ce rapport là,*** $\boldsymbol{n_T(j)/100}$? Et pourquoi dire – dans les deux formulations, de *R1* et de *R2* – *presque* certainement et non certainement tout court ?

Parce que, se dit-on, dès lors qu'après chaque tirage le carré tiré est remis dans l'urne et que l'on est libre de répéter un tirage autant de fois qu'on voudra, *il n'existe aucune base* pour exclure a priori strictement d'obtenir, dans une suite de longueur *N* arbitrairement grande, n'importe laquelle parmi les différentes possibilités $\{j\equiv 1,2,...q\}$; ni d'ailleurs n'importe lequel parmi les différents *ordres de succession* des valeurs de *j* appartenant à l'ensemble $\{j\equiv 1,2,...q\}$ des valeurs possibles ; ni n'importe laquelle parmi toutes les différentes *distributions statistiques globales* $\{n(j)/N, j\equiv 1,2,...q\}$, $\sum_j n(j)/N\equiv 1$, de fréquences relatives constructibles pour un *N* donné, avec les valeurs de *j* appartenant à l'ensemble $\{Dj\}$, $j\equiv 1,2,...q$. Dans les conditions de recommencement indéfini que nous avons posées, tout ce pour quoi il n'existe aucune base pour l'exclure a priori strictement, il faut l'admettre *a priori* comme possible, puisque cela revient à la *même* supposition. Sinon il s'introduirait une contradiction. Par exemple, rien ne permet d'exclure absolument la possibilité d'émergence de la distribution statistique tout à fait 'déséquilibrée' qui comporte sur toutes les places une et même valeur *j*, (par exemple *22222222……* (*N* fois *2*), ce qui veut dire que pour $j=2$ l'on y trouve la fréquence relative $n(j)/N\equiv 1$, c'est à dire qu'on a $n(j)\equiv N$ et $n(j')\equiv 0$ pour tout $j'\neq(j=2)$. Car, puisque avant le deuxième tirage on se retrouve dans exactement la même situation dans laquelle on se trouvait avant le premier tirage, si le premier tirage a pu donner $j\equiv 2$, rien n'empêche que le deuxième tirage donne lui aussi $j\equiv 2$, etc., etc.. Mais rien non plus n'empêche que l'on obtienne $j\neq 2$. Ceci conduit vers la réponse *R1* à la question *Q1*. Néanmoins – sachant que le nombre des carrés, comme aussi celui des couleurs considérées dans l'aspect *Vc*, sont finis – à chaque fois, *avant* qu'un tirage ait été opéré, il est tout aussi 'normal', se dit-on, de s'***attendre*** *plus* à voir sortir une valeur de *j* qui – ***dans le tableau T*** – intervient sur *10* carrés différents, que de voir sortir une valeur de *j* qui n'y intervient que sur *2* carrés différents. Ce qui *après* un tirage se trouvera effectivement réalisé, ne change rien au fait que l'attente d'*avant* le tirage qui vient d'être précisée, paraisse 'raisonnable'. Il ne faut pas confondre *a posteriori* et *a priori*, ni possible tout court, avec plus ou moins probable. Or en conséquence du dernier argument, se dit-on, puisque avant chaque tirage ***le tableau T est toujours dans l'urne lui seul et tout entier***, il est 'normal', si N est très grand, de s'attendre *a priori* à ce que, dans une suite assez longue, chaque valeur de j se trouve réalisée d'un nombre de fois à peu près proportionnel au nombre d'occurrences de cette même valeur j ***dans le tableau T***. C'est à dire, de s'attendre à ce que, cependant que N s'accroît, chaque fréquence relative $n(j)/N$ évolue en convergeant presque parfaitement vers le rapport $n_T(j)/100$ qui se réalise dans le tableau T pour cette même valeur j. Car dès lors que, de par les règles du 'jeu probabiliste' auquel nous avons convenu de jouer, on ne tient plus aucun compte ni de la paire de coordonnées spatiales (x_k,y_h) inscrite sur tout carré, ni des formes-de-couleur liées aux événements-description élémentaires Dj, ce ne sont plus maintenant ***que*** les rapports $n_T(j)/100,\ j\equiv 1,2,...q$ qui, ensemble, caractérisent la forme globale du tableau T qui avant chaque tirage se trouve dans l'urne seul et complet, même s'il y est en morceaux. Et il paraît presque certain que lorsque le nombre d'essais N s'accroît sans limitation cette forme globale de T finira par *manifester* sa présence dans l'urne, à chaque fois renouvelée. A savoir de *l'**unique** manière qui reste possible pour manifester cette forme globale, c'est à dire précisément par la convergence affirmée*. Mais bien entendu on ne peut pas être *sûr* que la convergence de toute fréquence relative $n(j)/N$ vers le rapport correspondant $n_T(j)/100$ se réalisera, ni, d'autant moins, qu'elle se réalisera *strictement*, car ce serait contraire aux conditions que nous mêmes avons instaurées. En effet ***toute*** suite de résultats $\sigma(j,N),\ j\equiv 1,2,....q$ de longueur N donnée mais q*uelconque*, est possible, même la suite *22222222…….* de N tirages $j\equiv 2$. Or une telle suite ***'consomme'*** certaines valeurs de la *méta*-probabilité associable à telle ou telle suite ordonnée ***entière*** considérée comme un méta-événement éléméntaire et qu'on peut dénoter σ_ω où *'ω'* est un indice de *structure d'ordre-et-structure-statistique* admettant un nombre de valeurs ν qui est très grand mais *fini*. Comme toute distribution de probabilités, celle-ci, $\{\ p(\sigma_\omega),\omega=1,2,\ldots\ldots\nu\}$, est normée elle aussi à *1*. Donc, dans un méta-espace de probabilité adéquat, cette valeur globale *1* de probabilité doit être répartie parmi *toutes* les possibilités d'ordre-et-structure-statistique $\{\sigma_\omega\ ,\ \omega=1,2,....\ \nu\}$ qui existent *a priori*. Toutefois il semble bien quasi certain que si l'on persévère assez pour accomplir un N assez grand, la coïncidence entre $n(j)/N$ et $n_T(j)/100$ se réalisera à très peu près, pour tout j. Donc c'est bien la formulation de *R2* qui s'impose comme réponse à *Q2*.

Voilà ce qui se montre lorsqu'on cherche la motivation intuitive des réponses *R1* et *R2*.

III.2.6. La définition effective d'une loi factuelle de probabilité dans le cas du 'jeu de probabilité' avec le tableau T. La motivation des réponses *R1* et *R2* qui vient d'être explicitée produit sans doute une impression de trivialité. Pourtant il s'en dégage une conclusion qui, elle, est bien loin d'être triviale. Car dans *R1* et *R2* il s'est subrepticement construite une définition effective et « à partir de faits réels », de la loi factuelle de probabilité qui agit dans le cas particulier du jeu de probabilité avec le tableau T. C'est la référence au théorème des grands nombres qui entraîne cette conclusion. En effet lorsqu'on l'écrit :

$$\forall j,\ \ \forall(\varepsilon,\delta),\qquad \exists N_0:\ \ \forall(N\geq N_0)\ \Rightarrow\ P[\,|(n(e_j)/N-p(e_j)\,|\leq\varepsilon\,]\geq(1-\delta)\qquad (2)$$

il apparaît clairement que *ce théorème représente une transcription mathématique rigoureuse de précisément le discours de motivation mi-intuitive mi-raisonnée que l'on vient d'expliciter*, à la seule condition d'une identification de termes. Les événements abstraits e_j sont à identifier aux [événements-élémentaires-***descriptions-relativisées*** factuelles] *Dj* ; et ***chaque probabilité abstraite*** $p(e_j)$ ***est à identifier au rapport factuel*** $n_T(j)/100$ ***ayant le même indice j***. En effet les nombres $\{p(Dj) \equiv n_T(j)/100)\}$, $j \equiv 1,2\ldots q$ obéissent à toutes les conditions imposées à une loi factuelle de probabilité (normation (on a $\Sigma_j\ p(Dj) \equiv 1$) ; ce sont bien des réels positifs $0 \leq p(Dj) \leq 1$ (ici des *rationnels*, par construction) ; etc.). L'ensemble des rapports $\{n_T(j)/100, j \equiv 1,2,...q\}$ dont chacun spécifie le nombre d'interventions dans le tableau *T*, de l'un parmi les *q* événements-élémentaires-descriptions-rélativisées factuelles *Dj* – juste le *nombre* d'interventions, *qui ne comportent* ***rien*** *d'aléatoire*, notons-le bien – l'ensemble de ces rapports donc, définit une *loi de probabilité* ***factuelle*** sur l'univers d'événements élémentaires $\{Dj\}$, $j \equiv 1,2\ldots q$. A savoir la loi

$$\{(p(Dj) \equiv n_T(j)/100), j \equiv 1,2,\ldots.q\}$$ [18]

Les rapports $\{n_T(j)/100, j \equiv 1,2,\ldots.q\}$ définissent d'une manière ***finie, effective***, une loi de probabilité factuelle ayant une ***structure*** bien spécifiée.

> *Ainsi, dans ce cas, la définition factuelle de la structure de la loi de probabilité à poser, est accomplie.*

Cette conclusion, avec les questions et réponses qui y conduisent, s'accompagne de la solution, également, à la question de la signification assignable, en ce cas, à *l'****existence*** d'une loi factuelle de probabilité. En effet dans *R1* et *R2* l'*existence* d'une loi factuelle de probabilités $p(Dj)$, $j \equiv 1,2\ldots q$ – l'existence *seule*, abstraction faite, maintenant, de la *structure* de la loi – est induite en tant qu'*une 'expression' de la présence dans l'urne, avant chaque tirage, du tableau T tout entier et de ce tableau seulement*, mais qui ne parvient à notre perception que progressivement et en état cryptique, via des fréquences relatives aléatoires et évolutives de suites de 'signes' $\{Dj\}$, $j \equiv 1,2\ldots q$ où les indications de formes dans l'espace (tracées à l'aide de couleurs) portées par les descriptions relativisées $D/G,\kappa(x_k,y_h),Vc \cup V(E) \cup V\phi c/$, $k=1,2,\ldots 10$, $h=1,2\ldots 10$, spécifiées au départ, sont *perdues* par le filtrage de l'aspect de forme $V\phi c$. Je dis 'signes' parce que – en conséquence de ce filtrage – les descriptions élémentaires *Dj*, $\{j \equiv 1,2..q\}$ sont restées dépourvues de toute suggestion d'une participation à une 'signification' globale qui déborderait chacun de ces événements-élémentaires-descriptions-relativisées et les engloberait tous en une méta-*structure* d'espace-aspects.

Bref, dans le cas considéré la forme-de-couleurs *globale* inscrite sur le tableau *T*, et aussi le mode de morcellement du tableau *T*, déterminent d'une façon *effective* aussi bien l'*existence* que la *structure* d'une loi factuelle de probabilité agissant sur l'ensemble d'événements-élémentaires-

[18] Le rapport *n(j)/100* est un nombre rationnel, cependant qu'une probabilité *p(Dj)* peut être un nombre réel. Je signale ce fait sans tenter ici d'en approfondir la signification ou les conséquences formelles.

descriptions-relativisées *Dj,* $\{j\equiv 1,2..q\}$ déterminés par les référentiels *locaux* (G_k, Vc). Nous disposons d'un modèle factuel de la loi de probabilité qu'il convient de poser.

> ***En ce cas particulier le concept abstrait de mesure de probabilité PEUT donc être doté d'une interprétation factuelle effective : l'ensemble des rapports*** $\{n(j)/100\}, j\equiv 1,2\ldots q$.

Pour accéder à cette interprétation il a fallu *quitter* le niveau de conceptualisation où sont confinées les manifestations directement observables de la situation probabiliste proprement dite créée *à partir* du tableau *T*. Il a fallu monter au niveau d'ordre supérieur où se trouve placé le tableau *T lui-même*. Or – face aux fragments *Dj* isolés qui, eux *exclusivement*, interviennent d'une manière directe dans la situation de 'jeu probabiliste' –, le tableau *T* tout entier est *une forme-de-couleurs* ***NON-observable***. Car dans la définition du paragraphe *III.2.5* d'un 'jeu probabiliste' cette forme globale a disparu du domaine du perçu, bien qu'elle règle ce qu'on perçoit. Elle est passée en dessous de l'horizon comme un soleil qui vient de se coucher mais continue de d'illuminer le ciel et les nuages. Cela suggère la voie vers une généralisation.

III.2.7. Deux généralisations : 'probabilisation' d'une description relativisée individuelle quelconque, versus 'déprobabilisation' d'une 'situation probabiliste'. Dans le cas du jeu de probabilité du paragraphe *III.2.5* nous sommes parti d'une description *individuelle* (au sens du point *5* de *III.1*), celle du tableau *T*, et, sur la vase de la connaissance de cette description individuelle, il a été possible d'en accomplir une 'probabilisation' où la loi de probabilité *factuelle* correspondante a pu être spécifiée, en termes finis. Ce résultat admet-il des généralisations ? On peut en spécifier deux, en quelque sorte inverses l'une de l'autre, et qui semblent couvrir l'entier domaine où les concepts de probabilité et de complexité se trouvent reliés. La première de ces généralisation concerne les 'probabilisations' d'une description individuelle quelconque, et la deuxième, les déprobabilisations d'une 'situation probabiliste' quelconque.

Probabilisations. Il semble clair que, dans son essence, l'analyse faite plus haut s'applique à toute description *individuelle* (au sens du point *5* de *III.1*) qu'on souhaite 'probabiliser', i.e. à laquelle on veut associer une loi factuelle de probabilité.

> *Selon MCR toute description individuelle est une forme d'espace-physique-valeurs-d'aspects, et toute telle forme peut être 'probabilisée' par un procédé analogue à celui employé avec le tableau T, ce qui conduit à un espace de probabilité correspondant où la mesure de probabilité possède par construction une interprétation* ***factuelle*** *tout à fait définie, finie, effective.*

Cette première conclusion, on le verra, représente d'ores et déjà une avancée significative pour le traitement de la question générale de la définition de mesures de complexité. Mais elle ne fournit pas une réponse générale à la question de savoir comment on définit une loi de probabilité factuelle lorsqu'une 'situation probabiliste' factuelle constitue la donnée première.

Déprobabilisations. Supposons donc maintenant que l'on ***commence*** par se donner une 'situation probabiliste'. C'est à dire, on se donne un phénomène aléatoire *[𝒫,U]* où la procédure *𝒫,* dont on dit qu'elle est reproductible 'identiquement', néanmoins engendre tout un univers *U* d'événements élémentaires mutuellement distincts lorsqu'elle est reproduite. Selon *MCR* la procédure *𝒫* consiste en une succession d'opérations *[G.V]* où : l'opération *G* introduit une entité-objet à décrire $œ_G$; la vue *V* – d'une manière analogue à ce qui se passe dans le cas d'une 'vue' de mesure-physique – est 'active', en ce sens qu'*elle **crée** l'entière action de réalisation de 'l'expérience aléatoire'* qui implique l'entité-objet-de-description introduite par *G*. Une réalisation de l'expérience aléatoire comporte divers objets (ou appareils) et une grille de qualification *en termes de valeurs gk de un ou plusieurs aspects g bien définis* de l'état qui a résulté de cette réalisation, pour l'entité-objet $œ_G$, (cf. la réf. 16E) [19]. C'est cette qualification finale, dénotée *Dr*, qui spécifie quel événement-élémentaire-*description*-relativisée de l'univers *U*, a été produit par la réalisation considérée de l'expérience aléatoire. On a donc *U={Dr}, r=1,2,….s*. Au départ, dans une situation probabiliste donnée, on ne connaît *que* le phénomène aléatoire *(𝒫,U)*. On ne connaît pas la loi de probabilité à poser sur *U*, ni, a fortiori, une 'forme globale' *F* qui permette de déterminer cette loi d'une façon analogue à celle mise en évidence dans *III.2.6* concernant un jeu de 'puzzle probabiliste'. Mais dire que l'on se trouve en 'situation probabiliste' implique que l'on *pose* néanmoins qu'une loi de probabilité *'existe'*.

On veut associer à cette loi inconnue une définition-interprétation finie, effective, liée à une forme *F* consistant en une description *individuelle* qui 'corresponde' *globalement* au phénomène aléatoire *(𝒫,U)*. On veut 'géométriser' la situation probabiliste indiquée par *(𝒫,U)* et, sur le forme globale *F* qui réalise cette géométrisation, définir une fragmentation finie convenable qui permette de *déterminer* ***factuellement*** la loi de probabilité qu'il convient d'y affirmer ; donc aussi la mesure *abstraite* de probabilité à introduire dans l'espace de probabilité mathématique qui représente la situation probabiliste en termes purement mathématiques. Une telle procédure, si elle était spécifiée, reviendrait à associer à la situation probabiliste de départ, une '*dé*-probabilistion correspondante', puisqu'elle la représente à l'aide d'une description globale *individuelle* bien définie. Ceci répondrait en toute généralité aux exigences d'interprétabilité factuelle de Kolmogorov. La probabilisation d'une description individuelle connue en tant que donnée première, fondée sur l'exemple du tableau *T*, n'était qu'un premier pas tenté avec ce but final. Mais est-il possible de réaliser ce but ?

[19] Dans le cas du jet d'un dé la vue *V* impliquée dans la 'procédure' *𝒫* comporte le dé, une surface plane, le mécanisme de jet (par homme ou machine) et une vue-aspect dont les valeurs sont les nombres de *1 à 6* lisibles sur le *6* faces du dé.

L'entière conceptualisation scientifique ***classique*** *comporte le 'postulat déterministe' selon lequel la question formulée plus haut* **admet** *une réponse positive* [20,21].

Toutefois ce postulat déterministe ne s'associe pas à l'indication d'une méthode de construction effective d'une telle réponse. C'est précisément l'absence d'une telle méthode qui fait problème. Dans ce qui suit j'indique les grands traits d'une méthode constructive.

Avant de délinéer la construction il est important de réaliser tout de suite que la question de 'savoir' si, relativement à une situation probabiliste donnée, une forme globale F 'existe' ou n''existe pas', est une question illusoire. Concevoir une telle forme comme une 'vérité' qu'il s'agirait de 'découvrir', ne peut conduire qu'à une impasse. Il ne s'agit pas de découvrir mais de ***construire un MODELE*** du concept de loi factuelle de probabilité correspondant à une situation probabiliste donnée, qui soit satisfaisant autant du point de vue sémantique que du point de vue de la cohérence logique, et qui puisse donc constituer une interprétation générale du concept abstrait de mesure de probabilité.

L'exemple de la probabilisation d'une description individuelle $D/G_T,T,Vc\cup V(E)\cup V\phi c/$ du tableau T, que l'on se donne au départ, servira comme un premier élément de référence, de la manière suivante. En s'appuyant sur cet exemple l'on tentera de délinéer une démarche qui, en un sens qui reste à être précisée, soit une 'inversion' de la démarchedes points *III.2.5* et *III.2.6*.

Dans le cas du tableau probabilisé, l'univers des événements-élémentaires-descriptions-relativisées $\{Dj\}$, $j=1,2,...q$ a pu être obtenu par un simple morcellement d'une description individuelle globale qui, elle, était *pré-constituée,* ***actuelle*** *et* ***connue***. Cette description individuelle globale constituait la donnée première. Le phénomène aléatoire correspondant que nous avons construit au point *III.2.5* en partant de cette description globale individuelle, ne consistait que dans une manipulation superficielle des descriptions locales Dj, pré-existantes elles aussi à l'intérieur de la forme globale – toutes à la fois et actuelles – en ce sens qu'elles ont pu être transformées en 'événements élémentaires' par simple découpage de la description globale pré-donnée qui déjà les contenaient toutes. Tandis que, lorsque la donnée *première* consiste en une situation probabiliste, en un phénomène aléatoire $(\mathcal{P},U)$, les circonstances sont très différentes. Les éléments de l'univers d'événements-élémentaires-descriptions-relativisées $U=\{Dr\}$, $r=1,2,...s$ sont engendrés un à un par la procédure $\mathcal{P}$, à tour de rôle, au fur et à mesure que s'effectuent successivement les essais comportés par $\mathcal{P}$. Et chacun de ces événements est engendré en un sens beaucoup plus créatif, et quelquefois même radicalement créatif, comme lorsqu'il s'agit de descriptions de microétats. En effet dans le cas des phénomènes aléatoires qui engendrent les espaces de probabilité quantiques, *l'entité-objet* ***elle-même*** – un microétat

[20] Longo,G., "Laplace, Turing et la géométrie impossible du « jeu de l'imitation » ", *Intellectica* 35, 2002.

[21] La microphysique moderne implique le postulat ***contraire***. La description quantique des microétats est construite directement à partir de données qui court-circuitent les disciplines de la physique classique. Le caractère probabiliste des prévisions concernant des microétat y apparaît *comme une* ***donnée première***. En ce sens on peut dire que les probabilités quantiques sont *'primordiales'*. J'emploie ce terme parce que, à la différence de celui, usuel, de probabilités 'essentielles', il est libre de toute connotation ontologique. En outre, l'interdiction de modélisation introduite par la mécanique quantique orthodoxe, exclut également les modèles déterministes *a posteriori*. Ainsi en mécanique quantique orthodoxe la question de la détermination et de l'interprétation du concept de loi factuelle de probabilité, reste béant.

– *ne préexiste pas*. L'opération de génération *G* mise en œuvre la re-crée entièrement cette entité-objet lors de chaque nouvel essai. Cependant que l''expérience' (la réalisation de la procédure *𝒫*) qui implique l'exemplaire d'entité-objet engendré par une réalisation *donnée* de l'opération de génération *G* – i.e. un acte de mesure subséquent accompli sur cet exemplaire-là – engendre *entièrement* les manifestations observables qui constituent l'événement-élémentaire-description-relativisée observé. Cet événement élémentaire est produit de fond en comble par *interaction* entre l'entité-objet *inconnue* créée par *G*, et les appareils impliqués dans l'acte de mesure. Il s'agit là d'un cas limite de degré de créativité d'un phénomène aléatoire [22]. Mais il existe aussi des phénomènes aléatoires qui ne sont créatifs que partiellement, en ce sens que l'entité-objet pré-existe et qu'on lui assigne des propriétés pré-existentes indépendamment des réalisations de la procédure *𝒫*. On peut par exemple penser au cas d'un jet de dé sur une table : la manière dont s'est accompli chaque jet, la structure de la surface de la table, les détails de la structure du dé (ses coins, son centre de gravité, etc.), le degré de planéité de la surface de la table, et ainsi de suite, déterminent les issues une à une. Chacune de ces issues est l'effet d'une véritable genèse d'espace-temps-valeurs-d'aspect. Les remarques qui précèdent convergent avec la célèbre interprétation 'propensionnelle' de K. Popper. Je cite (ma propre traduction de l'anglais) [23] :

> «Considérons par exemple une planche de billard ordinaire symétrique construite de telle façon que si nous laissons rouler un nombre de petites billes, elles formeront (idéalement) une courbe normale de distribution. Cette courbe représentera la distribution de probabilités pour chaque expérience individuelle, avec chaque bille individuelle, d'atteindre un endroit de repos. Maintenant 'donnons un coup' à cette planche ; disons, en soulevant légèrement son côté gauche. Alors nous 'donnons un coup' également à la propension, et à la distribution de probabilités….Ou bien, au lieu de cela, déplaçons *une aiguille*. Cela altérera la probabilité de chaque expérience individuelle avec chaque bille individuelle, que la bille approche effectivement ou non l'endroit duquel nous avons enlevé l'aiguille……Nous pouvons demander : ″Comment la bille peut-elle 'savoir' qu'une aiguille a été enlevée si elle ne s'approche jamais de l'endroit ?″ La réponse est : ″la bille ne 'sait' pas ; mais la planche en tant qu'un tout 'sait', et elle change la distribution de probabilités, ou la propension, pour *chaque* bille; un fait qui peut être testé par des tests statistiques″».

Ces considérations montrent que cette fois la forme globale *F* à construire doit incorporer des vues-aspect foncièrement ***processuelles*** concernant *l'entier* phénomène aléatoire *(𝒫,U). Le temps doit donc intervenir aussi, amplement, et une vue à dimensions de qualification dynamiques*. Il s'agira de construire les éléments d'un jeu de puzzle dans un espace de représentation extrêmement complexe, incorporant les deux aspects-cadre, d'espace *et* de temps. En ce sens il n'y a pas de symétrie face au cas de probabilisation d'une description globale individuelle qui est connue dès le départ en tant que donnée première.

Pour simplicité admettons, comme dans le cas du tableau *T*, que les événements-élémentaires-descriptions-relativisées de l'univers *U={Dr}, r=1,2,...s* impliqué dans la situation probabiliste à déprobabiliser, ne font intervenir qu'une seule vue-aspect *Vg*. Les *s* événements élémentaires sont

---

[22] C'est dans cette situation d'extrême limite, où les concepts classiques d''objet' et de 'propriétés *d'un objet*' n'ont plus aucun sens de manière que l'entier substrat conceptuel de la physique classique s'évanouit, qu'émergent les probabilités à caractère 'primordial' dont il est question dans la note 21.

[23] Popper, K., "Quantum Mechanics without the Observer", in :
A. "Quantum Theory and Reality", Mario Bunge ed., *Springer* 1967 ; **B**. "A World of Propensities", *Thoemmes*, 1980.

donc étiquetés par les valeurs d'aspect $[r \equiv gk]$, $k=1,2,...s$, après avoir éliminé par filtrage tout autre aspect que l'on pourrait y percevoir.

Dans le cas du tableau nous avions défini au départ une vue $V \equiv V(E) \cup Vc \cup V\phi c$ qui déterminait la forme globale de valeurs de couleurs en laquelle consistait la description relativisée individuelle du tableau *T* entier. Cette *même* vue intervenait ensuite dans les descriptions individuelles *locales*, $D/G,\kappa(x_k,y_h),Vc \cup V(E) \cup V\phi c/$, $k=1,2,...10$, $h=1,2...10$, portées par les *100* carrés obtenus en découpant le tableau. Ce sont précisément ces ***formes*** locales d'espace-valeurs-de-couleur engendrées à l'aide de la vue-aspect $V\phi c$ de *forme*-de-couleur, qui permettaient aux points *III.2.3* et *III.2.4* préliminaires à la probabilisation de *T*, de jouer à des jeux de puzzle fondés sur des attractions par continuité aux frontières spatiales, entre les formes locales d'espace-valeurs-de-couleur. Cependant que lors de la probabilisation du point *III.2.5* ces formes locales se sont ***perdues*** par la suppression de l'aspect de l'aspect-cadre d'espace physique *E* et de l'aspect de forme d'espace-couleur $V\phi c$, ce qui a conduit aux descriptions-étiquettes *Dj* : avant de probabiliser il y a eu ***simplification*** des descriptions locales $D/G,\kappa(x_k,y_h),Vc \cup V(E) \cup V\phi c/$ qui constituaient les pièces d'un puzzle.

En toute généralité, les descriptions probabilistes impliquent la dissolution de la topologie d'espace-temps-valeurs-d'aspects de la forme globale F qu'on peut leur associer.

Afin de pouvoir, à la fois, inverser le sens de la démarche probabilisation, *et* la généraliser, il conviendra maintenant de ***complexifier*** les descriptions de l'univers d'événements élémentaires $U=\{Dr\}$, $r=1,2,...s$ qui définissent la situation probabiliste donnée, de manière à transformer chaque description *Dr*, d'un simple signe d'étiquetage d'une valeur 'in-forme' *gk* d'un seul aspect $g \equiv r$, dans une ***forme*** locale d'espace-t*emps*-valeurs-d'aspects : C'est le pivot du procédé d'inversion délinéé plus bas.

Détaillons ce procédé. La vue développée dans la théorie des catastrophes de René Thom, est particulièrement appropriée : une modification morphogénétique est induite dans le 'substrat' correspondant à la procédure $\mathscr{P}$, et cette modification subit vers sa fin une 'attraction catastrophique' vers l'un ou l'autre d'un ensemble des 'bassins d'attraction' constitués par les événements élémentaires *Dr* de l'univers de départ *U*. Ces derniers sont tous posés être de nature *physique*, et observables. Donc chaque description *Dr* implique de départ nécessairement une localisation sur quelque domaine de l'espace physique (cf. la fin du point *5* de *III.1*). Mais en fait la description-étiquette *Dr* n'indique que l'issue catastrophique hypostasiée d'une *autre* description, d'espace-***temps***-valeur***s***-d'aspect***s*** : toute une morphogenèse, face à laquelle la seule valeur $r \equiv g$ avec sa localisation spatiale finale, n'est qu'une brutale simplification. Il s'agit d'élaborer des représentations de ces morphogenèses, par une complexification descriptionnelle qui fasse intervenir aussi des valeurs d'autres aspects, différents de *r,* impliquant le substrat stable ainsi que ses modifications. Les 's' des pluriels que souligne l'écriture 'espace-***temps***-valeur***s***-d'aspect***s***' se rapportent aux aspects nouveaux qui permettront d'exprimer cette complexification descriptionnelle.

Il s'agit d'élaborer une représentation des **genèses** *des événements de l'univers $U \equiv \{Dr\}$, $r=1,2,\ldots s$, qui soient telles qu'elles puissent éliminer le 'hasard' comporté par le phénomène aléatoire ($\mathscr{P}$,U) – i.e. la non-prédictibilité des issues catastrophiques $\{Dr\}$, $r=1,2,\ldots s$ – par une 'déprobabilisation' de ($\mathscr{P}$,U) en une forme individuelle globale F de morphogenèses 'locales' d'espace-temps-valeurs-d'aspects-différents-de-r.*

On cherche donc un diagramme morphogénétique multidimensionnel à la Thom. Rappelons de nouveau que le postulat déterministe qui sous-tend l'entière pensée classique, affirme précisément la possibilité de spécifier un tel diagramme morphogénétique. La complexification recherchée revient à introduire une vue complexifiée V^c convenable. La nouvelle description-événement obtenue à l'aide de V^c en partant d'un événement-description final et hypostasié *Dr* donné, peut être dénotée $D^c r'(r)$ et dénommée *une complexification morphogénétique de Dr*, où *r'* est un indice indifférencié distinguant l'une de ces complexifications. On obtiendra ainsi un nouvel univers d'événements élémentaires morphogénétiques consistant en un ensemble $U^c=\{ D^c r'(r)\}$, $r'=1,2,\ldots s'$ d'évolutions catastrophiques dont le cardinal *s'* est par construction plus grand que le cardinal *s* de l'univers initial $\{Dr\}$, $r=1,2,\ldots s$) des bassins statiques d'attraction : $s' \geq s$.

L'univers U' = $\{D^c r'(r)\}$, $r'=1,2,\ldots s'$ de morphogenèses individuelles, peut être regardé comme l'ensemble des pièces d'un jeu de puzzle auquel on peut tenter de jouer à l'intérieur de l'espace de représentation de la vue complexifiée V^c.

Comme dans un jeu de puzzle habituel [24] chacun des événements élémentaires morphogénétiques mutuellement distincts $D^c r'(r)$ doit maintenant trouver sa place *propre* dans l'espace de représentation de la vue complexifiée V^c, sur la base de la recherche de continuités sémantiques aux frontières entre événements. Celles-ci seront liées particulièrement aux valeurs des aspects *de substrat* qui ne participent ***pas*** aux morphogenèses individuelles (les paramètres dynamiques qui décrivent les morphogenèses individuelles forment des qualifications 'fermées' sur elles-mêmes délimitant une sorte de fibres, en conséquence de leur dimension temporelle). Cette sorte d''attractions' par continuités sémantiques de bord impliquant des paramètres de substrat, détermineront progressivement ***une topologie globale d'espace-temps-valeurs-d'aspects***. Les morphogenèses individuelles à développements 'avoisinants' formeront une sorte de gerbe affluant vers un bassin d'attraction donné. Entre les morphogenèses individuelles et entre les gerbes qu'elles

[24] En jouant à ce jeu de puzzle probabiliste, on identifierait sans doute, par essais et erreurs, certaines optimalités en ce qui concerne le meilleur choix de la vue complexifiée V^c. En outre, les procédures courantes dans les opérations de 'scanning', dans les méthodes de prévision météorologique, etc., qui appartiennent à des familles de démarches analogues, pourraient apporter des suggestions importantes pour la spécification exacte de la procédure générale recherchée ici.

forment, se constitueront des distances et des orientations relatives. Cette topologie fera progresser vers la constitution de la forme globale *F* recherchée d'espace-temps-valeurs-d'aspects associable à la description probabiliste de départ.

Mais la forme globale *F* ne s'ébauchera qu'au sein du processus de *génération* ***entremêlée*** d'un nombre non-défini d'exemplaires de *F*. Au cours de ce processus il faudra déterminer aussi exactement que possible le nombre total $n_F(r')$ de descriptions morphogénétiques complexifiées $D^cr'(r)$ qui est *nécessaire* et *suffisant* pour construire ***un*** exemplaire complet de *F* et *un* s**eul**, extrait de l'océan informe du nombre *N* de réalisations d'une expérience aléatoire qui produirait des suites illimitées d'éléments de l'univers complexifiée de morphogenèses $\{D^cr'(r)\}$, $r'=1,2,...s'$. Car *le nombre* $n_F(r')$ *est une* ***inconnue***. Et c'est une inconnue d'importance cruciale pour la détermination effective de la loi factuelle de probabilité recherchée. Une façon d'acquérir la connaissance de ce nombre $n_F(r')$, est d'essayer de simuler sur ordinateur l'expérience aléatoire qui engendre l'univers $\{D^cr'(r)\}$, $r'=1,2,...s'$, tout en jouant au puzzle correspondant à cet univers. A partir d'un nombre *N* d'expériences assez grand, on devrait constater que pour tel ou tel exemplaire parmi l'ensemble des exemplaires de *F* qui se constituent progressivement et de manière entremêlée, l'on ***cesse*** de trouver une morphogenèse individuelle D^cr' qui puisse encore y trouver une place, selon les critères d'attraction de continuité de bord : ce sera alors un exemplaire achevé de la forme globale *F*. D'autres exemplaires achevés devraient ensuite se former, plus tôt ou plus tard, en jouant un rôle de confirmation. Un programme informatique appropriée aboutirait peut-être assez vite à ce genre de stabilité. Par cette voie on doit donc pouvoir identifier l'ensemble $\{n[D^cr'(r)]/n_F(r')\}$ normé à *1* de tous les rapports rationnels $n[D^cr'(r)]/n_F(r')$, $r'=1,2,.....s'$, $\Sigma_r n[D^cr'(r)]=n_F(r')$, qui interviennent dans *un* exemplaire complet de la forme globale *F*, et un seul.

Une fois ces rapports établis, on pourra ***revenir*** *aux descriptions-étiquettes Dr de départ*, en faisant maintenant ***abstraction*** a posteriori de tous les aspects génétiques complexifiants qui, en définissant des morphogenèses $D^cr'(r)$, ont permis de jouer à un puzzle et déterminer ainsi une forme globale *F*. Lors de cette rétro-simplification par abstraction, chaque description-étiquette *Dr* 'collectivisera' ou 'absorbera' en elle la classe $D^cr'(r)$ de toutes les morphogenèses qui basculent de manière catastrophiques dans le bassin d'attraction désigné par la valeur de l'indice *r* qui intervient dans *Dr*.

> Les cardinaux des classes $D^cr'(r)$, associés aux nombres $n_F(r')$ établis précédemment, permettront de calculer le nombre $n_F(r)$ d'occurrence dans *F* de toute description-étiquette *Dr* donnée.

Ainsi l'on obtiendra finalement aussi l'ensemble $\{n(Dr)/n_F(r)\}$ normé à *1* des rapports rationnels $n(Dr)/n_F(r)$, $r\equiv 1,2,.....s$, $\Sigma_r n[Dr]=n_F(r)$, qui correspond aux événements-élémentaires-descriptions-étiquettes de l'univers ***de départ*** $U=\{Dr\}$, $r=1,2,...s$.

Cet ensemble de rapports rationnels $\{n(Dr)/n_F(r)\}$, $r \equiv 1,2,\ldots.s$ peut représenter d'une manière 'effective' la loi de probabilité factuelle $\{p(Dr)\}$, $r \equiv 1,2,\ldots.s$ à affirmer dans la situation probabiliste donnée au départ.

On pourrait percevoir comme un 'problème' le fait que les rapports de l'ensemble $\{n(Dr)/N\}$, $r \equiv 1,2,\ldots.s$ sont rationnels tandis que les probabilité de la loi de probabilités 'correspondante' $\{p(Dr)\}$, $r \equiv 1,2,\ldots.s$ sont en général des nombres réels. Mais cette différence est inévitable. Elle est due au fait que l'on substitue une approche factuelle qui est par construction discrète, finie, définie – comme aussi nos observations directes – , à une approche de mathématiques continues qui, elle, est par construction non-effective et approximative dans son essence même – face à ce qu'elle est censée représenter – parce qu'on veut pouvoir représenter à son aide *toute* situation probabiliste factuelle, quelles que soient ses caractéristiques non-spécifiées. On est en présence d'une illustration frappante du genre de questions que peuvent soulever les représentations mathématiques continues, et soumises à une exigence de généralité maximale, lorsqu'elles sont confrontées à des procédures foncièrement discrètes et effectives, comme notamment celles de l'informatique moderne (voir à ce sujet le très intéressant travail de Giuseppe Longo indiqué à la référence 20).

Enfin, en *comparant* les différentes formes globales F liées à des complexifications V^c différentes de la vue-aspect Vg, $g \equiv r$, utilisée au départ, on devrait pouvoir détacher certains *invariants* qui puissent coder pour l'entité-objet ***unique*** $œ_G$ qui intervient de façon stable dans toutes les descriptions-événements-élémentaires Dr. Ces invariants offriraient un équivalent d'un 'objet' au sens du langage courant [25].

Sylvie Leleu-Merviel exprime dans ses travaux [26] des attentes tout à fait consonantes avec les idées soumises ici, et elle y signale d'autres travaux qui vont dans des directions analogues [27].

La procédure qui vient d'être indiquée sera dénommée ***un algorithme d'intégration sémantique*** *d'une situation probabiliste donnée*. Cet algorithme permet de définir de façon effective, aussi bien le sens de *l'existence*, que celui de la *structure* de la loi *factuelle* de probabilité à affirmer dans toute situation probabiliste donnée. L'objection de Kolmogorov est dissoute, du moins en principe. Je montrerai maintenant que ce résultat clarifie la manière spécifique d'utiliser la théorie de l'information pour des estimations de complexité.

---

[25] Personne ne perçoit jamais un 'objet' lui-même – une table, une chaise – en dehors de tout 'point de vue' comporté par une situation observationnelle particulière. Le mot qui dénomme un 'objet', on le sait bien, n'est qu'une étiquette verbale pour tout un ensemble de descriptions relativisées qu'on peut en donner de points de vue différents (à l'aide de vues différentes).
[26] Leleu-Merviel, S., :**29A**. "La structure du Aha Aha. De la fulgurance comme percolation", *H2PTM'05* ;
29B. "Le désarroi des 'Maîtres du sens' à l'ère numérique"., *H2PTM'03*.
[27] Fauconnier G., Turner M., "The Way we Think,. Conceptual Blending and the Mind's Hidden Complexities", *Basic Books*, 2002.

IV. LE ‘SENS’ DANS LA THEORIE DE SHANNON DE LA *COMMUNICATION DE SIGNES CODANTS*

On dit couramment que la théorie de ‘l‘information’ de Shannon (appelée aussi ‘théorie des communications’, ou de la transmission d’informations), serait « vide de tout sens », qu’elle serait « purement syntaxique », rien que des algorithmes. Or lorsqu’il s’agit d’une structure formalisée qui possède une si étonnante potentialité d’***applications*** diverses, cette affirmation est *a priori* aberrante. La seule question pertinente est celle de savoir *quelles* significations sont impliquées dans cette théorie et où et comment elles interviennent.

IV.1. Rappels

L’un des concepts de base de la théorie de Shannon est celui de *source aléatoire de* ***signes***, dénotée usuellement S [28]. Une source aléatoire de signes S est définie comme le lieu d’émission d’un *alphabet de signes* $A\equiv\{a\}$, $i\equiv1,2,\ldots.q$, soit des signes de nature physique (signaux sonores, électromagnétiques, etc.), soit des signes de nature *conceptuelle* (comme dans le cas des mots d’une langue courante), soit enfin des signes strictement conventionnels. Chaque signe a_i émis par la source S est associé à une *probabilité* ‘locale’ $p_i\equiv p(a_i)$ d’être émis, qui est élément d’une loi de probabilité $\{p_i\}$, $i\equiv1,2,\ldots.q$ où selon la condition de normation on a $\Sigma_ip_i\equiv1$.

La quantité numérique de forme entropique $H(S)\equiv\Sigma_ip_ilog(1/p_i)$ formée avec les éléments p_i de la loi de probabilité $\{p_i\}$, $i\equiv1,2,\ldots.q$, est dénommée *l’entropie de la source S*. On dit couramment que $H(S)$ représenterait ‘la quantité d’ ‘information’ contenue dans la source S [29]. La quantité de forme entropique $H(S)\equiv\Sigma_ip_ilog(1/p_i)$ est *omniprésente* dans la syntaxe de Shannon et son rôle y est déterminant.

Selon la théorie initiale de Shannon, l’alphabet de signes $A\equiv\{a_i\}$, $i\equiv1,2,\ldots.q$ est utilisé *pour transmettre des ‘****messages****’*, d’où la dénomination initiale de théorie des communications d’ ‘informations’. Pourtant la syntaxe de la théorie ne dit strictement *rien* sur le ***sens*** de ces messages, ce qui revient à l’absence de toute restriction à cet égard. Il est seulement présupposé que tout ‘message’ possède un sens.

Mais ***ce sens là****, celui des messages transmis et reçus, n’intervient en effet nulle part dans la syntaxe de Shannon.*

---

[28] Je refuse la dénomination de ‘source d’ ‘***information***’ : l’information à transmettre se trouve dans la conscience de l’émetteur d’un message, et l’information reçue se constitue dans la conscience du récepteur du message. Une ‘source’ aléatoire au sens de Shannon est juste une source de *signes*. Etant dépourvue d’une conscience, elle ne peut contenir aucune ‘information’. Par contre il apparaîtra qu’elle comporte une *forme globale* qui associe un ‘sens’ à la loi de probabilité posée sur l’alphabet des signes qu’elle émet.

[29] Je prie le lecteur d’oublier l’expression courante d’ ‘entropie ***informationnelle***’, car elle crée confusion, pour les raisons exprimées dans la note 28.

Voilà pourquoi on dit qu'il n'y aurait aucune sorte de sens dans la théorie de Shannon. Etant donné qu'ici cette affirmation est refusée a priori, il s'agit de chercher l'élément de la théorie où se cache du sens, et de spécifier ce sens.

La *communication* d'un message se réalise toujours à travers tel ou tel *canal de transmission C*. Les signes a_i de A doivent être *codés* avant de pénétrer dans C. On doit donc introduire un deuxième alphabet, un *alphabet codant* $X \equiv \{x_j\}$, $j \equiv 1,2,\ldots.r$. Pour des raisons pragmatiques chaque signe a_i est d'habitude codé par tout un *mot de code* X_i ('codage en blocs') qui consiste en une suite *finie* de signes de code x_j. L'on introduit en outre des modes divers de 'ponctuation' qui séparent les mots les uns des autres. Un mot de code a donc la forme $X_i \equiv x_s x_k \ldots x_w$. La *longueur* l_i de X_i est par définition le nombre de lettres de code dans le mot X_i. En *particulier on peut avoir* $l_i \equiv 1$.

Un *message* est un mot de code ou une suite de mots de code.

Un signe d'entrée a_i est en général *transformé* au cours de sa traversée du canal C, et de manière *aléatoire*. Donc à la sortie de C on se trouve en général en présence d'une lettre b_j d'un autre alphabet, 'de sortie', $B \equiv \{b_j\}$, $j \equiv 1,2,\ldots.n$, chaque lettre de B étant associée à une probabilité d'apparition conditionnelle, $p(b_j/a_i) \equiv p_{ji}$. C'est à partir de l''output' exprimé en lettres de B qu'il faudra, *malgré les effets* ***aléatoires*** *du canal C*, reconstituer le message émis en mots de code $X_i \equiv x_s x_k \ldots x_w$. Cela exige des méthodes spécifiques de 'décodage' (de reconstruction de la forme de départ du message reçu) qui puissent faire face au caractère aléatoire de la déformation de X_i lors de sa transmission par le canal C.

Via le fonctionnelle $H(S) \equiv \Sigma_i p_i \log(1/p_i)$ d'entropie de la source S, c'est la loi de probabilité $\{p_i\}$, $i \equiv 1,2,\ldots.q$ sur les signes de la source qui joue le rôle *majeur* dans toutes les phases de la théorie de Shannon, celle de codage, celle de transmission, et celle de décodage. C'est donc sur cette loi $\{p_i\}$, $i \equiv 1,2,\ldots.q$ que nous nous concentrons plus bas pour identifier de quelle façon la syntaxe de Shannon comporte du sens.

IV.2. Le *'sens'* dans la théorie de Shannon

Soit donc une source aléatoire de signes

$$S \equiv \left| \begin{array}{l} \{a_i\},\ i \equiv 1,2,\ldots.q \\ \{p_i\},\ i \equiv 1,2,\ldots.q \end{array} \right|$$

Dans l'exacte mesure où les signes a_i de l'alphabet A émis par S sont de quelque manière perçus *physiquement* par des hommes ou des appareils et communiqués *en tant que 'signaux', ils sont à regarder comme un ensemble de* ***descriptions*** *d'entités-objet* ***physiques****, et chaque telle description introduit quelque référentiel épistémique où elle est élaborée* [30].

[30] On peut 'montrer' ce qu'on veut décrire, au lieu de le dire, mais cela ne change rien au fait qu'il intervient nécessairement quelque triade $(G, œ_G, V)$ et la description relativisée correspondante. Car *toute connaissance communicable est description* (cf. le point *8* de III.1).

Cette affirmation vaut même lorsqu'il s'agit de signaux 'conceptuels', comme des mots, et même lorsqu'il s'agit de simples signes conventionnels écrits ou lus, comme des lettres. En tant que signal perceptible et transmissible, toujours, en dernière essence, il s'agit plus ou moins explicitement d'une ***description*** d'une entité-objet *physique*. Donc la source S peut être regardée comme une description probabiliste où le générateur G introduit comme entité-objet un émetteur S de signes qui, lui, par quelque phénomène aléatoire, produit l'alphabet $A=\{a_i\}$, $i\equiv 1,2,\ldots.q$ en tant qu'univers $U=\{Di\}$ où $Di\equiv a_i$, d'événements-élémentaires-descriptions-relativisées 'locales' *Di*, *simplifiées à l'état d''étiquettes' i* (comme dans le cas des descriptions *D(r)* de *III.2.7*). Selon le modèle interprétatif d'une loi de probabilité installé dans le chapitre III, la loi de probabilité $\{p_i\}$, $i\equiv 1,2,\ldots.q$ de la définition de la source *S*, exprime une *forme globale F* constructible à partir de l'univers $U=\{Di\}$ par le procédé de déprobabilisation de *III.2.7*.

La forme globale F construite ainsi constitue un 'sens' global relativisé associable à la loi de probabilité de la source S. L'entropie $H(S)\equiv\Sigma_i p_i \log(1/p_i)$ ***de la source S est une fonctionnelle de ce 'sens' global de S. Via la forme*** $H(S)\equiv\Sigma_i p_i \log(1/p_i)$***, la syntaxe de la théorie de Shannon est fortement déterminée par ce 'sens' global de la source S*** [31].

En effet les théorèmes de la théorie de Shannon dépendent d'une façon *cruciale* de ce sens-***là***, un sens relativisé ***de la source S***, exprimé par la fonctionnelle *H(S)*. Je ne donnerai qu'un seul exemple, particulièrement simple et frappant.

Un théorème fondamental de la théorie de Shannon établit que la valeur numérique de *H(S)* définit la borne inférieure de la longueur moyenne de codage $L\equiv\Sigma_i p_i l_i$. En d'autres termes, quel que soit le codage accompli, on a toujours

$$L \geq [H(S)]/\log r$$

où r est le nombre de lettres de code x_j dans l'alphabet de code X et par conséquent $1/\log r$ n'est qu'un simple facteur numérique de proportionnalité. Donc *H(S)* introduit une indépassable limitation de la possibilité de raccourcir la longueur moyenne de codage des signes de source a_i. Cette limitation concerne *la dépense moyenne de* ***matière*** *porteuse de messages et de* ***temps*** *de transport.* C'est une dépense d'importance pragmatique.

La borne inférieure $[H(S)]/\log.r$, si elle ne peut pas être dépassée, peut quelquefois être atteinte strictement, à savoir lorsqu'il est possible de réaliser ce qu'on appelle un codage *spécial*. Un codage spécial est tel que la longueur l_i du mot de code X_i que l'on associe au signe de source a_i – forcément un nombre *entier et positif* – satisfait pour tout i l'égalité

[31] Lorsque *S* consiste en une langue courante et l'alphabet *A* de signes est un ensemble assez grand de mots de cette langue (disons, tous les mots usuels) la forme globale *F* correspondant à la loi de probabilité d'apparition des mots dans un très grand ensemble de messages formés avec ces mots, introduit une topologie globale de la langue en question : certains mots, 'mère' et 'amour' y seront sans doute plus voisins l'un de l'autre que 'mère' et 'haine', etc. Une telle forme globale d'une langue donnée est un construit intéressant en soi : il exprime probablement des traits psychologiques spécifiques de la population qui parle cette langue.

$$l_i \equiv 1/log_r\, p_i$$

Évidemment cette condition n'est pas réalisable toujours dans la forme stricte requise par l'égalité $l_i \equiv 1/log_r p_i$, car une fois que les p_i et r sont posés, le rapport $1/log_r\, p_i$ n'est pas nécessairement toujours un entier. Mais quand cela *est* possible et que la condition est respectée par un choix convenable des longueurs de mots l_i, alors on atteint effectivement le codage le plus bref qui est imaginable ***avec la source S** et l'alphabet de code X dont on dispose*. C'est à dire, en ce cas heureux on a l'égalité

$$L \equiv [\Sigma_i\, p_i\, log\ (1/p_i)]/log.r \equiv [H(S)]/log\ r$$

Cette valeur de L est celle qui réalise *l'économie maximale de matière porteuse des messages et de temps dépensés*, compatible avec les moyens que l'on s'est donnés.

Mais ***pourquoi*** ? Quoi de tellement singulier y a-t-il dans cette sorte de codage 'spécial' ? Une fois la question soulevée, la réponse est identifiée facilement. Il s'agit là d'un *double* codage simultané. Cependant que les lettres de code $x_k x_s....x_m$ choisies dans le mot de code $X_i \equiv x_k x_s....x_m$ codent pour le signe a_i qu'on veut communiquer à un interlocuteur, la longueur l_i du mot X_i, *via* la relation $l_i \equiv 1/log_r\, p_i$, code elle aussi, de son côté, pour *autre chose*, à savoir pour la probabilité p_i du signe a_i. Et elle accomplit ce codage-là de la manière qui 'assortit' au maximum, d'une part la prédisposition structurelle globale que la source S possède pour émettre le signe a_i, avec, d'autre part, la longueur l_i du mot de code $X_i \equiv x_k x_s....x_m$ choisi pour transmettre le sens assigné au signe a_i : *Le codage est le plus bref pour les signes a_i ayant la probabilité p_i la plus grande, et le plus long pour les signes a_i ayant la probabilité p_i la plus petite, etc.* La relation particulière entre réalisée dans un codage spécial entre la longueur de codage du signe a_i transmis, et sa probabilité, élimine strictement toute longueur de codage inutile.

Les probabilités $\{p_i\}$, $i \equiv 1,2,...q$, sont utilisées comme un deuxième alphabet émis par S, et les *longueurs l_i* des mots X_i jouent le rôle d'un second alphabet de codage $Y \equiv \{y_s, s \equiv 1,2...t\}$ pour les signes de ce deuxième alphabet, chaque signe p_i devant toujours être codé par un 'mot' Y_i d'une seule lettre $y_s \equiv l_i$, avec l_i la longueur du mot X_i qui code pour le signe a_i. Ainsi, cependant que le mot de code X_i 'dit a_i' par son contenu $x_k x_s....x_m$ de lettres de code, il 'dit p_i' par sa propre longueur $l_i \equiv 1/log_r p_i$. ***Une sorte de double codage auto-référentiel***.

Mais en 'disant' p_i par sa longueur, le mot de code X_i reflète de manière cryptique un trait de la forme globale liée à la loi entière de probabilité $\{p_i\}$, $i \equiv 1,2,...q$, dans l'espace de représentation de la vue V comportée par le référentiel (C,V) de l'arbre de probabilité $\mathbf{T}(C,V)$ équivalent à la donnée de S. En effet selon le postulat que nous avons posé concernant l'interprétation du concept abstrait de mesure de probabilité, $p_i \leftrightarrow l_i$ *désigne le nombre d'interventions, dans la forme globale F liée à la loi $\{p_i\}$, $i \equiv 1,2,...q$, de l'événement- élémentaire-description-relativisée de probailité p_i*. Il reflète ce trait de la forme globale liée à la loi entière $\{p_i\}$, $i \equiv 1,2,...q$ comme une goutte d'eau jetée par une vague sur une roche peut refléter la couleur de l'océan tout entier d'où elle provient. Car la probabilité p_i ne possède *pas* une existence isolée de celle des autres probabilités p_j, $j\neq i$ de la loi entière de probabilité $\{p_i, j\equiv 1,2,...q\}$. Elle fait corps organiquement avec l'entière loi $\{p_i, j\equiv 1,2,...m\}$, ***en conséquence de la condition de normation*** $\boldsymbol{\Sigma_i p_i \equiv 1}$. Ainsi le signe p_i est porteur d'un élément du *sens* ***global*** de *S* exprimé de façon morcelée et cryptique par la loi de probabilité $\{p_i, j\equiv 1,2,...q\}$.

On perçoit la relation flagrante avec le 'problème' signalé dans *III.2.7*, à savoir le fait que les rapports $n(Dr)/n$ de l'ensemble $\{n(Dr)/n\}$, $r=1,2,...s$ qui y intervenait, sont des nombres rationnels, tandis que la loi de probabilité factuelle $\{p\ n(Dr)\}$ à introduire dans une 'situation probabiliste', représentée à l'aide de ces rapports, est un ensemble de nombres réels. Comme *tout* ce qu'on connaît de manière communicable est description relativisée, et *toute* telle description peut être regardé comme le 'codage d'un sens relativisé', on voit que :

> ***L'algorithme développée dans le chapitre III, de définition de la loi factuelle de probabilité à affirmer dans une situation probabiliste donnée, précise la signification du théorème*** $\boldsymbol{L\geq[H(S)]/\log r}$ ***tout en insérant ce théorème dans une représentation GENERALE des connaissances communicables, c'est à dire de 'descriptions', de morceaux de sens relativisés***.

Dans la théorie de Shannon, la signification portée par la longueur $l_i\equiv 1/\log_r p_i$ du mot de code X_i concernant la forme globale liée à *S*, *n'est pas* ***voulue*** *en tant qu'un élément de message, à communiquer d'une façon* ***énoncée***. Elle n'est qu'un moyen muet pour le but utilitaire, latéral à la transmission des messages, de minimiser les dépenses de matière et de temps nécessaires pour cette transmission. Néanmoins cet élément de sens crypté porté par l'expression $l_i\equiv 1/\log_r p_i$, bien que muet, bien que dépourvu de toute possibilité de décodabilité aménagée dans la théorie, se trouve ***là*** lui aussi, encodé dans le message transmis.

> *Le message transmis possède une composante de sens* ***à communiquer*** *d'une manière* ***énoncée et*** *une* ***autre*** *composante* ***à SENS MUET*** *mais qui est* ***déterminante d'un point de vue pragmatique***.

IV.3. La théorie de Shannon *versus* 'information'

Celle-ci est la situation dans la théorie de Shannon proprement dite, celle qui traite exclusivement de la transmission de messages conçus par un émetteur *conscient* à l'intention d'une receveur *conscient*. Mais dans les tentatives tellement diverses d'application de la structure syntaxique de cette théorie, à d'autres domaines, tout à fait *différents* de celui de la communication délibérée de messages, l'existence d'un sens comporté implicitement par la source *S* joue souvent un rôle *explicite*, direct et ***majeur***, d'une nature tout à fait distincte du rôle muet est latéral, exclusivement pragmatique, qui lui est assigné dans la théorie de Shannon. Corrélativement, en de tels cas les conditions impliquées, et même les *éléments* requis dans la syntaxe de Shannon, sont modifiés plus ou moins explicitement, ce qui conduit à des ambiguïtés et des confusions.

Par exemple, une mesure physique dont les résultats sont dispersés, est quelquefois traitée comme un canal de transmission de 'messages' qui ferait parvenir à une conscience humaine un message venant 'de la part' d'une entité-objet-d'étude *physique* traitée comme une 'source d'informations'. Les signes de source sont parfois *entièrement* ***inconnus*** et donc on ne les code pas [32]. Le physicien théoricien code a priori *tout* signe qui peut être ***reçu*** à travers le 'canal' de la mesure, en employant un alphabet de codage lié à la syntaxe d'une théorie physique. Le physicien expérimentateur, lorsqu'il reçoit les signes codés de cette façon – des marques physiques observables sur l'enregistreur de quelque appareil de mesure – les 'lit' dans le langage codant de la même théorie physique. En ce cas donc – contrairement à ce qui se passe dans la syntaxe de Shannon – ***c'est bien le 'sens' comporté PAR LA SOURCE S qui détermine le CONTENU SEMANTIQUE des 'messages'***.

Rappelons dans ce contexte, en tant qu'un second exemple, que les applications de la syntaxe de Shannon, à des estimations de la complexité d'une 'entité' ***individuelle*** donnée (d'habitude de nature bio-chimique ou biologique, cellule organique, organe, animal) [33] – qui ont scandalisé Kolmogorov et Chaitin – reposent elles aussi sur le sens implicite, muet, d'une 'source de signes' *S*, qu'on peut associer à une description individuelle de cette entité, en la *'probabilisant'* et en l'exprimant à l'aide d'une structure mathématique entropique.

Dans les deux exemples mentionnés plus haut le rôle de l'émetteur de messages coalesce avec le rôle de source *S* de signes : il n'y a pas d'émetteur conscient de messages conçus délibérément. On voit que l'organisation conceptuelle qui sous-tend la syntaxe de Shannon se trouve subrepticement déformée comme un matériau élastique soumis à des tensions.

Un autre exemple, courant dans la littérature des sciences chimiques, bio-chimiques, économiques, est celui de l'interaction entre deux entités matérielles représentées comme des 'agents' qui transmettent l'un vers l'autre des 'messages', quelquefois à travers un 'milieu' auquel on assigne le rôle d'un 'canal de transmission'.

On imagine facilement toutes les déformations implicites auxquelles on soumet la syntaxe de Shannon dans toutes ces 'applications'.

Enfin, plus généralement, en l'absence de relativisations descriptionnelles explicites, la syntaxe de Shannon, comme celle de Kolmogorov, reste entachées de flous rédhibitoires qui entachent toutes les applications.

La méthode de conceptualisation relativisée peut être elle-même regardée comme une représentation qualitative, mais *formalisée* et *générale*, de la genèse et la communication de morceaux de sens élaborés délibérément et explicitement par un concepteur : des *descriptions* relativisées, donc *des* ***'informations'*** *au sens propre du terme*. En combinant cette méthode avec la syntaxe de Shannon et en distinguant toujours clairement entre source de signe et message à communiquer, l'on pourra élaborer un ensemble de mathématisations très intéressantes de certains points de vue *pragmatiques*, mais aussi conceptuels,

---

[32] C'est le cas lors de mesures sur un microétat.

[33] En fait, d'une *description* individuelle d'une entité biologique donnée.

puisque dès qu'on décrit on 'code'. L'on obtiendra ainsi une structure conceptuellement bien dominée, souvent mathématisée,une véritable théorie de l''information' incorporant la théorie de la transmission de messages à contenus sémantiques non-spécifiées. Cette théorie pourra faire face pleinement à l'ensembles des domaines auxquels on a voulu appliquer telle quelle le théorie de la communication de messages de Shannon, sans vraiment y réussir.

V. ESTIMATIONS DE COMPLEXITE SELON *MCR*

On vient de re-fonder d'une façon bien définie, dans le cadre de *MCR*, l'utilisation des concepts de probabilité et d'entropie probabiliste. Ces concepts sont maintenant disponibles pour élaborer des modes d'estimation de *'complexités **relativisées**'*. Une méthode-*MCR* spécifiquement appropriée à ce but n'est pas encore entièrement élaborée. Mais quelques principes d'une telle méthode – tout à fait nouvelle dans ce domaine – sont clairs dès maintenant. Dans ce qui suit j'esquisse ces principes.

V.1. Quelques principes concernant l'estimation de 'complexités relativisées'

V.1.1. L'entité-objet d'une estimation de complexité relativisée. Selon *MCR* le concept de 'complexité' possède foncièrement le statut d'une ***méta-****description via* une *méta-**vue** de 'complexité'*, d'une entité-objet consistant en une ***description*** accomplie précédemment : seule une description déjà accomplie peut 'exister' relativement aux qualifications de complexité. Je m'explique.

Imaginons une feuille de papier sur laquelle est inscrit un relevé de compte bancaire. Quelle est la complexité de cette entité-objet ? Dans *l'absolu* cette question n'est pas définie, elle est tout simplement dépourvue de sens. Tout d'abord, personne ne percevra 'une feuille de papier sur laquelle est inscrit un relevé de compte bancaire' *en dehors de* ***toute*** *vue*. La locution entre guillemets est déjà une description impliquant une vue relativisante. C'est une description énoncée dans le langage courant et portant donc toutes les marques trompeuses d'une désignation qui indiquerait un 'objet' par des 'propriétés intrinsèques' qu'il possèderait de façon intrinsèque, indépendamment de toute relativité à des actions descriptionnelles. Mais en fait d'ores et déjà il s'agit d'une description, *pas d'une entité-objet non-décrite*. Un bébé d'un an considérera l'entité-objet de cette description, à travers une vue qui n'introduira que certains aspects physiques (la forme, les couleurs, le nombre et la disposition des signes). *Via* cette vue-là, il en percevra – de manière non-explicite, bien entendu – une certaine description de cette entité-objet, différente de celle indiquée par la locution adulte 'relevé de compte bancaire'. Un psychologue arriverait probablement à identifier la description du bébé. Il pourra même peut-être la méta-qualifier en termes de complexité. Il essaierait probablement d'abord d'estimer séparément la complexité de la description du bébé *via* l'aspect 'forme', puis d'en estimer aussi la complexité face à l'aspect couleur', etc.. Et s'il dispose de principes pour totaliser des complexités différentes d'une et même entité-objet-description-relativisée, il pourra finalement tenter d'intégrer *une* méta-description de complexité relativisée globale. Mais imaginons un adulte français normal à qui on donnerait la même feuille. Il *pourrait* la considérer *via* la même vue que le bébé. Mais il est vraisemblable qu'il ne le fera pas. En un certain sens il ne verra même pas la description du bébé. Il considérera cette entité-objet plus ou moins exclusivement à travers une vue dont le bébé ne dispose pas et qui *filtre* les aspects physiques pour ne retenir que des aspects 'bancaires' (débit, crédit, etc.). Si on lui proposait donc d'estimer la complexité de cette 'feuille', il considérerait probablement les ***descriptions*** relatives aux aspects de la vue bancaire qui se manifestent dans le texte inscrit sur la feuille (les informations bancaires de divers types, leur organisation, etc.). Puis il chercherait à former des *méta*-aspects de complexité de ***ces*** descriptions relativisées *là*, et d'estimer les complexités de celles-ci face à ces méta-aspects, puis face à l'entière méta-vue qu'ils constituent ensemble. Donc en ce cas aussi la complexité recherchée se rapporterait à une *description* accomplie précédemment, la description bancaire. Enfin, imaginons un physicien qui vient d'engendrer par une opération G de génération de microétat, un microétat me_G qui n'a encore jamais été soumis à quelque opération de mesure à effets observables. Si l'on demandait à ce physicien : « quelle est la complexité du microétat me_G ? », que répondrait-il ? Il dirait sans doute : « je ne sais pas encore, pour répondre je dois d'abord décrire ce microétat par des mesures, et ensuite je verrai comment exprimer de la complexité correspondante ». Dans ce cas également, l'estimation de complexité sera spontanément placée sur un *méta*-niveau descriptionnel. Bref :

L'entité-*objet* d'une description de complexité est une ***description*** *accomplie précédemment*. Une description *de complexité* possède foncièrement le statut descriptionnel d'une ***méta***-description.

La mesure de complexité d'une entité-objet n'est pas une caractéristique *en soi*, 'ontologique', de cette entité. C'est une caractéristique *épistémologique* foncièrement soumise à des relativités descriptionnelles.

Parler de 'la' complexité d'une entité-objet, sans autres spécifications, est une fausse absolutisation. C'est du NON-SENS.

V.1.2. 'Emergence' ou 'exmergence'. Selon *MCR* une 'émergence' est la dénomination qu'on donne à un certain type de *changement* d'une entité-objet-*description-relativisée*. C'est donc une *méta*-qualification d'une *description* accomplie précédemment, comme dans le cas d'une qualification de complexité. Considérons un exemple. Supposons un pot qui contient une plante. J'ai regardé cette plante hier matin et je conserve en mémoire une description exacte de l'apparence qu'elle avait face au référentiel épistémique *(G,V)* qui était alors en action dans mon esprit (*G* : le générateur d'entité-objet – le champ de mon attention – qui introduisait en tant qu'entité-objet le pot avec la plante ; *V* : ma vue biologique contenant les aspects, disons, de formes-de-couleur et d'odorat). Ce matini je regarde de nouveau le pot avec la plante et je vois un bourgeon qui n'existait pas dans la description enregistrée hier à l'aide de *(G,V)*. Je peux me dire : « tiens, quelle émergence soudaine ! », et ce sera une *méta*-

description de différence entre la nouvelle *description* qui s'accomplit aujourd'hui dans mon esprit, et la *description* d'hier.

Dans cet exemple j'ai supposé que mes sens m'ont permis de mettre en jeu *spontanément* des vues-aspect nouvelles qui n'intervenaient pas dans la vue *V* de *(G,V)*, par exemple une vue-aspect de toucher qui me permet de sentir que ce bourgeon est si neuf qu'il est encore un peu humide et collant. Mais s'il s'agissait de la répétition d'une analyse du sang chez un patient donné, tel ou tel aspect nouveau de la vue qui serait nécessaire pour pouvoir percevoir les éventuelles émergences face à l'analyse précédente, ne serait en général pas toujours automatiquement disponible et *active* : c'est l'une des sources majeures des difficultés de diagnostic.

La perceptibilité d'une émergence, donc la descriptibilité de *ce* qui émerge, sont foncièrement relatives aux référentiels épistémiques dont dispose l'observateur-concepteur, et de l''intuition' qui détermine chez lui le choix de mettre en œuvre tel référentiel épistémique plutôt qu'un autre. Toute l'immense question de l'aptitude à percevoir des émergences – des questions, des problèmes, des traitements de problèmes – est foncièrement relative aux référentiels épistémiques dont on dispose, *ou de l'idée et la capacité de les construire*.

Une émergence peut aussi être 'négative'. Elle peut consister dans la *disparition* d'un élément d'une description précédente, relativement au référentiel *(G,V)* où celle-ci avait été construite. En ce cas on pourrait parler de 'exmergence'. Lorsqu'il s'agit d'une émergence on peut parler de *complexification* relative à la description précédente, et lorsqu'il s'agit d'une exmergence on peut parler de *simplification* de cette description.

> Les concepts d'émergence ou d'exmergence sont eux aussi ***foncièrement*** *relatifs, épistémologiques,* ***non-ontologiques***. Emergence ou exmergence dans l'absolu, dans un sens ontologisé, sont *du non-sens*.

V.1.3. Méta-vue de complexité. Comment définir une méta-vue de complexité à introduire dans le référentiel épistémique où s'accomplit une méta-description de complexité, ou de complexification (émergence), ou de simplification (exmergence) ? Dénotons *a priori* par $V(\mathbf{c})^{(2)}$ une méta-vue de complexité. De quelles sortes de méta-***aspects***-de complexité convient-il de munir $V(\mathbf{c})^{(2)}$? Cette question n'est pas simple. Disons-en d'abord ce qui peut être dit tout de suite avec certitude. Quelle que soit une méta-vue-*aspect* appartenant à $V(\mathbf{c})^{(2)}$, – dénotons-la $V(\mathbf{c}a)^{(2)}$ – en accord avec la définition générale d'une vue-aspect *quelconque*, elle doit donc comporter : ***(a)*** un méta-aspect de complexité $(\mathbf{c}a)^{(2)}$ muni d'un nombre fini de 'valeurs' bien définies, et ***(b)*** la spécification d'un examen de complexité correspondant – un $(\mathbf{c}a)^{(2)}$-examen – qui soit *effectif* et qui incorpore une règle explicite spécifiant pour tout résultat possible du $(\mathbf{c}a)^{(2)}$-examen, comment il est à *coder* en termes de telle ou telle valeur bien définie de $V(\mathbf{c}a)^{(2)}$.

Mais comment construire les méta-vues-*aspect* $V(\mathbf{c}a)^{(2)}$ dont il convient de munir une méta-vue de complexité $V(\mathbf{c})^{(2)}$? L'entité-*objet* d'une méta-description de complexité est une description relativisée $D/G,œ_G,V/$ accomplie précédemment. On pourrait donc se dire au premier abord qu'il faut estimer successivement les complexités de G, de $œ_G$, et de V, et totaliser de quelque façon ces estimations partielles, afin d'estimer la complexité globale de $D/G,œ_G,V/$. Mais dès qu'on s'est dit cela, on s'arrête. Car on vient de montrer que la complexité d'une entité-objet qui n'est *pas* une entité-objet-***description***-*relativisée*, est du non-sens. Donc *en général* on ne peut pas estimer la complexité de l'élément $œ_G$. Toutefois on pourrait estimer la complexité correspondante à l'opération de génération d'entité-objet G, car celle-ci est supposée être toujours décrite, et d'une manière communicable et inter-subjective. Mais selon cette voie on initierait une régression sans fin : pour estimer la complexité d'une *description* de départ $D/G,œ_G,V/$, il faudrait commencer par estimer la complexité de [la *description* de l'opération G qui agit dans la description $D/G,œ_G,V/$], etc.. Or *MCR* bannit toute régression infinie, par méthode.

Dans MCR on est donc acculé à construire une méta-vue de complexité $V(\mathbf{c})^{(2)}$ *dont les méta-vues-aspects* $V(\mathbf{c}a)^{(2)}$ *sont toutes relatives* ***exclusivement*** *aux aspects g de la vue V qui agit dans l'entité-objet-description-relativisée* $D/G,œ_G,V/$.

Cette conclusion est non-triviale. Elle conduit à re-noter, pour tout aspect g appartenant à la vue V, : $V(\mathbf{c}a)^{(2)} \equiv V(\mathbf{c}g)^{(2)}$ et $[(\mathbf{c}a)^{(2)}$-examen$] \equiv [(\mathbf{c}g)^{(2)}$-examen$]$. Ainsi l'on dote la méta-vue de complexité, d'un méta-aspect de complexité $\mathbf{c}g$ correspondant à chaque aspect g de V. Comment, maintenant, définir, pour une description relativisée $D/G,œ_G,V/$ donnée, la *valeur* de complexité déterminée par un méta-aspect $V(\mathbf{c}g)^{(2)}$ de complexité et par le $(\mathbf{c}g)^{(2)}$-examen correspondant, lorsque l'aspect g contribue à la vue V de pour une description relativisée $D/G,œ_G,V/$? En d'autres termes, comment *mesurer* la complexité d'une description relativisée $D/G,œ_G,V/$ face à l'aspect g qui y intervient ?

V.1.4. Complexité d'une description relativisée $D/G,œ_G,V/$.

A. Supposons d'abord que la description considérée est une description *probabiliste* $D/G,œ_G,V/$ d'une entité-objet $œ_G$ physique, i.e. que les reproductions d'une séquence $[G.V]$ ne produisent pas toutes une et même valeur gk de g pour tout aspect g qui intervient dans la vue V. Alors pour tout aspect g qui intervient dans la vue V, la description $D/G,œ_G,V/$ affirme une loi de probabilité, disons $\{p(Dgk)\}$, $k=1,2,\dots q$, posée sur l'univers $\{Dgk\}$, $k=1,2,\dots q$ des événements-élémentaires-descriptions-étiquettes [34]. Alors on peut procéder de la façon suivante. On introduit dans la méta-vue de complexité $V(\mathbf{c})^{(2)}$ à

[34] Aux notations près l'univers $\{Dgk\}$, $k=1,2,\dots q$ est l'équivalent des univers $\{Dj\}$, $j=1,2,\dots q$ ou $\{Dr\}$, $r=1,2,\dots s$ du chapitre III.

utiliser, un méta-aspect 'entropique' de complexité à savoir *($\mathcal{C}g)^{(2)}$* $\equiv$ *[$\Sigma_k p(Dgk)\log(1/p(Dgk))$, $k=1,2,\ldots q$]* où le contenu du deuxième membre de la définition peut-être dénoté plus synthétiquement *$H(D(æ_G)/Vg)$* et dénommé *la fonctionnelle de complexité de la description $D/G,æ_G,V/$ de $æ_G$ relativement à l'aspect g*. Les valeurs de ce méta-aspect entropique de complexité sont par définition les valeurs ***numériques*** de la forme entropique *$\Sigma_k p(Dgk)\log(1/p(Dgk))$, $k=1,2,\ldots q$*.

L'on obtient pratiquement le même ensemble de *q* valeurs de complexité face à *g* de la description *probabiliste $D/G,æ_G,V/$*, si l'on *déprobabilise $D/G,æ_G,V/$* selon la procédure indiquée dans le chapitre III ; c'est à dire, si l'on construit la forme globale *F* qui correspond à *$D/G,æ_G,V/$*, en y estimant l'ensemble des rapports rationnels *$\{n(Dgk)/n_F(gk)\}$, $k=1,2,\ldots q$* [35] et en calculant les valeurs numériques des *q* formes entropiques *$\Sigma_k n(Dgk)/n_F(gk)\log(1/(n(Dgk)/n_F(gk))$, $k=1,2,\ldots q$*.

B. Supposons maintenant que *$D/G,æ_G,V/$* est une description relativisée *individuelle* d'une entité-objet physique. En ce cas, par définition, cette description consiste elle-même, directement, en une forme *F* déterminée dans l'espace de représentation de la vue *V* par des valeurs *gk* distribuées sur le domaine d'espace-temps incorporé par la vue *V*. On peut donc estimer directement sur cette forme *F* les valeurs des rapports *$n(Dgk)/n_F(gk)$, $k=1,2,\ldots q$* pour tous les aspects *g* qui interviennent dans la vue *V*, puis procéder comme au point A.

On peut ensuite essayer de 'totaliser' les estimations numériques de complexité d'une description relativement à un aspect g donné, proposées aux points A et B. Une telle totalisation peut être tentée en additionnant tout simplement toutes ces estimations concernant la même entité-objet *$æ_G$*.

On peut également explorer l'expressivité, pour ce but, ***d'entropies conjointes ou conditionnelles****, ce qui rapprocherait de l'intégration de la forme globale F associée à l'ensemble des lois de probabilité qui interviennent*. C'est de cette manière, en fait, que pourraient intervenir avec toutes leurs ressources, les différents concepts entropiques de la syntaxe de Shannon, *tous relativisés*.

Par la voie indiquée l'on peut élaborer un traitement *mathématique* des questions de complexité, qui soit ***foncièrement non-destructeur des contenus sémantiques*** [36].

Par contre, l'utilisation du concept shannonien d'entropies de sources de signes, en dehors du cadre de *MCR* où chaque pas est guidé, a une chance quasi-nulle d'échapper à tous les pièges de fausse absolutisation, d'ambiguïté, etc., qui guettent les démarches descriptionnelles spontanées.

V.1.5. Complexité d'une vue V. Une vue *V* est elle aussi une description : une description-*définition* du concept général de vue, adaptée au cas de la particulière considérée. On peut donc en

[35] Le même principe de notation que dans *III.2.7*.

[36] Une autre démarche fondée sur *MCR*, plus sophistiquée du point de vue mathématique, se trouve esquissée dans :

Schächter, V., "Complexity Measures Viewed Through the Method of Relativised Conceptualisation", in "Quantum Mechanics, Mathematics, Cognition and Action : Proposals for a Formalised Epistemology', M. Mugur-Schächter and A. van der Merwe, eds., *Kluwer Academic Publishers* (2003).

donner une méta-estimation de complexité. Pour commencer, on peut tenter d'estimer la complexité de V par la somme des aspects g qu'elle introduit, chaque aspect étant multiplié par le nombre de valeurs gk correspondant : $\Sigma_g\Sigma_k gk$. Un tel départ apparaîtra probablement comme simpliste lors de l'examen de cas spécifiés. Par la suite on pourrait introduire aussi d'autres critères, par exemple la distance entre deux valeurs gk successives, lorsqu'une telle distance est définie pour V, etc. Cela peut paraître arbitraire. Mais notons bien que *toute manière de mesurer ou de 'repérer' des valeurs de quelque chose, se construit plus ou moins arbitrairement* : on se *donne* une unité ou une échelle, on se *donne* les modes d'utilisation de cette unité, etc.

V.1.6. Complexité de la description relativisée d'une opération de génération G d'une entité-objet-de-description. Quant l'entité-objet de la méta-description de complexité est, en particulier, la description relativisée d'une opération de génération *G* d'une entité-objet-de-description – considérée séparément de la description où *G* intervient – alors on peut en estimer la complexité selon l'esquisse générale du point *V.1.4.*.

VI. CONCLUSION

Selon *MCR* l'entité-objet d'une qualification de complexité possède *systématiquement* le statut d'une ***DESCRIPTION*** *relativisée*. En d'autres termes, une estimation de complexité, qualitative ou numérique, possède toujours le statut d'*une **méta-description** relativisée de la complexité d'une **description** (d'une entité-objet de départ)*. Lorsque les deux formalismes, probabiliste et shannonien, sont incorporés à *MCR* via des relativisations descriptionnelles systématiques, ces formalismes s'unifient et s'épurent de toute ambiguïté (cf. la réf. 16 E). Ceci dote les estimations quantitatives de complexité accomplies dans *MCR*, de nuances et de précisions qui leurs faisaient défaut dans les formulations ensemblistes classiques, absolutisées « afin de maximiser la généralité ». On peut donc prévoir que la démarche délinéée ici conduira à la construction de mesures de complexité aussi non-réductrices qu'on voudra, d'autant plus que le caractère du concept *MCR* de 'vue' est entièrement non-restrictif, ce qui permet d'introduire toute dimension sémantique souhaitée. Il deviendra ainsi possible enfin d'associer aux innombrables analyses profondes d'Edgar Morin [37] et de Jean-Louis Le Moigne [38], des instruments conceptuels quantitatifs de précision, qui n'en amputeront pas le *sens*. Et il sera peut-être possible également, de doter d'outils véritablement scientifiques l'approche pionnière de Georges-Yves Kervern [39] dans le domaine de l'étude des situations de danger, éminemment complexes. Ce serait un progrès notable dans le développement de 'la pensée complexe'.

---

[37] Morin, E., "La Méthode", t. I, II, III, IV, *Seuil*, 1977-1991.
[38] Le Moigne, J-L., "La théorie du système général, théorie de la modélisation", *PUF*, 1977 ; in " Modélisation de systèmes complexes", *Dunod*, 1990.
[39] Kervern, G.Y., "Latest Advances in Cindynics", *Economica*, 1994.