

**Contribution à l'étude de  
quelques mesures du contenu d'information des objets**  
*Vers une théorie multicruciale de l'information*

Gérard Pinson

(manuscrit version 1.0)

email : [pinson.g@wanadoo.fr](mailto:pinson.g@wanadoo.fr)

## **Avant-propos**

La mesure du contenu d'information d'un message ou d'un objet peut s'effectuer à l'aide de grandeurs définies sous différents angles théoriques : quantité statistique d'information au sens de Shannon, entropie de Von Neumann en théorie quantique, complexités algorithmiques ou complexités de Kolmogorov de chaînes alphabétiques, profondeur logique au sens de Bennet, voire spectre de Fourier en analyse harmonique, etc. L'analyse comparée des quantités d'information statistique, quantique, algorithmiques, montre qu'il est nécessaire de préciser les notions qui sont à la base du concept de contenu d'information, en cherchant à dégager les invariants apparaissant dans les différentes théories mais aussi leurs dissemblances. Le concept de contenu d'information est une notion délicate à manier : de la confrontation entre les différentes théories on peut espérer dégager les caractéristiques et propriétés essentielles de cette notion telle qu'elle apparaît à l'analyse, et peut-être ainsi rejoindre ses caractères principaux tels qu'ils apparaissent à l'expérience.

## Table des matières

Liste des symboles

Introduction

### Section 1

INTRODUCTION : CONTENU D'INFORMATION & CONTENU D'IDENTIFICATION

- 1.1. Analyse ternaire de l'information
- 1.2. Structure ternaire de l'information et sémantique
- 1.3. Information et compressibilité
- 1.4. Résumé : caractéristiques de l'information et des processus informationnels
- 1.5. Conclusion : vers une approche multicruciale du concept d'information

### Section 2

DÉFINITIONS

- 2.1. Les objets
- 2.2. Les symboles
  - 2.2.1. Du symbole au message
  - 2.2.2. Source symbolique
- 2.3. Les mots
- 2.4. Les textes
  - 2.4.1. La transcription alphabétique des messages en textes
  - 2.4.2. Source alphabétique
  - 2.4.3. Source lexicale
  - 2.4.4. Source faible
  - 2.4.5. Source préfixée
  - 2.4.6. Transcodage d'une source alphabétique vers une autre
  - 2.4.7. Multicrucialité des textes
  - 2.4.8. Résumé : classification des sources
- 2.5. Extensions au continu de la théorie de l'information
- 2.6. Conclusion : sources conjointes

### Section 3

SOURCES SYMBOLIQUES CLASSIQUES

- 3.1. Introduction
- 3.2. Entropie statistique d'une source unique
  - 3.2.1. Approche intuitive
  - 3.2.2. Approche axiomatique
  - 3.2.3. Calcul approché
  - 3.2.4. Propriétés
- 3.3. Entropies statistiques de sources conjointes
  - 3.3.1. Information composée

- 3.3.2. Information mutuelle
- 3.3.3. Écart entropique
- 3.4. Exemple : structure ternaire et théories de l'information
  - 3.4.1. Point de vue du récepteur : analyse statistique du message
  - 3.4.2. Point de vue de l'émetteur : construction algorithmique du message
- 3.5. Conclusions
  - 3.5.1. Interprétation ensembliste
  - 3.5.2. Structure heuristique de la théorie statistique classique
  - 3.5.3. Résumé des propriétés de l'entropie discrète
- 3.6. Extension au continu : entropie différentielle
  - 3.6.1. Source infinie discrète ou source dénombrable
  - 3.6.2. Entropie différentielle d'une source échantillonnée
  - 3.6.3. Entropies différentielles de sources échantillonnées conjointes
  - 3.6.4. Conclusions
    - 3.6.4.1. Quantification
    - 3.6.4.2. Résumé des propriétés de l'entropie différentielle

#### *Section 4*

#### SOURCES SYMBOLIQUES QUANTIQUES

- 4.1. Introduction
  - 4.1.1. Rappel succinct des axiomes de la mécanique quantique
    - 4.1.1.1. États
    - 4.1.1.2. Observables
    - 4.1.1.3. Mesure
    - 4.1.1.4. Dynamique
  - 4.1.2. Bit quantique
  - 4.1.3. Paires corrélées de bits quantiques
    - 4.1.3.1. États quantiques d'une paire corrélée. Base de Bell
    - 4.1.3.2. Création d'une paire corrélée
    - 4.1.3.3. Exemple : photons corrélés
- 4.2. Entropies quantiques d'une source unique
  - 4.2.1. Matrices de densité
  - 4.2.2. Entropie de Von Neumann
  - 4.2.3. Propriétés de l'entropie de Von Neumann
  - 4.2.4. Théorème de Holevo
- 4.3. Sources symboliques quantiques : entropie classique quantique
  - 4.3.1. Exemple
  - 4.3.2. Structure heuristique de la théorie classique quantique
- 4.4. Sources faibles quantiques : entropie quantique quantique
  - 4.4.1. Sources corrélées
  - 4.4.2. Entropies de Von Neumann de sources corrélées
  - 4.4.3. Exemple

#### 4.5. Conclusion : résumé des propriétés de l'entropie de Von Neumann

### Section 5

#### SOURCES ALPHABÉTIQUES

##### 5.1. Introduction

###### 5.1.1. Source alphabétique lexicale

###### 5.1.1.1. Description

###### 5.1.1.2. Propriétés

###### 5.1.2. Source alphabétique préfixée

###### 5.1.2.1. Description

###### 5.1.2.2. Propriétés

###### 5.1.3. Transcodage d'une source lexicale vers une source préfixée : exemples

##### 5.2. Complexités algorithmiques des sources alphabétiques

###### 5.2.1. Présentation générale et classification

###### 5.2.2. Complexité simple d'une source lexicale

###### 5.2.2.1. Source unique

###### 5.2.2.2. Sources conjointes

###### 5.2.3. Complexité uniforme d'une source préfixée

###### 5.2.3.1. Source unique

###### 5.2.3.2. Sources conjointes

###### 5.2.4. Complexité préfixée d'une source lexicale

###### 5.2.4.1. Source unique

###### 5.2.4.2. Sources conjointes

###### 5.2.4.3. Complexité préfixée explicite d'une source lexicale

###### 5.2.5. Complexité monotone d'une source préfixée

##### 5.3. Extension au continu des complexités préfixées

##### 5.4. Résumé

###### 5.4.1. Structure heuristique de la théorie algorithmique

###### 5.4.2. Résumé des propriétés des complexités algorithmiques

##### 5.5. Conclusion : énumérabilité du message et du contexte

###### 5.5.1. Complexités des processus informationnels

###### 5.5.1.1. Problématique générale

###### 5.5.1.2. Exemple

###### 5.5.2. Trois classes de calculabilité

### Section 6

#### SOURCES ÉCHANTILLONNÉES

6.1. Introduction : classification élémentaire de quelques formalismes du traitement de l'information

##### 6.1.1. Fini dénombrable

###### 6.1.1.1. Niveau primaire

###### 6.1.1.2. Niveau booléen

##### 6.1.2. Infini dénombrable

- 6.1.2.1. Niveau arithmétique
- 6.1.2.2. Niveau algorithmique
- 6.1.3. Infini non dénombrable
  - 6.1.3.1. Niveau harmonique
  - 6.1.3.2. Niveau topologique
- 6.1.4. Résumé
- 6.2. Structure algébrique (  $C, \delta, \mathbf{1}, *, \cdot, \mathcal{F}, =$  )
  - 6.2.1 Restriction du formalisme harmonique au calcul booléen
    - 6.2.1.1. Échantillonnage et quantification : discussion préliminaire
    - 6.2.1.2. Calculs
  - 6.2.2. Extension du calcul booléen au formalisme harmonique
    - 6.2.2.1. Signe
    - 6.2.2.2. Position
    - 6.2.2.3. Précision
  - 6.2.3. Conclusion

*Section 7*

CONCLUSION GENERALE

Annexe 1

Annexe 2

Bibliographie

Résumé

## Liste des symboles

(les notations mathématiques usuelles sont omises)

$i$	indice (notamment : index de bibliothèque et d'alphabet)
$j$	indice (notamment : index de mot)
$k$	indice (notamment : index de message et de texte)
$k_C$	complexité préfixée explicite (complexité de Chaitin)
$l(.)$	longueur d'une chaîne (i.e. nb de caractères)
$m$	nombre de symboles dans un message = nombre de mots dans un texte
$n$	nombre de caractères dans un mot
$o$	élément d'un objet dénombrable
$o$	élément d'un objet non dénombrable
$p, q$	probabilité
$r, s$	mot
$\bar{s}$	chaîne inverse (l'ordre des caractères dans la chaîne est inversé)
$t$	texte
$v, w$	mot
$x, y, z$	caractère (variable de lettre) ; variable générique
$A$	alphabet
$A^*$	vocabulaire
$B$	alphabet binaire $\{0,1\}$
$\mathcal{B}$	écriture binaire pondérée des nombres entiers
$C$	source lexicale
$C_C$	complexité simple
$C_K$	complexité uniforme
$C_F$	complexité harmonique
$C_G$	complexité du graphe d'une fonction
$C$	ensemble des nombres complexes
$E$	émetteur
$E(..)$	encodage
$\mathcal{E}$	encodage
$\mathcal{F}$	transformation de Fourier (TF)

$F$	TF discrète (TFD)
$F$	TFD numérique (FFT)
$F_1$	FFT normalisée à 1
$H$	entropie statistique
$I(A)$	quantité (i.e. mesure du contenu) d'information d'un objet $A$
$I(A,B)$	contenu d'information de $A \cup B$
$I(A:B)$	contenu d'information mutuel de $A \cap B$
$I(A B)$	contenu d'information conditionnel de $A$ sachant $B$
$I_H$	information de Holevo
$K$	source préfixée
$K_C$	complexité préfixée
$K_K$	complexité monotone
$L$	longueur moyenne des mots d'un langage
$L$	langage
$\mathcal{L}$	application de $\mathbf{N}$ vers $\mathbf{B}^*$ (encodage lexicographique)
$\mathbf{M}$	observable (opérateur de mesure)
$N$	nombre d'éléments d'une bibliothèque
$\mathbf{N}$	ensemble des entiers naturels
$\mathbf{N}^+$	ensemble des entiers naturels $> 0$
$\mathcal{O}$	objet dénombrable
$\mathcal{O}$	objet non dénombrable
$\emptyset, \mathbf{I}$	valeurs booléennes 0, 1
$\mathcal{P}(\cdot)$	ensemble de parties
$Q$	nombre d'éléments d'un alphabet
$\mathbf{R}$	récepteur
$\mathbf{R}$	ensemble des nombres réels
$\mathbf{R}$	opérateur de rotation
$\mathcal{R}(\cdot)$	rotation
$\mathbf{S}$	source alphabétique
$\mathcal{T}(\cdot)$	transcription
$\dot{U}, \dot{V}$	vecteur dans $\mathbf{R}^n$
$X$	lettre

$\mathbb{Z}$	ensemble des entiers relatifs
$\chi, \pi$	symbole
$\delta$	distribution de Dirac
$\varepsilon$	chaîne vide
$\phi$	application
$\gamma, \mu, \sigma$	message
$\eta$	variable symbolique
$ \phi\rangle$	vecteur d'état
$\theta$	angle
$\rho$	matrice de densité
$\sigma_i$	matrices de Pauli
$\hat{\sigma}$	vecteur de spin
$\omega_{1:n}$	suite des $n$ premiers caractères d'une chaîne infinie
$\xi$	signe (variable symbolique)
$ \psi\rangle$	vecteur d'état
$\Phi$	machine de Turing, fonction récursive partielle
$ \Phi^\pm\rangle$	base de Bell
$\Gamma, \Sigma$	source symbolique
$\Lambda$	nombre d'éléments d'un objet
$\Pi, \Xi$	bibliothèque symbolique
$\Theta$	quantité totale d'information dans un message
$\Xi^*$	dictionnaire
$ \Psi^\pm\rangle$	base de Bell
$\equiv$	relation de correspondance entre chaînes et nombres entiers
$<$	ordre lexicographique croissant
$>$	ordre lexicographique décroissant
$\langle \cdot   \cdot \rangle$	produit scalaire
$ \cdot\rangle \langle \cdot $	projecteur

$\langle \cdot, \cdot \rangle$	encodage
$\otimes$	produit tensoriel
$\otimes$	corrélation
$*$	convolution
$*$	conjugaison complexe
$\mathbb{W}_a$	peigne de Dirac de pas $a$
$\square$	point fixe
$\lrcorner$	marque élémentaire
$\neg$	négation logique
$\diamond$	modalité
$\square$	modalité

## Introduction : résumé général

De même que l'énergie subit dans les processus physiques divers changements de forme, dans le schéma :

$$\text{objets} \rightarrow \text{symboles} \rightarrow \text{mots} \rightarrow \text{textes-source} \rightarrow \text{textes-code}$$

l'information apparaît sous différents états dont il convient d'analyser les transformations. En toute généralité, appelons *processus informationnel* de telles transformations (dont les processus communicants sont un cas particulier). Et, de même que les processus physiques s'appliquent à des systèmes physiques, convenons d'appeler *sources* les systèmes de traitement de l'information auxquels s'appliquent les processus informationnels.

*L'expérience montre que la discernabilité est une condition première de l'analyse de l'information, qui est une entité nécessairement "diacritique" [Oswald, 1986] c'est-à-dire distinguable (par un opérateur de disjonction). On considère donc tout d'abord un objet  $O$  en tant qu'ensemble fini ou infini dénombrable d'éléments discernables. A chaque élément  $o_i$  de l'objet on associe bijectivement un symbole  $\chi_i$ . L'ensemble des symboles, discernables par définition, constitue une bibliothèque  $\Xi = \{\chi_1, \dots, \chi_i, \dots\}$  avec  $N = \text{card}(\Xi)$  si la bibliothèque est finie. Une source symbolique  $\Sigma$  lit ou écrit des suites finies de symboles appelées *messages* ou *chaînes* :*

$$\sigma = \xi_1 \dots \xi_k \dots \xi_m ; \xi_k \in \Xi ; \sigma \in \Xi^m, \text{ longueur } l(\sigma) = m.$$

*Une source alphabétique comprend une source symbolique et un alphabet fini  $A = \{X_1, \dots, X_Q\}$  de valence  $Q$  ; elle encode chaque symbole en une suite finie de caractères formant un mot :*

$$s = x_1 \dots x_j \dots x_n ; x_j \in A ; s \in A^n, \text{ longueur } l(s) = n,$$

et chaque message en une suite finie de mots formant un *texte* :

$$t = s_1 \dots s_k \dots s_m ; t \in \text{vocabulaire } A^* = \bigcup_{n \geq 1} A^n$$

Si l'encodage est *singulier* (c'est-à-dire non injectif : un mot est associé à plus d'un symbole), la source alphabétique sera dite *source faible*. Sinon, un encodage régulier (i.e. injectif) est ou n'est pas déchiffrable. Dans la négative, la source produit des mots sans délimiteur, équivalents à des suites de nombres qu'il n'est pas possible de distinguer les uns

des autres. Cette source sera dite *lexicale*. Dans le cas contraire, la source produit des mots délimités, donc discernables. S'ils sont autodélimités (c'est-à-dire s'ils incluent leur propre délimiteur, que l'on supposera situé au début du mot), une telle source alphabétique déchiffirable est dite *préfixée*. Enfin, un transcodage est une application qui fait passer d'une source alphabétique à une autre.

*L'expérience montre que la bipolarité ou relativité est une autre condition première de l'analyse de l'information : un processus informationnel est effectif si les contenus d'information sont identifiables les uns par rapport aux autres, ce qui suppose qu'un processus informationnel réunisse conjointement (par un opérateur de conjonction) au moins deux contenus  $s$  et  $r$ , qui tiennent l'un pour l'autre le rôle de référence. Le concept d'identification de l'information peut faire l'objet d'une analyse épistémologique très générale, selon un schéma ternaire [Deman, 1975, Pinson, 1994] : à la donnée, c'est-à-dire au contenu d'information proprement dit, intrinsèque, est associée une référence ou contenu d'identification explicite. Cet identificateur est lui-même une donnée : un signal de synchronisation émis lors d'une transmission d'image par exemple mobilise une partie des moyens de transmission, et est sujet à erreur, comme la donnée. Donnée et identificateur sont de même type, qui est défini par un ensemble de ressources contextuelles constituant le contenu d'identification implicite de la donnée. Ce contenu, par définition, n'est pas échangé ni communiqué, et ne peut donc faire l'objet d'erreurs de transmission. C'est le cas par exemple de l'information contenue dans le protocole d'un processus communicant, qui assure la réception correcte des données grâce à l'émission d'identificateurs de référence (la compression d'information consistant d'ailleurs à rejeter dans le protocole un maximum d'information).*

Soient  $I(s)$  et  $I(r)$  les mesures des contenus d'information ou *entropies* ou *densités diacritiques* des chaînes  $s$  et  $r$  symbolisant les objets  $A$  et  $B$ ,  $I(s,r)$  l'information liée à  $s \cup r$ ,  $I(s:r)$  l'information mutuelle échangée ou partagée conjointement entre  $s$  et  $r$  dans  $s \cap r$ . L'entropie conditionnelle  $I(s|r)$  (à laquelle est attachée la représentation ensembliste  $s \cap \neg r$ ) est la quantité d'information contenue dans  $s$  sachant  $r$ . Nous partons des hypothèses suivantes :

- *symétrie* : la mesure du contenu mutuel d'information est une opération commutative :

$$I(s:r) = I(r:s)$$

- *additivité* : les contenus d'information s'ajoutent :

$$I(s,r) = I(r) + I(s|r) = I(s) + I(r|s)$$

$$I(s) = I(s:r) + I(s|r)$$

$$I(r) = I(s:r) + I(r|s)$$

On note respectivement :  $H$  : entropie d'une source symbolique classique ;  $S$  : entropie d'une source symbolique quantique ;  $C$  : entropie ou *complexité algorithmique* ou *complexité de Kolmogorov* d'une source lexicale ;  $K$  : entropie ou complexité d'une source préfixée. Un symbole étant une donnée "brute", à la fois insécable et supposée indépendante de l'objet symbolisé, la seule approche possible pour évaluer l'information liée à un ensemble de symboles est de nature statistique, ce qui conduit à la théorie classique de l'information [cf notamment : Shannon, 1948, Rényi, 1966 ou, plus récemment : Cover et Thomas, 1991, Battail, 1997], avec :

$$H = - \sum_{i=1}^N p(i) \log p(i) \quad \text{avec} \quad p(i) = \text{prob}(\chi_i)$$

Toutefois, la prise en compte au cours de cette symbolisation de l'existence et des propriétés de l'objet en tant qu'objet physique transforme la nature purement statistique de cette évaluation en une approche de nature quantique [cf notamment : Von Neumann, 1955, Holevo, 1973 ou, plus récemment : Bennet, 1998, Cerf et Adami, 1998, Gruska, 1999, Landauer, 1991, Preskill, 1998]. La théorie quantique renseigne sur les caractéristiques de certaines des probabilités (les probabilités quantiques) qui entrent dans le calcul de l'entropie, alors que la théorie classique est muette concernant l'origine des probabilités classiques. Toutefois ces indications ne proviennent pas de l'objet lui-même, symbolisé par la bibliothèque  $\Xi = \{|\chi_i\rangle\}$  formée d'une base orthonormée d'états quantiques  $|\chi_i\rangle$ , mais de l'application d'un observable  $\mathbf{M}$  sur la source  $\Sigma$ , source dont la matrice de densité est  $\rho = \sum_i p(i) \langle \chi_i | \chi_i \rangle$ . L'entropie de Von Neumann est alors définie par ("tr" désignant l'opérateur trace) :

$$S = - \text{tr} (\rho \log \rho)$$

Quant aux sources alphabétiques, la procédure qui consiste à encoder une suite de symboles en une suite de nombres ou de mots est de nature algorithmique [cf notamment : Kolmogorov, 1965, Solomonoff, 1964, Chaitin, 1969, Loveland, 1969b, ou, plus récemment : Delahaye, 1994, Li et Vitányi, 1993, Uspensky, 1992, Uspensky et Chen, 1996]. Selon le type de source, deux schémas sont applicables, en "étoile" dans le cas d'une source lexicale où tous les symboles sont transcrits par des nombres indépendants les uns des autres, ou en

"arbre" dans le cas où les symboles sont transcrits par des mots classés selon un ordre préfixe. Une fonction récursive partielle  $\Phi$  calculée par une machine de Turing universelle  $U$  étant donnée, la complexité de  $s$  est évaluée par la longueur du plus petit programme  $w$  qui calcule  $s$  :

$$C_{\Phi}(s) = \min \{l(w) : \Phi(w) = s\}$$

La transcription des symboles est plus ou moins efficace. Cette efficacité est mesurée par le degré de compression des messages ou *chaînes* transcrits. Deux messages portant la même information pouvant avoir des caractéristiques différentes (être notamment de longueurs différentes), il existe quatre mesures de l'entropie algorithmique, selon la nature des chaînes-sources qui encodent les symboles et la nature des chaînes-code qui encodent les chaînes-source (l'indice désignant la chaîne-source) :  $C_C$  (complexité simple),  $C_K$  (complexité uniforme),  $K_C$  (complexité préfixée),  $K_K$  (complexité monotone).

Le tableau ci-dessous résume quelques propriétés des différentes entropies évoquées :

Sources : Propriétés			Symbolique			Lexicale		Préfixée		
			H	Hd	S	Cc	Ck	Kc	kc	Kk
<i>Enumérabilité</i>			O	N	O *2	O	O	O	N	O
<i>Minimum</i>	L'entropie d'un message connu est nulle	$I(0) = I(1) = 0$	O	N	O	O	O	O	O	O
<i>Majoration</i>	Il existe une fonction de majoration de l'entropie	$I(s) \leq f(s)$	O	N	O	O	O	O	O	O
<i>Concavité</i>	L'entropie d'une moyenne est supérieure à la moyenne des entropies élémentaires.	$I(\sum k.s) \geq \sum k.I(s)$	O	O	O	O	O	O	O	O
<i>Monotonie</i>	L'entropie du tout est supérieure à l'entropie des parties.	$I(s, r) \geq I(s)$ $I(s, r) \geq I(r)$	O	N	N	N	O	O	O	O
<i>Additivité</i>	Etant donné s et r, la connaissance de s ne peut que diminuer l'incertitude sur r.	$I(s   r) \leq I(s)$ $I(s, r) = I(s) + I(r   s)$ $= I(r) + I(s   r)$	O	O	O	N	O	N	O	?
<i>Sous-additivité</i>	Les corrélations entre 2 parties composant un objet ne peuvent que diminuer l'entropie totale	$I(s, r) \leq I(s) + I(r)$	O	O	O	N	N	O	O	?
<i>Sous-additivité "forte"</i>	Si 2 sources sr et rt (union srt) se recouvrent partiellement en leur intersection r, alors :	$I(s, r, t) + I(r) \leq I(s, r) + I(r, t)$	O	O	O	?	?	?	?	?
<i>Signes</i>	$I(s)$	L'entropie est toujours $\geq 0$	O	N	O	O	O	O	O	O
	$I(s, r)$	""	O	N	O	O	O	O	O	O
	$I(s   r)$	L'entropie conditionnelle est toujours $\geq 0$	O	N	N	O	O	O	O	O
	$I(s : r)$	L'entropie mutuelle est toujours $\geq 0$	O	O	O	O	O	O	O	O
<i>Symétrie de l'entropie mutuelle</i>	Information contenue dans s à propos de r = information contenue dans r à propos de s	$I(s : r) = I(r : s)$	O	O	O	N	N	N	O	?
<i>Changement de référentiel</i>	L'entropie est invariante dans un changement de coordonnées	$s \rightarrow s'$ $\Rightarrow I(s) = I(s')$	O	N	O *3	N *4	N *5	N *5	N *5	N *5
<i>Aléatoire</i>	Il existe une loi de probabilité.		O	O *1	O	N	N	O	O	O *6

Notes : O : oui ; N : non ; ? : propriété non examinée dans cette étude

\*1 : sous réserve que les intégrales convergent

\*2 : la mesure (projection) est discrète, mais le contenu d'information (état d'un bit quantique) est continu

\*3 : par changement de base

\*4 : permutation récursive ; Oui pour la complexité conditionnelle sachant la longueur  $C_c(s | I(s))$

\*5 : permutations récursives

\*6 : extension au continu impossible

L'analyse des processus informationnels qui s'appliquent aux différentes sources conduit alors à l'argumentation suivante.

1°) Si l'on ne fait aucune hypothèse préalable sur les objets et leur symbolisation (hormis la discernabilité et la bipolarité), aucun symbole, comme aucun nombre, ne contient "plus d'information" ou n'est "plus significatif" qu'un autre : tous les éléments des ensembles étudiés sous cet angle sont "sur un pied d'égalité", ce qui implique une structuration en étoile, propre aux sources lexicales.

2°) Mais pour qu'il y ait transfert de plus d'un élément d'information entre deux ensembles, il faut nécessairement que les chaînes qui permettent ce transfert soient discernables les unes des autres, donc structurées en arbre, autrement dit préfixées. Par conséquent la mesure adaptée aux ensembles informationnels alphabétiques est la complexité  $K_C$ . Rappelons que la quantité  $K_C(s)$  est la longueur minimale de la chaîne-code préfixée  $w^*$  (i.e. le programme le plus court) qui produit la chaîne-source  $s$ . N'étant pas récursive,  $K_C(s)$  n'est pas calculable à partir de  $s$ . Par contre, la chaîne  $\langle s, K_C(s) \rangle$  contient la même information que  $w^*$  : par énumération diagonale (1er pas du programme 1, 1er pas du prog. 2, 2ème pas du prog.1, 1er pas du prog. 3, etc), on peut retrouver  $w^*$  à partir de  $\langle s, K_C(s) \rangle$  (resp.  $v^*$ ,  $\langle r, K_C(r) \rangle$ ). La réciproque est triviale.

3°) Comme l'information est bipolaire, c'est-à-dire ne saurait avoir un contenu significatif que si l'on prend en compte l'association des objets entre eux (au moins deux,  $s$  et  $r$ ), il faut traiter non seulement des quantités  $K_C(s)$  et  $K_C(r)$ , mais aussi des quantités  $K_C(s,r)$ ,  $K_C(s|r)$ ,  $K_C(r|s)$ ,  $K_C(s:r)$  et  $K_C(r:s)$ . La chaîne  $\langle s, K_C(s|r) \rangle$  (resp.  $\langle r, K_C(r|s) \rangle$ ) contient la même information que la chaîne  $\langle u^*, r \rangle$  (resp.  $\langle t^*, s \rangle$ ), où  $u^*$  est le plus petit programme qui calcule  $s$  sachant  $r$  (resp.  $r$  sachant  $s$ ).

4°) Or, il n'y a pas de raison de préférer dans un couple informationnel un membre du couple à l'autre membre, ce qui se traduit par le postulat de symétrie  $I(s:r) = I(r:s)$ .

5°) Une telle égalité n'est pas possible avec la mesure  $K_C$ . Pour cela, il faut expliciter les données qui permettent la construction effective de la connaissance de  $r$  utilisée dans l'évaluation de  $s$  (et réciproquement), en substituant aux quantités  $K_C(s|r)$  et  $K_C(r|s)$  les quantités définies par  $k_C(s) = K_C(s | \langle r, K_C(r) \rangle)$  et  $k_C(r) = K_C(r | \langle s, K_C(s) \rangle)$ , pour obtenir l'égalité  $k_C(s:r) = k_C(r:s)$ . Alors la chaîne :

$$\langle s, K_C(s | \langle r, K_C(r) \rangle) \rangle = \langle s, K_C(s | v^*) \rangle + O(1) = \langle s, k_C(s|r) \rangle + O(1)$$

$$(\text{resp. } \langle r, K_C(r | \langle s, K_C(s) \rangle) \rangle = \langle r, K_C(r | w^*) \rangle + O(1) = \langle r, k_C(r | s) \rangle + O(1) )$$

est calculable à partir de la chaîne  $\langle u^{**}, v^* \rangle$  (resp.  $\langle t^{**}, w^* \rangle$ ), où  $u^{**}$  est le plus petit programme qui calcule  $s$  sachant  $v^*$  (resp.  $r$  sachant  $w^*$ ). La connaissance contextuelle invoquée dans la description de  $s$  (resp. de  $s$  pour décrire  $r$ ) est le programme de calcul de  $r$  et non  $r$  elle-même. Seule cette connaissance permet la construction effective de la description de  $s$  sachant  $r$  (resp. de  $r | s$ ). C'est pourquoi nous appelons *complexité préfixée explicite* la quantité  $k_C$ , également nommée *complexité de Chaitin*.

6°) Mais la complexité algorithmique préfixée explicite conditionnelle  $k_C(s | r) = K_C(s | v^*)$  n'est pas co-énumérable. La quantité  $k_C(s : r) = k_C(s) - k_C(s | r)$  n'est donc pas énumérable. L'explicitation de l'implicite entraîne la non-énumérabilité de la mesure du contenu d'information mutuel. On ne peut plus comparer les chaînes  $\langle s, K_C(s | v^*) \rangle$  et  $\langle u^{**}, v^* \rangle$ . Non seulement il n'est pas possible de calculer  $u^{**}$  et donc  $K_C(s | v^*) = l(u^{**})$  à partir de  $s$  sachant  $v^*$  pour cause de non récursivité, mais encore il n'est pas possible de retrouver  $u^{**}$  par énumération à partir de  $\langle s, K_C(s | v^*) \rangle$  pour cause de non énumérabilité.

L'irruption du continu dans le champ informationnel entraîne à son tour nombre de conséquences et de difficultés. La définition initiale des objets, qui conditionne tout l'édifice, s'en trouve bouleversée. Par définition même du continu, la notion de discernabilité disparaît.

1°) Les contenus d'identification  $| w^*$  et  $| v^*$  de  $r$  et de  $s$  respectivement sont des chaînes préfixées alors que les contenus  $| s$  et  $| r$  ne le sont pas. Si l'on veut encoder ces contenus, il faut donc substituer au schéma de transcodage sources lexicales  $\rightarrow$  encodages préfixés, le schéma sources préfixées  $\rightarrow$  encodages préfixés, dont la mesure est la *complexité algorithmique monotone*  $K_K$ . Or l'extension au continu de cette complexité conduit à une impasse. Il faut en effet distinguer  $K_{K_1}$ , ou *complexité a priori*, fondée sur une semi-mesure de probabilité définie sur un ensemble continu :

$$K_{K_1}(s) = -\log(\Pi(s))$$

et  $K_{K_0}$ , ou complexité monotone, dont découle la définition d'une probabilité algorithmique  $\pi(s)$  par l'intermédiaire de l'extension au continu du théorème de codage :

$$K_{K_0}(s) = -\log(\pi(s))$$

Ces deux versions de la complexité monotone diffèrent d'un ordre de grandeur égal à

l'inverse de la fonction d'Ackermann [Gács, 1983], fonction qui peut être interprétée en termes de cardinaux transfinis [Odifreddi, 1989], ce qui permet de préciser l'origine exacte du non dénombrable dans la théorie de l'information. En suivant la croissance de la fonction d'Ackermann, l'encodage devient arbitrairement difficile.

2°) Le fait que l'extension au continu de la théorie statistique classique à travers l'entropie différentielle  $H_d$  conduite à des entropies négatives est bien connu. Seule demeurent positifs dans tous les cas le gain d'information et donc l'information mutuelle [Kolmogorov, 1956]. L'entropie de Von Neumann manifeste un comportement allant dans le même sens dans le cas de systèmes quantiques corrélés : l'information mutuelle demeure positive, mais l'information conditionnelle peut être négative dans une corrélation. Cela conduit pour les quantités d'information à des violations d'inégalités qui sont par ailleurs bien connues du domaine expérimental [Bell, 1964, Clauser *et al.*, 1969, Aspect *et al.*, 1982]. L'équivalent "entropique" de l'inégalité BCHSH est :

$$H(\hat{A}_1:\hat{B}_1) + H(\hat{A}_2:\hat{B}_1) + H(\hat{A}_1:\hat{B}_2) - H(\hat{A}_2:\hat{B}_2) \leq 2$$

Le calcul numérique de la partie gauche de cette inéquation sur les entropies mutuelles présente un maximum de violation de l'ordre de 2,22 pour  $\theta \approx 17,4^\circ$  ( $\theta$  étant l'angle  $(\hat{A}_i, \hat{B}_j)$  séparant les détecteurs de polarisation des photons corrélés). Ainsi la "téléportation" d'un bit quantique, application de la corrélation, ne permet pas l'échange (i.e. ni le transport ni la mesure) de symboles discrets mais seulement le partage d'une bibliothèque continue représentée par un vecteur d'état qui n'offre que la potentialité d'un échange de données (celles-ci ne pouvant être mesurées qu'en détruisant l'état quantique du système).

On peut conclure de ce qui précède qu'un processus informationnel est discret sur le plan de l'effectivité, au niveau opérationnel où les données sont explicitées, mais qu'il s'accompagne d'une part contextuelle implicite de nature continue que l'on ne peut qu'approcher. Aux deux classes de calculabilité bien connues fondées sur les automates finis (fini dénombrable) et les machines de Turing (infini dénombrable), l'analyse de tels processus conduit à poser l'existence d'une troisième classe de calculabilité [Anceau, 1997], fondée sur les nombres réels [Blum *et al.*, 1989], encore appelée *super-Turing* par certains auteurs [Siegelmann, 1999], dont la propriété essentielle est de calculer des fonctions qui ne sont pas calculables par les machines de Turing.

Pour examiner les propriétés et les limitations de l'extension au continu des processus de calcul, nous proposons ici de modéliser celle-ci par le biais de l'analyse harmonique, qui apparaît comme une suite naturelle dans le schéma :

CLASSES :	Fini dénombrable	Infini dénombrable	Infini non dénombrable
	<i>Automates</i>	<i>Machines de Turing</i>	<i>Processus informationnels</i>
	Circuits Booléens	Arithmétique	Analyse harmonique
OPÉRATEURS :	<i>bits</i>	<i>chaînes</i>	<i>formes</i>
Disjonction (Op)	NON	Logarithme, exponentielle	Transformation de Fourier
Conjonction (*, •)	ET, OU	Multiplication, Addition	Multiplication, Convolution
$Op(A \bullet B) = Op(A) \bullet Op(B)$	$\neg (A \wedge B) = \neg A \vee \neg B$	$\text{Log}(A \times B) = \text{Log}A + \text{Log}B$	$F(A \times B) = F(A) * F(B)$
$Op^{\sim}(A * B) = Op^{\sim}(A) \bullet Op^{\sim}(B)$	$\neg (A \vee B) = \neg A \wedge \neg B$	$\text{Exp}(A + B) = \text{Exp}A \times \text{Exp}B$	$F(A * B) = F(A) \times F(B)$
Eléments neutres	$\emptyset, 1$	0,1	$\delta(v), 1(v)$

Nous nommons *super-booléen* [Pinson, 1996] un tel principe de calcul. On sait que la réduction de l'analyse harmonique discrète au formalisme booléen est possible.

Réciproquement, l'extension de celui-ci à l'analyse harmonique discrète puis continue fait apparaître un certain nombre de propriétés logiques nouvelles inconnues du monde booléen (point fixe, modalité, principe d'indétermination, complexité en  $n \log n$ , etc).

La question suivante, ouverte, se pose : le continu ne pouvant être traité qu'abstraitement, et les processus informationnels étant supposés effectifs, comment et à quelles conditions ceux-ci approchent-ils le continu ? Notre modèle apporte un premier indice : en analyse harmonique discrète, la durée d'observation, le nombre d'échantillons, le pouvoir séparateur, nécessairement finis, se soldent par un fenêtrage du signal (et un échantillonnage dans l'espace dual), ce qui a une double conséquence : 1°) la longueur  $n$  de la chaîne traitée est une donnée implicite fondamentale supposée connue préalablement à l'analyse ; 2°) l'information sur l'origine (la phase) du signal est perdue, ou plus exactement devient arbitraire. Or, certaines indications théoriques semblent aller dans ce sens.

Respectivement :

1°) Il est possible de fonder une classe de calculabilité *nonuniforme* avec les nombres réels à partir d'un modèle de calcul dit "ATM" (*Advice Turing Machine*) où le "conseil" externe apporté au calcul concerne la longueur  $n = l(s)$  de la chaîne traitée  $s$ , sous la forme (avec délimitation)  $\langle s, n \rangle$ . Siegelmann montre qu'un réseau neuronal analogique fondé sur ce modèle possède une propriété intéressante : bien qu'un nombre réel soit une donnée de précision infinie, la précision demandée par le calcul neuronal est linéaire. Pour  $q$  pas de calculs, seulement les  $O(q)$  premiers bits des paramètres réels sont nécessaires ; les bits de poids faible sont sans effet sur la sortie.

2°) Si l'on se donne une fonction récursive  $f$  et  $n$  segments initiaux de son graphe  $G$ , avec  $G_f^n = \{ \langle x, f(x) \rangle, x \leq n \}$ , et si l'on compare l'information contenue dans  $f$  et celle contenue dans

$G_f^n$ , on constate que la complexité  $C_C(G_f^n | n)$  des segments initiaux du graphe approche celle de la fonction  $C_C(f)$ , mais pas toujours [Durand, Porrot, 1998]. Or  $C_C(G_f^n | n)$  est une complexité conditionnelle sachant la longueur : une telle forme de complexité est indépendante d'une permutation récursive [Loveland, 1969a] – cela constitue d'ailleurs une exception dans les différentes formes de complexités algorithmiques. Donc en passant de la fonction au graphe on perd une part de l'information, celle qui concerne l'ordre de sa description.

Précision linéaire mais perte d'information sur l'ordre : telle pourrait être le bilan de l'approximation discrète du continu. En tout état de cause, c'est en croisant de façon *multicruciale* les différentes approches possibles de la mesure du contenu d'information des objets que l'on peut espérer cerner efficacement ce concept dans toute sa complexité, par le biais de modèles de calcul qui en montrent les particularités, les possibilités et les limitations.

### Bibliographie résumée

(bibliographie complète en fin de volume)

1. ANCEAU F. : Vers l'hyper-calculabilité, *coll. XVIème Congrès International de Bio-Mathématique*, Paris, 11-13 septembre 1997.
2. ASPECT, A., DALIBARD, J., ROGER, G. : Experimental tests of Bell's inequalities using time-varying analysers, *Physical Review Letters*, vol.49, p 1804-1807, 1982.
3. BATTAIL, G. : *Théorie de l'information*, Masson, Paris, 1997.
4. BELL, J.S. : On the Einstein-Podolsky-Rosen paradox, *Physics I*, p 195-200, 1964.
5. BENNET, C.H. : Quantum information, *Physica Scripta*, vol.76, p 210-217, 1998.
6. BLUM, L., SHUB, M., SMALE, S. : On a theory of computation and complexity over the real numbers : NP completeness, recursive functions, and universal machines, *Bulletin of the American Mathematical Society*, n°21, p 1-46, 1989.
7. CERF, N. J., ADAMI, C. : Quantum information theory of entanglement and measurement, *Physica D*, vol.120, p 62-81, 1998. eprint quant-ph/9605039 .
8. CHAITIN, G. : On the Length of Programs for Computing Finite Binary Sequences : Statistical Considerations, *Journal of the ACM*, vol.16, n°1, p 145-159, 1969.
9. CLAUSER, J.F., HOLT, R.A., HORNE, M.A., SHIMONY, A. : A proposed experiment to test local hidden-variable theories, *Physical Review Letters*, vol.23, p 880-884, 1969.
10. COVER, T., THOMAS, J.A. : *Elements of Information Theory*, John Wiley, New York, 1991.

11. DELAHAYE, J.-P. : *Information, complexité et hasard*, Hermès, Paris, 1994.
12. DEMAN, P. : La compression d'information à la division télécommunications, *Revue Technique Thomson-CSF*, vol 7, n°4, p 723-730, dec. 1975.
13. DURAND, B., PORROT, S. : *Comparison between the complexity of a function and the complexity of its graph*, Research Report n°98-27, Ecole Normale Supérieure de Lyon, 1998. eprint <ftp://ftp.lip.ens-lyon.fr/Rapports/RR/RR1998/RR1998-39.ps.Z>
14. GÀCS, P. : On the relation between descriptive complexity and algorithmic probability, *Theoretical Computer Science*, n° 22, p 71-93, 1983.
15. GRUSKA, J. : *Quantum Computing*, Mc Graw Hill, London, 1999.
16. HOLEVO, A. S. : Bounds for the quality of information transmitted by quantum communication channel, *Problems of Information Transmission*, vol.9, p 177-183, 1973.
17. KOLMOGOROV, A.N. : On the Shannon Theory of Information Transmission in the Case of Continuous Signals, *IEEE Transactions on Information Theory*, vol IT-2, p 102-108, 12/1956.
18. KOLMOGOROV, A.N. : Three Approaches to the Quantitative Definition of Information, *Problems Information Transmission*, vol.1, n°1, p 1-7, 1965.
19. LANDAUER, R. : Information is physical, *Physics Today*, n°44, p 23-29, mai 1991.
20. LI M., VITÀNYI P. : *An Introduction to Kolmogorov Complexity and Its Applications*, Springer-Verlag, Berlin, 1993, 1997.
21. LOVELAND, D.W. : On minimal-program complexity measures, *Conference Record of the ACM Symposium on theory of computing*, p 61-65, mai 1969a.
22. LOVELAND, D.W. : A variant of the Kolmogorov concept of complexity, *Information and Control*, vol.15, p 510-526, 1969b.
23. OSWALD, J. : *Théorie de l'information ou Analyse diacritique des systèmes*, Masson, Paris, 1986.
24. ODIFREDDI, P. : *Classical Recursion Theory - The Theory of Functions and Sets of Natural Numbers*, North-Holland, Amsterdam, 1989.
25. PINSON, G. : Cognition et Multicrucialité, *Revue Internationale de Systémique*, vol 8, n°2, p 183-210, 1994. *Suivi de* : Une théorie cognitive de l'information, *Revue Internationale de Systémique*, vol 9, n°1, p 27-66, et vol 9, n°5, p 541-544, 1995.
26. PINSON, G. : Beyond Boole Algebra with TMS320 DSP Multiprocessing, *Proceedings of The First European DSP Education and Research Conference*, Texas Instrument, ESIEE, Paris, 25-26 sept. 1996.
27. PRESKILL, John : *Advanced Mathematical Methods of Physics*, Physics 229, UCLA, Los Angeles, 1997-98 et 1998-99.  
<http://www.theory.caltech.edu/~preskill/ph229>.
28. RENYI, A. : *Calcul des probabilités*, Dunod, Paris, 1966 (Original : Deutscher Verlag der Wissenschaften, 1962 ; Réédition : Jacques Gabay, 1992).
29. SHANNON, C.E. : A Mathematical Theory of Communication, *AT&T Bell Laboratories Technical Journal*, vol.27, p 379-423 et p 623-656, 1948.

30. SIEGELMANN, H., T. : *Neural Networks and Analog Computation : Beyond the Turing Limit*, Birkhäuser, Boston, 1999.
31. SOLOMONOFF, R.J. : A formal theory of inductive inference, Part. I., *Information and Control*, n°7, p 1-22, 1964.
32. USPENSKY V.A. : Complexity and Entropy : An Introduction to the Theory of Kolmogorov Complexity, in WATANABE O. (dir.), *Kolmogorov Complexity and Computational Complexity*, EATCS, Springer-Verlag, Berlin, 1992.
33. USPENSKY V.A., SHEN, A. : Relations Between Varieties of Kolmogorov Complexities, *Mathematical Systems Theory*, vol.29, n°3, p271-291, 1996.
34. VON NEUMANN, J. : *Mathematical Foundations of Quantum mechanics*, Princeton University Press, Princeton, NJ, 1955.