Revue Internationale de

systēmique

ISBN 2-10-000151-5

Vol. 5, N° 5, 1991

afcet

DUNOD

AFSCET

Revue Internationale de systémique

Revue Internationale de Sytémique

volume 05, numéro 5, page 561 - 581, 1991

Les théories de l'information

Jean-Paul Delayaye

Numérisation Afscet, août 2017.

[65] J.-P. LUMINET, J. CHASTANG et Th. MOULIN, Structures de séquences engendrées par des relateurs arithmétiques, *Cahiers Systema*, 10, 1983, p. 69-149.

[66] Cl. VALLET, H. LE GUYADER et Th. MOULIN, Ambiguity and imprecision in arithmetical models of natural systems, in *Applied Systems and Cybernetics*, G. E. LASKER éd., VI, Pergamon Press, p. 3070-3075.

[67] L. ZADEH, The birth and evolution of fuzzy logic, *Internat. J. Gen. Systems*, 17, 2-3, 1990, p. 95-105.

[68] G. J. KLIR, A principle of uncertainty and information invariance, *Internat. J. Gen. Systems*, 17, 2-3, 1990.

[69] T. YAMAKAWA, Hardware implementation of fuzzy logic systems, *3rd Int. Conf. on Information Processing and Management of Uncertainty in knowledge-based systems* (*IPMU*), Paris, 2-6 juillet 1990 (à paraître).

[70] G. GENTZEN, *Recherches sur la déduction logique*, P.U.F., Paris, 1955 (version originale en allemand, 1934).

[71] P. LORENZEN, *Métamathématique*, Gauthier-Villars, Paris, 1967 (version originale en allemand, 1962).

[72] B. PAULRÉ, *La Causalité en Economie*, Presses Universitaires de Lyon, Lyon, 1985.

[73] C. P. BRUTER, *Sur la nature des mathématiques*, Gauthier-Villars, Paris, 1973.

[74] R. THOM, *Stabilité structurelle et morphogénèse*, W. A. Benjamin, USA, 1973.

[75] R. THOM, *Modèles mathématiques de la morphogénèse*, Collection 10×18, n° 887, U.G.E., Paris, 1974.

[76] Cl. VALLET, Modélisation arithmétiques de formes biologiques, *Cahiers Systema*, 9, nov. 1982, p. 5-40.

[77] J. MAURIN, *Simulation déterministe du hasard*, Masson, Paris, 1975.

[78] J. BASS, *Fonctions de corrélation, fonctions pseudo-aléatoires et applications*, Masson, Paris, 1984.

[79] G. CHAUVET, *Traité de Physiologie Théorique*, 3, *Physiologie intégrative : champ et organisation fonctionnelle*, Masson, Paris, 1990.

[80] V. ARNOLD, A. VARCHENKO et S. GOUSSEIN-ZADE, *Singularités des applications différentiables*, Editions MIR, Moscou, 1986 (version originale en russe, 1982).

[81] P. SLODOWY, Simple singularities and simple algebraic groups, *Lectures Notes in Math.*, 815, Springer Verlag, Berlin.

[82] T. POSTON et I. STEWART, *Catastrophe Theory and its Applications*, Pitman, Londres, 1978.

[83] Cl. VALLET, Th. MOULIN, H. LE GUYADER et L. LAFRENIÈRE, Emergence et imbrication de niveaux d'organisation dans les systèmes complexes, *VIII^e Congrès Int. de Cybernétique*, Namur, 1976, p. 187-199.

[84] E. STUDY et E. CARTAN, *Nombres complexes*, in *Œuvres complètes*, partie II, I, Gauthier-Villars, 1953, p. 107-246 (texte original, paru dans l'Encyclopédie des Sciences, en 1908, p. 329-468).

REVUE INTERNATIONALE DE SYSTÉMIQUE
Vol. 5, N° 5, 1991, pp. 561 à 581

LES THÉORIES DE L'INFORMATION

J.-P. DELAHAYE

Université des Sciences et Techniques de Lille,
Laboratoire d'Informatique Fondamentale de Lille[1]

Résumé

La notion d'information dispose-t-elle d'une ou de théories satisfaisantes? Mathématiquement elle en dispose de plusieurs dont nous tentons de montrer qu'aucune n'est en mesure de constituer une théorie complète et définitive de l'information. Nous insistons sur l'intérêt de théories récentes comme celle de Chaitin-Kolmogorof et celle de Bennett. En physique le concept d'information est difficile à identifier quel que soit l'importance de ses liens avec la thermodynamique, liens qui semblent avoir subi une refonte de première importance très récemment grâce aux travaux de Bennett et Zurek. Quant à la biologie elle cumule toutes les difficultés, car de toute évidence un sens pragmatique doit être attribué au concept d'information biologique qui rend inutilisables ou incomplètes les théories mathématiques de l'information et peu probable l'utilité des théories thermodynamiques de l'information.

Abstract

Is there a satisfactory theory of information? There are several mathematical theories of information, and we try to show that none is the real complete and ultimate theory we are looking for. We insist on the interest of several new propositions in the field : the algorithmic theory of information of Chaitin-Kolmogorof, and the theory of logical depth of Bennett. In physics the concept of information is obviously related to thermodynamics, but a new point of view (W.H. Zurek) about this relationship has been very recently proposed that forces us to reconsider many of the old (and sometime not sufficently cautious) uses of simple identifications concerning information and physical entropy. In biology, the theory of information is in a more difficult position because elementary reflexions prove that a pragmatic sense to information cannot be avoided, and must be added to the mathematical sense and the physical sense of information.

1. U.A. CNRS 369, Bât. M3, 59655 Villeneuve d'Ascq Cedex, France.

Rev. intern. systémique. 0980-1472 Vol. 5/91/05/ 561 /21/$ 3.10/© Afcet Gauthier-Villars

Introduction

Le mot information est employé dans des phrases et des contextes très variés. Par exemple on dit : « Les informations contenues dans ce livre », « L'information dont il dispose sur le problème », « Les informations codées dans le génome »; « Le peu d'informations qu'a apporté son long discours ».

Lorsqu'on parle d'information on pense souvent « information ayant une certaine valeur », ou « information pouvant servir à... », ou « contenu en information ». On sent bien que derrière ce mot se cache quelque chose de complexe, quelque chose de variable peut-être, quelque chose en tout cas qui mérite qu'on y réfléchisse. On est donc amené à se poser la question :

Peut-on faire une théorie scientifique générale de l'information? et si oui comment s'y prendre?

Il n'est pas facile de répondre sérieusement, et il est très facile de répondre mal, car certaines théories mathématiques (ou physiques) utilisent déjà le mot information, énoncent des théorèmes et donnent l'illusion que le problème est résolu et que donc on peut parler mathématiquement de l'information.

La théorie de l'information de Shannon ([Shannon, Weaver, 1949], voir aussi [Brillouin, 1959] [Ash, 1965] [Cullmann *et al.*, 1970] [Martin, England, 1981] [Welsh, 1988]) pendant longtemps a prétendu détenir le sens de ce qu'est l'information. L'intimidation a d'ailleurs si bien réussi que certains croient que l'information c'est ce qu'en dit cette théorie et rien d'autre. Lorsqu'on a voulu parler précisément de l'information génétique on a essayé de faire des liens avec cette théorie mathématique; les résultats n'ont pas été à la mesure des espoirs et des ambitions annoncées ([Atlan, 1972, 1979] et on en a même contesté la pertinence ([Tonnelat, 1978]). On s'est parfois consolé en disant que ce n'était qu'une théorie de la transmission de l'information. Aujourd'hui il semble clair que la théorie de Shannon a assez peu enrichi la compréhension que les biologistes ont par exemple de la façon dont le génome se trouve inscrit dans les chromosomes de chaque cellule d'un être vivant.

De même, alors qu'à un moment on a pu penser que la théorie de l'information allait être « la théorie de l'informatique », on a vite découvert qu'il n'en était rien, et les programmes des licences et maîtrises d'informatique d'aujourd'hui laissent bien peu de place à cette théorie.

Heureusement depuis une vingtaine d'années une autre théorie de l'information, dite « théorie algorithmique de l'information » ou « théorie de l'information de Chaitin-Kolmogorof », est apparue ([Kolmogorof, 1965] [Chaitin, 1966, 1987*a*, 1987*b*] [Li, Vitanyi, 1990, 1991]). Parfois on la présente comme un substitut à la théorie de Shannon, mais puisqu'une théorie mathématique ne saurait être fausse du point de vue des mathématiques (et nul ne prétend qu'il y a de fausses théories de l'information) on se trouve en présence de deux théories mathématiques de l'information, et donc l'une est de trop. A moins que ce ne soit les deux.

C'est ce que nous allons soutenir ici. Précisément, nous allons tenter de démontrer que ni la théorie de Shannon, ni la théorie de Chaitin-Kolmogorof, ni aucune autre, ne disent tout de ce qu'est l'information, et que de plus, chacune de ces théories en prétendant être « la théorie de l'information » empêche de progresser. L'information doit être conçue sous des formes générales et variées dont il est peu probable qu'on fasse le tour rapidement, et dont aucune théorie mathématique ne fournira le secret ultime : la notion d'information est « sémantique », c'est-à-dire riche de sens : vouloir la réduire à quelques équations ou définitions est déraisonnable.

En cours de chemin nous reconnaîtrons aux théories mentionnées un intérêt réel, celle de Chaitin-Kolmogorof d'ailleurs méritant une attention particulière qu'elle n'a peut-être pas encore reçue, à cause sans doute des déceptions produites par les prétentions excessives de la théorie de Shannon qui ont laissé un goût amer à ceux qui s'y sont épuisés pour peu de résultats.

Sans prétendre donner « la solution » à un problème qui nous semble n'en avoir aucune de simple, nous proposerons une vue possible de ce qu'on peut entendre par *information relativement à un but*, et par *information descriptionnelle* (qui en est un cas particulier).

Beaucoup des propos que nous allons tenir sont de simples remarques de bon sens, mais il nous a semblé important de les formuler, tant les approximations, excès de simplification, identifications injustifiées et abus naïfs sont nombreux dans la littérature scientifique et de vulgarisation dès qu'on parle d'information. Les cas de la physique et de la biologie seront examinés à part et là encore nous soutiendrons que la prudence est de rigueur, quoique pour la physique des travaux récents de W.H. Zurek ont fait tout d'un coup avancer considérablement notre compréhension des liens entre entropie physique et information.

1. Les théories de l'information

(a) Des exemples à prendre au sérieux

Précisons le problème.

Pour simplifier les choses nous allons supposer que nous avons une suite de symboles s. Et nous allons essayer de réfléchir à ce que nous entendons lorsque nous parlons de son *contenu en information* ou de sa *valeur en information*.

Pour rendre les choses un peu concrètes, nous prêterons une attention particulière aux exemples suivants de chaînes de caractères :

– la suite des caractères composant l'Introduction à la Psychanalyse de Freud,

– la liste dactylographiée des emplacements des lances-missiles américains dans le monde,

– une table de logarithme,

– le génome complet d'un virus HIV du Sida,

– un « compact disc » avec les concertos pour piano de Chopin interprétés par Samson François,

– le programme du traitement de texte que j'utilise pour taper ce texte tel qu'il est en ce moment dans la mémoire de mon ordinateur,

– le programme de ce même traitement de texte avant qu'il n'ait été compilé, qu'on appelle « programme source » (l'éditeur du logiciel ne le diffuse pas),

– le résultat de tel sondage politique commandité par un parti et qu'il se garde bien de rendre public.

Tous ces exemples correspondent bien à des objets ayant, ou ayant eu à un moment donné, un contenu en information d'une certaine valeur. La preuve en étant tout simplement qu'ils ont pu être vendus et achetés, ou qu'on a dépensé de l'argent pour les produire, ou qu'on continue d'en dépenser pour les conserver. Une théorie générale de l'information qui ne reconnaît pas à l'un de ces objets une valeur en information ne peut qu'être incomplète ou inadéquate. Une théorie de l'information qui ne donne pas un moyen de comparer entre elles les valeurs du contenu en information de ces objets ne peut pas prétendre être « la théorie de l'information ».

(b) Le contenu brut en information

Accordons qu'il y a un *contenu brut* d'information pour chacun de ces objets, qu'il est donné en digits (ou bits) et qu'on peut le définir comme le nombre de mémoires élémentaires que la chaîne **s** occupe dans la mémoire d'un ordinateur quand on ne lui fait subir aucun traitement particulier autre que la mise sous un format compatible avec le système de l'ordinateur.

Le contenu brut d'information d'une chaîne de caractères **s** de longueur **n** est donnée par **n** si **s** ne comporte que des 0 et des 1 et c'est : $n \log(m)/\log(2)$ si la chaîne comporte des caractères pris parmi **m**, au lieu de pris parmi 2.

Dans nos exemples l'objet ayant le contenu brut d'information le plus grand est sans doute le « compact disc », celui ayant le plus petit contenu brut d'information est la liste des emplacements des lance-missiles. Celui ayant le plus de valeur marchande est le programme non compilé de traitement de texte ou la liste des emplacements des lance-missiles (bien que la valeur de cette liste diminue tous les jours !). Celui ayant le moins de valeur aujourd'hui est la table de logarithme, à moins que ce ne soit le résultat du sondage s'il est ancien et devenu inutile.

Il est clair que le contenu brut en information ne détermine pas la valeur de l'information. C'est une évidence : la valeur d'une information est quelque chose de plus compliqué et de relatif.

Et c'est bien parce que la valeur en information d'une chaîne de caractères est relative à un certain but et à une certaine situation qu'il y a plusieurs théories de l'information et qu'aucune ne peut vraiment traiter tous les problèmes que pose la notion d'information.

Si un but B est donné nous noterons Val (**s**, B) la valeur de l'information contenue dans **s** relativement à ce but B. Pour l'instant il s'agit d'une notion imprécise, mais nous allons voir que dans un certain nombre de cas simples Val (**s**, B) possède une définition naturelle. Nous allons découvrir qu'en précisant le but B on obtient différentes théories, dont en particulier les deux théories mentionnées dans l'introduction : la théorie de l'information de Shannon, et la théorie algorithmique de l'information.

(c) La théorie algorithmique de l'information (de Chaitin-Kolmogorof)

Si on se fixe le but de compresser au maximum la chaîne de caractères **s** et on suppose qu'on dispose pour cela d'une machine M alors : la valeur de l'information de **s** est la longueur du plus petit programme (écrit en binaire) qui lorsqu'on le donne à M lui permet de reconstituer la chaîne **s**. La valeur en information c'est ce contenu incompressible (relatif à M) de **s**.

Il se trouve que la puissance de calcul des machines n'est pas sans limite. Elles peuvent aller plus ou moins vite, mais dès qu'on a affaire à des machines d'une certaine puissance ce qu'elles peuvent calculer est le maximum de ce que n'importe quelle machine puissante peut calculer. C'est là la découverte fondamentale de Turing ([Turing, 1936]), qu'il y a des mécanismes de calculs universels et qu'ils ne sont pas très compliqués (n'importe quel micro-ordinateur est un tel mécanisme universel). Pourvu donc qu'on se donne un mécanisme de calcul universel, la notion de plus petit programme pouvant engendrer **s** ne dépend pas de la machine universelle qu'on utilise. Ou plus précisément, ne dépend de cette machine que par une constante additive qu'on peut négliger en première approximation si on traite des suites suffisamment mieux, puisque la fonction qui donne la complexité de Chaitin-Kolmogorof d'une chaîne **s** en fonction de **s** n'est pas une fonction récursive, et que déterminer la complexité d'une chaîne de caractères **s** constitue dans tout système formel donné S, un problème indécidable sauf pour des chaînes **s** d'une complexité inférieure à une constante c(S) dépendant de S ([Chaitin, 1974, 1975] [Delahaye, 1988, 1989 *a*, 1989 *b*, 1990]).

En tant que théorie optimale du contenu descriptionnel de l'information, la théorie de Chaitin-Kolmogorof joue un rôle théorique particulier qui commence à être reconnu par les biologistes et les physiciens. Citons par exemple : [Chaitin, 1979] [Bennett, 1982, 1986] [Eriksson *et al.*, 1987] [Bennett, 1988] [Zurek, 1989 *a*, 1989 *b*, 1990 *a*, 1990 *b*] [Caves, 1990] [Küppers, 1990]). Nous y reviendrons.

(e) La théorie de l'information pour la transmission de Shannon

Si le but poursuivi est de transmettre une chaîne de caractère **s** à un récepteur disposant de certaines connaissances sur la fréquence des lettres (prises dans l'alphabet $\{a_1, a_2, \ldots, a_n\}$) de la chaîne de caractères **s** que je veux transmettre, on définira la valeur en information de **s** par la formule bien connue :

$$\mathrm{Val}(\mathbf{s}, [\text{transmettre } \mathbf{s} \text{ à un récepteur qui sait la probabilité de } a_i \text{ est } p(i)]) \\ = \mathrm{E}(\mathbf{s}, \{p(i)\}) = \text{longueur}(\mathbf{s})\ (-\textstyle\sum p(i) \log(p(i)))$$

Cette formule donne en fait le contenu *moyen* d'information de l'ensemble des chaînes de caractères **s** quand on tient compte des probabilités $p(i)$ des lettres utilisées. Le « théorème de la voie sans bruit » indique que l'on ne peut pas *en moyenne* compresser plus les chaînes de caractères **s**.

Implicitement dans cette conception de l'information on suppose que le récepteur est susceptible de faire certains calculs pour reconstituer **s** à partir de ce qu'on lui transmet véritablement. Implicitement donc, on suppose un certain pouvoir de calcul du récepteur. En définitive la machine M qui fait le décodage peut être prise comme référence et on découvre alors que la théorie de Shannon peut être vue comme une version probabiliste de la théorie algorithmique de l'information (on ne suppose pas de limitation aux machines M de décodage) et compatible avec elle dans le sens suivant :

le contenu algorithmique *moyen* d'information de Chaitin-Kolmogorof des chaînes de caractères **s** de longueur n (pondérées par les probabilités résultant des fréquences supposées $p(i)$ pour les lettres a_i) K (**s**, $\{p(i)\}$) vérifie la relation :

$$\mathrm{E}(\mathbf{s}, \{p(i)\}) = \mathrm{K}(\mathbf{s}, \{p(i)\}) + \mathrm{O}(1)$$

(ce résultat dû à Chaitin se démontre en utilisant le fait que rares sont les suites ayant un contenu algorithmique en information nettement plus petit que leur contenu brut).

La théorie de l'information de Shannon est donc aussi une théorie de l'information par compression, qui au lieu de considérer des suites quelconques suppose que les suites qu'on transmet vérifient certaines propriétés statistiques exploitées pour compresser les données. Et finalement la théorie de Shannon est une théorie du contenu en information relativement à un but de compression et à une certaine distribution statistique des suites. Ce n'est donc pas une théorie de l'information limitée à cause du fait qu'elle ne s'occupe que du problème de la transmission, comme on le dit parfois, c'est une théorie de l'information probabiliste compatible avec la théorie algorithmique de l'information et limitée simplement parce qu'elle est relative à des distributions probabilistes particulières des ensembles de suites traitées.

Plus encore que dans le cas de la théorie de l'information de Chaitin-Kolmogorof la notion de contenu en information proposée par cette théorie est dépendante de certaines données. D'une théorie individuelle du contenu incompressible d'information on est passé à une théorie du contenu incompressible d'information *moyen*. Cette fois non seulement la machine de reconstitution de la suite **s** doit être connue (le récepteur doit connaître l'algorithme de codage) mais de plus le contenu en information de **s** n'est donné par la formule indiquée plus haut qu'en moyenne sous l'hypothèse d'une certaine distribution probabiliste des lettres utilisées (et donc des chaînes de caractères).

Bien des erreurs dans l'application de la théorie de l'information de Shannon proviennent de l'oubli de la nature probabiliste de cette théorie : cette théorie ne mesure jamais la quantité d'information contenue dans une chaîne de caractères individuelle, elle ne mesure que la quantité moyenne d'information d'une *famille* de chaînes (satisfaisant une certaine distribution). Parler de l'information de Shannon du génome de tel être vivant précis, par exemple, n'a pas de sens, car ce génome est une chaîne unique : l'approcher par le contenu incompressible moyen d'un ensemble des chaînes auxquelles elle appartient peut conduire à une erreur grossière.

Par exemple si une chaîne de dix millions de 0 est vue comme un élément particulier de l'ensemble de toutes les chaînes de dix millions de 0 et de 1 (0 et 1 munis de poids égaux) on lui attribuera un contenu en information de dix millions de bits, ce qui est absurde car une telle suite peut être décrite avec quelques bits (c'est d'ailleurs ce que j'ai fait). Une telle suite n'est pas un élément « typique » de l'ensemble de toutes les suites de dix millions de 0 et de 1. C'est commettre une erreur que de lui attribuer comme contenu en information celui d'une suite « typique ».

La théorie de l'information de Shannon est compatible avec celle de Chaitin-Kolmogorof. Chacune a son rôle à jouer et par exemple dans une théorie de la mesure ou de la prise d'information. C'est le jeu complémentaire des deux théories qui doit être considéré : une situation sur laquelle on dispose de peu d'éléments ne peut être qu'assimilée à la situation typique de l'ensemble des situations compatibles avec ce que l'on sait (utilisation de la théorie de Shannon), et au fur et à mesure des précisions qu'on acquiert c'est la théorie de Chaitin-Kolmogorof qui doit devenir prépondérante. Cette idée est très précisément celle qui a servi de base à Zurek pour reformuler la théorie de la mesure et de l'entropie physique. Nous allons en reparler plus loin. Il ne fait aucun doute qu'elle doit donner des résultats intéressants en dehors de la physique.

Toutes les théories que nous venons de présenter sont des théories « compressionnelles » (ou « descriptionnelles ») de l'information. Compresser des données est certes important mais il est bien clair que ce n'est pas cela uniquement qui constitue la valeur d'une information. On sait par exemple qu'une suite tirée au hasard de 0 et de 1, est le plus souvent incompressible, mais cela ne signifie pas bien sûr que chaque suite de 0 et de 1 a de la valeur, ni que chaque suite de 0 ou de 1 est incompressible.

Une autre insuffisance évidente des théories descriptionnelles de l'information est mise en évidence par la suite des cent premiers millions de digits de Pi qui a une complexité de Chaitin-Kolmogorof assez petite (car il y a des algorithmes courts qui la génèrent). Ce qui fait la valeur de cette suite (et on a dépensé beaucoup d'argent pour la calculer !) comme ce qui faisait la valeur des tables de logarithmes d'autrefois, c'est qu'une telle suite *représente une quantité importante de calculs*. La notion de *profondeur logique* proposée par Bennett prend en compte cet aspect du contenu en information, et on peut dire de cette autre théorie de l'information que c'est une théorie de l'information computationnelle. Son intérêt est peut-être égal à celui de la théorie de Chaitin-Kolmogorof.

(f) La notion de profondeur logique de Bennett, une théorie computationnelle de l'informaion

Si on poursuit comme but de limiter au maximum le nombre de pas de calculs qu'on doit faire pour retrouver **s**, alors la « valeur en information de la suite **s** » sera le nombre minimum de pas de calculs qu'il faut pour produire **s** ».

Une théorie adaptée à cette notion de valeur de l'information est celle de Bennett ([Bennett, 1987, 1988 *b*, 1990] [Rucker, 1987] [Delahaye, 1991 *b*]) où on définit *la profondeur logique* d'une suite **s** par le temps de calcul qu'il faut à une machine universelle M pour produire **s** à partir de sa description minimale (de Chaitin-Kolmogorof). Dans son article de 1988 Bennett montre que cette notion est robuste, c'est-à-dire qu'elle dépend relativement peu de la machine universelle qu'on se donne et qu'elle varie assez peu si au lieu du programme minimal on part d'un programme non minimal mais proche du programme minimal. Là encore on a une définition très séduisante de la valeur en information d'une chaîne de caractères.

Nous pensons effectivement que comme la théorie de l'information de Chaitin-Kolmogorof, cette théorie est appelée à jouer un rôle important en physique ([Bennett, 1986, 1988 *b*] [Lloyd, Pagels, 1988]), en biologie et dans d'autres disciplines, mais que pas plus que les autres, elle ne définit la notion de « valeur absolue en information d'une chaîne de caractères ».

Bien sûr là encore de très nombreuses variantes sont possibles : limiter le pouvoir de calcul des mécanismes utilisés; ne pas partir de la description minimale; faire des hypothèses statistiques sur les fréquences d'apparition des lettres etc.

Bennett [Bennett, 1988 *b*] introduit aussi d'autres concepts liés à la profondeur logique bien que différents : la notion de chaînes de caractères « *ambitieuses* », celle de chaînes de caractères « *cryptiques* ». Ces définitions confirment encore à nos yeux que la valeur de l'information contenue dans une chaîne de caractères **s** (même lorsqu'on s'intéresse plus au temps de calcul qu'à la compressibilité) n'a pas de définition universelle.

(g) Un programme compilé est de l'information de valeur

Dans nos exemples du début nous avons considéré un programme de traitement de texte. Il est clair que la compilation d'un programme source en un programme exécutable crée de l'information de valeur.

Là le but est un but mixte : avoir une forme compacte d'un algorithme assez rapide réalisant une famille de calculs.

– Plus le code est rapide plus la compilation a de la valeur.

– Plus le code compilé est compact plus il a de la valeur (aujourd'hui, contrairement à ce qui se passait il y a 30 ans on préfère favoriser la rapidité aux dépens de l'espace, car le prix des mémoires a beaucoup baissé).

– Plus nombreux sont les problèmes traités par le programme compilé plus sa valeur est grande.

Ce qui fait la valeur de l'information contenue dans un programme compilé c'est un mélange de ces trois qualités (rapidité, compacité, taille du champ des problèmes traités), et d'autres encore comme l'originalité, la portabilité, etc. La valeur de l'information d'un programme compilé (et de bien des chaînes de caractères données en exemple au début) ne se laissera jamais réduire à un seul aspect : le contenu en information d'un programme compilé ce n'est ni l'information algorithmique de Chaitin-Kolmogorof, ni l'information donnée par la théorie de Shannon, ni l'information-profondeur de Bennett.

(h) Théories pragmatiques de la valeur de l'information

Si on considère que le but qu'on poursuit est de nature pratique (et déjà avec le programme compilé cet aspect est apparu), par exemple survivre dans un milieu donné, ou gagner le plus possible d'argent tel jour à la bourse de Paris, alors la valeur de l'information contenue dans une chaîne de caractères **s** se mesurera en fonction du but poursuivi. Une information de grande valeur ce sera par exemple l'endroit où aller chercher tel aliment, ou ce sera le nom de l'action boursière qu'il faut acheter car elle va monter.

Aucune théorie générale de l'information n'est possible qui prenne en compte tous les aspects pragmatiques qui déterminent la valeur d'une chaîne de caractères. Une fois reconnu le problème, c'est une évidence.

Que les théories générales puissent jouer un rôle particulier en physique ou en biologie c'est une hypothèse à ne pas écarter et que nous allons examiner maintenant. Mais il faut bien garder à l'esprit qu'aucune ne constitue « la théorie de l'information », que chacune est relative à un but, et que quand il sera question de biologie par exemple il est à prévoir que les aspects pragmatiques ne pourront pas être négligés.

2. La physique et les théories de l'information

Malgré l'évidence d'un lien entre l'information et l'entropie thermodynamique des difficultés diverses font obstacle à l'identification précise de ce lien.

Pour avoir une idée des usages plutôt flous qui sont faits du concept d'information en physique on pourra lire la presse de vulgarisation où une utilisation parfois délirante est faite du mot « information », souvent d'ailleurs associé aux mots : "complexité", "ordre", "désordre", "organisation". Même certains auteurs scientifiques se laissent entraîner par la « force suggestive » des mots. Dans [Reeves, 1986] une image cosmologique de l'évolution de l'information dans l'univers est décrite comme allant de soi, et cela bien que le concept d'information ne soit jamais rendu vraiment précis (dans ce livre comme dans bien d'autres textes il est vaguement assimilé à celui d'entropie physique). Un peu plus prudente est la présentation proposée dans [Barrow, Tipler, 1986] où on trouvera des spéculations intéressantes (car valables au moins pour le contenu brut d'information) sur la quantité maximum d'informations que peut contenir le système solaire (10^{70}) ou sur la possibilité en fonction des différents modèles cosmologiques pour l'univers de contenir ou pas un calcul infini ou une authentique machine de Turing Universelle (voir aussi [Tipler, 1986]). Plus étonnant est [Stonier, 1990] où sans même mentionner la théorie algorithmique de l'information, ce qui est un comble pour un ouvrage aux ambitions si généralistes, on va jusqu'à conjecturer l'existence d'une particule appelée "infon" qui serait le quanton de l'information comme le photon est celui de la lumière; tout cela sans qu'aucune réflexion de fond sur les théories mathématiques ou thermodynamiques de l'information ne soit présentée. Nous n'évoquerons pas certaines élucubrations littéraires (quoi que pensent ceux qui les commentent) où le nombre d'apparitions du mot "information" est inversement proportionnel à la rigueur adoptée pour manipuler le concept.

Plus sérieusement plusieurs problèmes se posent pour introduire le concept d'information en physique. Et il semble que même pour les plus simples d'entre eux aucune solution définitive ne se soit imposée aujourd'hui.

(a) Le problème de l'échelle de base, et du codage

Pour parler d'information (ne serait-ce que du contenu brut d'information) contenu dans un objet physique il faut choisir une échelle de description. Et plus cette échelle est petite plus le contenu en information est grand : les physiciens parlent de « divergence logarithmique » pour évoquer le fait que le contenu brut d'information tend vers l'infini quand on fait tendre vers 0 la résolution de la description physique d'un objet. Si le contenu en information dépend des choix que je fais pour le mesurer, c'est que ce contenu n'est pas "objectif", comme l'est la masse, la température, l'énergie. L'évocation de la mécanique quantique pour choisir une échelle de description est classique, mais ne résout pas complètement de problème. Une discussion très sérieuse de cette question peut être trouvée dans [Zurek, 1989 *b*] où une conclusion prudente est donnée : les notions de la théorie algorithmique de l'information résolvent une partie des problèmes bien qu'une part de subjectivité reste présente. Il est par ailleurs ridicule de se réjouir de cette subjectivité du concept physique d'information brut, car pour que la notion puisse avoir un sens véritable il faudrait au moins qu'elle ait une objectivité faible (intersubjectivité) ce que les concepts de la mécanique quantique possèdent tous (voir par exemple [d'Espagnat, 1985]) et sans laquelle rien n'est possible scientifiquement.

Une autre idée est proposée dans [Lloyd, Pagels, 1988], mais elle paraît difficile à évaluer et ne semble pas avoir été reprise.

(b) Le problème du coût thermodynamique de l'information

Une des premières questions qui se pose à la physique concernant l'information est celle du coût thermodynamique minimum qu'entraîne sa manipulation et son stockage. Sur ce point les choses en sont aujourd'hui tout simplement au stade de la *controverse*.

Pendant plusieurs décennies (à la suite de [Szilard, 1929] [Brillouin, 1959]), on a admis que l'explication du paradoxe du démon de Maxwell (un être microscopique qui en triant les particules d'un gaz réussit à extraire du travail de la chaleur ambiante, et contredit donc le second principe de la thermodynamique), devait être recherchée dans le *coût thermodynamique de toute mesure*.

A la suite des travaux de Landauer, Bennett, Toffoli et Fredkin sur la réversibilité thermodynamique du calcul ([Landauer, 1961, 1985] [Bennett, Landauer, 1985] [Bennett, 1973, 1982, 1988 *c*] [Fredkin, Toffoli, 1982] [Toffoli, 1980, 1982]) cette explication a été contestée : la résolution du paradoxe du démon de Maxwell doit être recherchée non pas dans le coût de la mesure, qui pourrait être rendu nul, mais dans le *coût thermodynamique de l'effacement* de la mémoire du démon des informations qui lui permettent de trier les molécules de gaz, effacement qui doit être opéré à la fin de chaque cycle élémentaire pour restaurer le système dans son état initial (sur ce problème [Bennett, 1988 *a*] est une très bonne introduction).

Il en est résulté une controverse assez violente ([Porod *et al.*, 1984 *a*, 1984 *b*] [Landauer, 1984] [Bennett, 1984] [Benioff, 1984] [Toffoli, 1984] [Barrow, Tipler, 1986]). D'un côté l'école de ceux qui pensent que tout calcul a un coût thermodynamique théorique minimum, et de l'autre ceux qui soutiennent le contraire et qui défendent l'idée que le seul coût théoriquement incompressible de la manipulation d'informations est la remise à zéro des mémoires, et qui proposent des schémas universels de calcul (des « machines de Turing Universelles ») réversibles et ne conduisant donc en théorie à aucune dissipation pour fonctionner (sur les ordinateurs réversibles [Bennett, Landauer, 1985] [Filotti, Mercouroff, 1984] sont de très bonnes introductions). L'intervention du prix Nobel de Physique R. Feynman ([Feynman, 1985]) dans le débat a sans doute donné un avantage aux partisans de la possibilité théorique d'un calcul sans consommation d'énergie.

Un doute reste donc concernant le coût minimal du transfert d'information, sur la solution correcte du paradoxe de Maxwell, et donc en définitive sur le sens thermodynamique de l'information.

Des progrès importants ([Zurek, 1989 *a*, 1989 *b*, 1990 *a*, 1990 *b*] [Caves, 1990] [Delahaye, 1991 *a*] semblent avoir eu lieu récemment qui donneraient raison aux partisans de la possibilité du coût théorique nul du calcul. Cette nouvelle approche de l'entropie physique associe l'entropie de Shannon S (classiquement évoquée par les thermodynamiciens pour fonder la notion d'entropie physique et appelée entropie de Boltzmann-Gibbs-Shannon) au contenu algorithmique en information de Chaitin-Kolmogorof K, ce qui donne une quantité $S+K$ que Zurek propose de reconnaître comme l'*entropie physique*. Elle permet de comprendre du point de vue de l'opérateur ce qu'est une mesure, ceci en échappant au paradoxe classique (que toute mesure fait baisser l'entropie de Boltzmann-Gibbs-Shannon). Cette nouvelle conception physique de l'information, outre qu'elle prouve à nouveau l'importance des concepts de Chaitin-Kolmogorof, risque si elle est acceptée par les thermodynamiciens théoriciens de bouleverser les fondements de la thermodynamique (qui il faut bien le dire ont toujours paru assez peu solides).

Il n'en reste pas moins vrai que même si les rapports entre la thermodynamique et les théories mathématiques de l'information semblent en passe de s'éclaircir, cela ne signifie pas que les progrès en terme de complexité et de richesse d'information dont l'univers semble le siège s'expliqueront seulement par la thermodynamique. D'autres notions d'information doivent être introduites en physique et le sont déjà (Lloyd et Pagels proposent ([Lloyd, Pagels, 1988]) d'introduire une notion de profondeur dérivée de celle de Bennett).

Tout cela est particulièrement clair lorsqu'on réfléchit au contenu en information du génome d'un être vivant.

3. La biologie et les théories de l'information

Il est évident que la biologie moléculaire s'occupe d'informations. La reproduction d'une cellule suppose la duplication d'une quantité substantielle d'informations situées dans le noyau. Le code génétique, les mécanismes de copie, de traduction chimique, de régulation etc., c'est certain comportent une composante « information ».

De nombreux travaux ont été menés pour tenter de rendre compte en termes mathématiques du contenu en information des objets et processus biologiques. Citons l'utilisation de la théorie de Shannon par Atlan [Atlan, 1972]. La théorie algorithmique de l'information a été aussi utilisée par Chaitin lui-même pour proposer une caractérisation formelle d'un être organisé [Chaitin, 1979] et par d'autres ([Küppers, 1990]). L'idée de base de Chaitin est que ce qui caractérise un être organisé E c'est que : la variation de la somme des quantités d'informations (algorithmiques) nécessaires pour décrire E quand on le sépare en morceaux de taille *d*, décroît très rapidement lorsque *d* croît. L'idée quoique très intéressante ne semble pas pour l'instant avoir eu la diffusion qu'elle mérite.

L'utilisation des concepts thermodynamiques en biologie a aussi eu beaucoup de succès. Cette utilisation est souvent très intéressante et ne repose pas forcément sur l'identification de l'entropie physique avec l'information en général, ni même avec l'information au sens de Shannon, identification dont nous avons vu plus haut toutes les difficultés. J. Tonnelat ([Tonnelat, 1978, 1979] et Prigogine ([Prigogine, 1966, 1972]) sont par exemple deux chercheurs ayant travaillé dans cette direction. Il ne semble pas pourtant qu'ils soient arrivés à une conception commune (voir par exemple [Tonnelat, 1979] page 622 qui conteste l'importance attribuée par Prigogine aux structures dissipatives pour expliquer l'existence des êtres vivants). Une certaine prudence est parfois souhaitée dans l'utilisation des concepts de la thermodynamique pour expliquer les phénomènes du vivant : « il est d'un intérêt tout à fait secondaire de savoir si l'entropie d'un être vivant augmente ou diminue au cours de certaines périodes de son existence » ([Tonnelat, 1979] tome II, page 162).

On le voit donc encore à propos de la biologie, ni les théories mathématiques de l'information, ni même les théories thermodynamiques ne résolvent tout.

Je crois en fait qu'il est absurde de vouloir réduire l'information biologique pour la faire entrer dans le cadre d'une théorie mathématique de l'information, ou même de tenter de la réduire à une théorie physique de l'information. On a vu que déjà pour la physique plusieurs théories mathématiques sont amenées à se compléter pour rendre compte des concepts de base de la thermodynamique (ce sont les propositions de Zurek). Il est à prévoir que pour la biologie la situation sera encore plus compliquée. De toutes façons il est évident que la valeur en information du génome par exemple comporte un aspect pragmatique : un être vivant terrestre (dont le « plan est codé dans son génome ») ne peut survivre que dans un contexte spécifié : l'information de son génome n'a pas de valeur absolue.

Est-ce que malgré cela certaines théories mathématiques ou physiques de l'information peuvent servir à la biologie ? La question est posée, et je ne crois pas qu'on ait répondu d'une façon satisfaisante à la question aujourd'hui.

Examinons quelques hypothèses.

Assez vraisemblablement la machine chimique est universelle et donc la complexité de Chaitin-Kolmogorof du génome possède biologiquement une signification. Cependant il n'y a pas de raison de croire que le code génétique d'un individu soit écrit sous la forme la plus compressée possible (on suppose d'ailleurs que de grandes parties du génome ne servent à rien). Le contenu en information algorithmique de Chaitin-Kolmogorof est donc sensiblement inférieur au contenu brut en information et au contenu pragmatique en information.

L'information du génome possède certainement une valeur "à la Bennett" car le génome est le résultat d'une sorte de calcul qui a duré des millions d'années et il n'y a pas à douter que ce calcul n'est pas simplifiable et réductible à quelque chose de court.

La valeur de l'information du génome est aussi une valeur "relative" et "pragmatique". Relative à la chimie et à la physique de notre monde (et pas absolue indépendante de la forme particulière des lois de notre monde physique comme l'est la complexité de Chaitin-Kolmogorof). Pragmatique aussi, car la sélection a joué, et joue encore, et ce qu'elle sélectionne est ce qui est adapté « à un moment donné dans une situation de compétition donnée, dans un écosystème donné » : deux génomes ayant la même complexité de Chaitin-Kolmogorof et la même profondeur de Bennett n'ont pas forcément la même valeur de survie.

Conclusion

Dans tout cela il n'y a peut-être que des évidences, mais il faut les dire car les théories mathématiques disponibles ou qui s'élaborent encore concernant l'information ont des prétentions à l'universalité et à l'applicabilité dont il faut se méfier. Elles sont parfois mal assimilées : celle de Shannon n'a pas apporté grand chose à la biologie et a engendré beaucoup de discours dont aujourd'hui on comprend qu'ils étaient insuffisamment prudents; celle de Chaitin-Kolmogorof semble parfois aussi induire en erreur (le chapitre 10 du livre de Küppers [Küppers, 1990] qui pourtant tente une intéressante utilisation de la théorie algorithmique de l'information apparaît extrêmement hasardeux pour ne pas dire faux [Chaitin, 1990]).

L'attitude de prudence que je préconise n'implique pas la stérilité : souvenons-nous en effet de J. Monod qui ressentant sans doute que tout n'était pas clair du côté des théories physiques de l'information, parlait à propos de ce qu'on leur faisait dire et en particulier de l'identification information = néguentropie « de généralisations et d'assimilations imprudentes » ([Monod, 1970] p. 212). Monod était prudent, avec instinct il se méfiait des spéculations trop hâtives, et pourtant son travail ne fut pas stérile !

En ces domaines risqués à vouloir aller trop vite on risque de ne pas tout voir; comme de plus, ainsi que nous l'avons montré ici, beaucoup d'idées très intéressantes ont été avancées ces dernières années, il serait dommage de passer à côté parce qu'on fonce la tête baissée.

Bibliographie

[ASH, 1965] R. ASH, *Information Theory*, Wiley, New York, 1965.

[ATLAN 1972] H. ATLAN, *l'organisation biologique et la théorie de l'information*, Hermann, 1972.

[ATLAN, 1979] H. ATLAN, *Entre le cristal et la fumée : essai sur l'organisation du vivant*, Editions du Seuil, 1979.

[BARROW, TIPLER, 1986] J.D. BARROW and F.J. TIPLER, *The Anthropic Cosmological Principle*, Oxford University Press, 1986, 1988.

[BENIOF, 1984] P. BENIOF, Comment on « Dissipation in Computation ». *Phys. Rev. Lett.*, *53*, 12, 1984, p. 1203.

[BENNETT, 1973] C.H. BENNETT, Logical Reversibility of Computation. *IBM J. Res. Development*, 1973, p. 525-532.

[BENNETT, 1982] C.H. BENNETT, The Thermodynamics of Computation — a Review, *Int. J. Theoret. Phys.*, *21*, 12, 1982, p. 905-940.

[BENNETT, 1984] C.H. BENNETT, Thermodynamically Reversible Computation, *Phys. Rev. Lett.*, *53*, 12, sept. 1984, p. 1202.

[BENNETT, 1986] C.H. BENNETT, On the Nature and Origin of Complexity in Discrete, Homogeneous, Locally-Interacting Systems, *Found. Phys.*, *16*, 6, 1986, p. 585-592.

[BENNETT, 1987] C.H. BENNETT, Information, Dissipation and the Definition of Organization, In *Emerging Syntheses in Science*, Edited by David Pine. Reading, MA, 1987, p. 215-231.

[BENNETT, 1988 *a*] C.H. BENNETT, Démons, Machines et Thermodynamique, *Pour La Science*, janvier 1988, p. 91-97.

[BENNETT, 1988 *b*] C.H. BENNETT, Logical Depth and Physical Complexity. In *The Universal Turing Machine : A Half-Century Survey*, R. Herken, Ed., Oxford University Press, 1988, p. 227-257.

[BENNETT, 1988 *c*] C.H. BENNETT, Notes on the history of reversible computation. *IBM J. Res. Development, 32*, 1, january 1988, p. 16-23.

[BENNETT, 1990] C.H. BENNETT, How to define complexity in physics, and Why. In *Complexity, Entropy and the Physics of Information*, SFI Studies in the Sciences of Complexity, vol. VIII, W.H. Zurek Ed., Addison-Wesley, 1990, p. 137-148.

[BENNETT, LANDAUER, 1985] C.H. BENNETT, R. LANDAUER, Les Limites Physiques du Calcul, *Pour La Science*, septembre 1985, p. 18-27.

[BRILLOUIN, 1959] L. BRILLOUIN, *La Science et la Théorie de l'Information*, Masson, Paris, 1959.

[CAVES, 1990] C.M. CAVES, *Complexity, Entropy and the Physics of Information*, SFI Studies in the Sciences of Complexity, vol. VIII, W.H. Zurek Ed., Addison-Wesley, 1990.

[CHAITIN, 1966] G.J. CHAITIN, On the Length of Programs for Computing Finite Binary Sequences, *J.A.C.M. 13*, 1966, p. 547-569. Repris dans [CHAITIN, 1987 *a*].

[CHAITIN, 1974] G.J. CHAITIN, Information Theoritic Limitations of Formal Systems, *J.A.C.M.* 1974, p. 403-424. Repris dans [CHAITIN 87 *a*].

[CHAITIN, 1975] G.J. CHAITIN, Randomness and Mathematical Proof. Scientific American, *232*, may 1975, p. 47-52.

[CHAITIN, 1979] G.J. CHAITIN, Toward a mathematical definition of "life". In *The Maximum Entropy Formalism*, R.D. Levine and M. Tribus Eds., MIT Press, 1979, p. 477-498. Repris dans [CHAITIN, 1987 *a*].

[CHAITIN, 1987 *a*] G.J. CHAITIN, *Information, Randomness and Incompleteness* : Papers on Algorithmic Information Theory, World Scientific, Singapore, 1987

[CHAITIN, 1987 *b*] G.J. CHAITIN, Algorithmic information theory. Cambridge Tracts in *Theoret. Comput. Sci.*, 1, Cambridge University Press, New York, 1987.

[CHAITIN, 1990] G.J. CHAITIN, Communication personnelle : 27 novembre 1990.

[CULLMANN *et al.*, 1970] G. CULLMANN, M. DENIS-PAPIN, A. KAUFMANN, *Cours de calcul informationnel appliqué : Introduction à l'étude de l'informatique*, Seconde édition, Albin Michel, Paris, 1970.

[D'ESPAGNAT, 1985] B. d'ESPAGNAT, *Une Incertaine Réalité : Le Monde Quantique, la Connaissance et la Durée*, Gauthier-Villars, Paris, 1985.

[DELAHAYE, 1988] J.-P. DELAHAYE. Une Extension Spectaculaire du Théorème de Gödel. *La Recherche, 200*, juin 1988, p. 860-862.

[DELAHAYE, 1989 *a*] J.-P. DELAHAYE, *Cinq Classes d'Idées*. Rapport. Laboratoire d'Informatique Fondamentale de Lille, Univ. Sc. Lille, Bât. M3, 59655 Villeneuve d'Ascq, avril 1989.

[DELAHAYE, 1989 *b*] J.-P. DELAHAYE, *Chaitin's Equation : An Extension of Gödel's Theorem*. Notices of The American Mathematical Society, october 1989, p. 984-987.

[DELAHAYE, 1990] J.-P. DELAHAYE, Randomness, Unpredictability and Absence of Order : the Identification of the Mathematical Notion of Random Sequence by the Theory of Recursivity, *Colloque International sur la Philosophie des Probabilités*, organisé par le CNRS et l'Institut d'Histoire des Sciences de Paris, 10-11-12 mai 1990, Paris, Kluwer, à paraître.

[DELAHAYE, 1991 *a*] J.-P. DELAHAYE, Thermodynamique et informatique théorique : une nouvelle définition de l'entropie physique, *Pour La Science* (édition française du Scientific American), avril 1991, p. 17-20.

[DELAHAYE, 1991 *b*] J.-P. DELAHAYE, Complexités. La profondeur logique selon Bennett, *Pour La Science* (édition française du Scientific American), août 1991, p. 102-104, 112.

[ERIKSSON *et al.*, 1987] K.-E. ERIKSSON, K. LINDGREN, B.A. MASSON, *Structure, Context, Complexity, Organisation : Physical Aspects of Information and Value*. World Scientific, Singapore, 1987.

[FEYNMAN, 1985] R. FEYNMAN, Quantum Mechanical Computers. *Optics News, 11*, 1985, p. 11-20.

[FILOTTI, MERCOUROFF, 1984] I. FILOTTI, W. MERCOUROFF, Des ordinateurs sans consommation d'énergie, *La Recherche*, juillet-août 1984, p. 1016-1018.

[FREDKIN, TOFFOLI, 1982] E. FREDKIN, T. TOFFOLI, Conservative Logic. *Int. J. Theoret. Phys., 21*, n° 3/4, 1982, p. 219-253.

[KOLMOGOROF, 1965] A.N. KOLMOGOROF, Three approaches for defining the concept of information quantity, *Inform. Transmis., 1*, 1965, p. 3-11.

[KOLMOGOROF, 1968] A.N. KOLMOGOROF, Logical basis for Information Theory and Probability Theory. *IEEE Trans. Inform. Theory, IT14*, 5, 1968, p. 662-664.

[KÜPPERS, 1990] B.-O. KÜPPERS, *Information and the origin of life*, MIT Press, Cambridge, Massachusetts, 1990.

[LANDAUER, 1961] R. LANDAUER, Irreversibility and heat generation in the computing process, *IBM J. Res. Development, 5*, 3, 1961.

[LANDAUER, 1984] R. LANDAUER, Dissipation in Computation, *Phys. Rev. Lett., 53*, 12, 1984, p. 1205.

[LANDAUER, 1985] R. LANDAUER, Fundamental Physical Limitations of the Computational Process. *Annals of the New York Academy of Sciences, 426*, 1985, p. 161-170.

[LI, VITANYI, 1990] M. LI, P.M.B. VITANYI, Application of Kolmogorof Complexity in the Theory of Computation. In *Complexity Theory Retrospective*, Alan Selman Ed., Springer Verlag, 1990, p. 147-203.

[LI, VITANYI, 1991] M. LI, P.M.B. VITANYI, *Kolmogorof Complexity and Its Applications*, Handbook of Theoritical Computer Science, J. van Leeuwen Ed., North-Holland, 1991.

[LLOYD, PAGELS, 1988] S. LLOYD and H. PAGELS, Complexity as Thermodynamic Depth, *Ann. Phys., 188*, 1988, p. 186-213.

[MARTIN, ENGLAND, 1981] N.F.G. MARTIN, J.W. ENGLAND, *Mathematical Theory of Entropy*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1981.

[MONOD, 1970] J. MONOD, *Le hasard et la nécessité : essai sur la philosophie naturelle de la biologie moderne*, éditions du Seuil, Paris, 1970.

[POROD *et al.*, 1984*a*] W. POROD, R. GRONDIN, D. FERRY, G. POROD, On the possibility of thermodynamically reversible computation. *Phys. Rev. Lett., 52*, 1984, 232.

[POROD *et al.*, 1984*b*] W. POROD, R. GRONDIN, D. FERRY, G. POROD, Dissipation in Computation. *Phys. Rev. Lett., 53*, 1984, 1205-1206.

[PRIGOGINE, 1966] I. PRIGOGINE, *Introduction of thermodynamics of irreversible processes*, J. Wiley, 1966.

[PRIGOGINE, 1972] I. PRIGOGINE, La thermodynamique de la vie, *La Recherche*, 1972, p. 547-558.

[RUCKER, 1987] R. RUCKER, *Mind Tools : The Five Levels of Mathematical Reality*, Houghton Mifflin Company, Boston, 1987, p. 243-247.

[SHANNON, WEAVER, 1949] C.E. SHANNON, W. WEAVER, *The Mathematical Theory of Communication*, Univ. of Illinois Press, Urbana, 1949.

[STONIER, 1990] T. STONIER, *Information and the Internal Structure of the Universe : An Exploration into Information Physics*, Springer-Verlag, Heidelber, 1990.

[SZILARD, 1929] L. SZILARD, Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen, *Z. Physik., 53*, 1927, p. 840-856.

[TIPLER, 1986] F.J. TIPLER. Cosmological Limits on Computation. *Int. J. Theoret. Phys., 25*, 6, 1986, p. 617-661.

[TOFFOLI, 1980] T. TOFFOLI, *Reversible Computating*, RR. Laboratory For Computer Science, Massachusetts Institute of Technology, 1980.

[TOFFOLI, 1982] T. TOFFOLI, Physics and Computation. *Int. J. Theoret. Physics, 21*, 12, 1982, p. 165-175.

[TOFFOLI, 1984] T. TOFFOLI. Comment on « Dissipation in Computation », *Phys. Rev, Letters, 53*, 12, sept. 1984, p. 1204.

[TONNELAT, 1978] J. TONNELAT. *Thermodynamique et Biologie*. Tome I : *Entropie, désordre et complexité*. Tome II : *L'ordre issu du hasard*, Maloine-Doin, Paris, 1978.

[TONNELAT, 1979] J. TONNELAT, Qu'est-ce qu'un être vivant? *La recherche, 101*, 1979, p. 614-622.

[TURING, 1936] A.M. TURING. On Computable Numbers, with an application to the Entscheidungsproblem. *Proceeding of the London Mathematical Society, 2*, 42, 1936-7, p. 230-265, corrections, 43, 1937, p. 544-546.

[WELSH, 1988] D. WELSH, *Codes and Cryptography*. Oxford Science Publications, Clarendon Press, Oxford, 1988.

[ZUREK, 1989*a*] W.H. ZUREK. Thermodynamic cost of computation, algorithmic complexity and the information metric, *Nature, 341*, 1989, p. 119-124.

[ZUREK, 1989*b*] W.H. ZUREK. Algorithmic randomness and the physical entropy, *Physical Rev. A., 40*, 1989, p. 4731-4751.

[ZUREK, 1990*a*] W.H. ZUREK, Algorithmic Randomness, Physical Entropy, Measurements and the Second Law. In *Proceedings of International Symposium on Quantum Mechanics*, edited by Y. Murayama, Tokyo, Physical Society of Japan, 1990, p. 115-123.

[ZUREK, 1990*b*] W.H. ZUREK, Algorithmic Information Content, Church-Turing Thesis, Physical Entropy, and Maxwell's Demon. Lectures in Complex Systems, SFI Studies in the Sciences of Complexity, Lect. vol. II. Erica Jen Ed., Addison-Wesley, 1990, p. 49-65. Aussi in : Complexity, Entropy and the Physics of Information SFI Studies in the Sciences of Complexity, vol. VIII, Ed. W.H. Information, SFI Studies in the Sciences of Complexity, Vol. VIII, W.H. Zurek Ed., Addison-Wesley, 1990, p. 73-89.