Séminaire de philosophie et mathématiques

Jean-Paul Delahaye

Le concept de suite aléatoire et la thèse de Church

Séminaire de Philosophie et Mathématiques, 1991, fascicule 3
« Le concept de suite aléatoire et la thèse de Church », , p. 1-35

<http://www.numdam.org/item?id=SPHM_1991___3_A1_0>

Novembre 1991

Le concept de suite aléatoire et la thèse de Church

Jean-Paul DELAHAYE.
Université des Sciences et Technologies de Lille
Laboratoire d'Informatique Fondamentale de Lille
(U. A. CNRS 369 Bât M3)
59655 Villeneuve d'Ascq Cedex. FRANCE
e-mail : delahaye@liff.lifl.fr

Résumé

La théorie de la calculabilité dont les bases furent posées entre 1930 et 1936 (incomplétude de Gödel 1930, thèse de Church 1934, machines de Turing 1936) permet de formuler une définition précise de la notion de suite infinie aléatoire. Cette définition réalise sans doute les objectifs que von Mises s'était fixé sans les atteindre d'une manière satisfaisante. Après avoir rappelé les termes du problème, nous retraçons brièvement cette histoire en en décrivant les phases successives qui sont : (i) la réflexion autour de la notion de collectif de von Mises, (ii) la mise au point de la première théorie algorithmique de l'information par Solomonof, Kolmogorov et Chaitin, (iii) la formulation de la bonne définition de suite aléatoire par Martin-Löf en 1966, (iv) la seconde théorie algorithmique de l'information par Levin, Chaitin, et Schnorr, mettant en évidence les liens entre entropie et complexité et fournissant des arguments définitifs en faveur de la définition de Martin-Löf. La situation actuelle est ensuite exposée et un parallèle est proposé entre la thèse de Church et ce que nous appelons la thèse de Martin-Löf. Les quelques doutes qui persistent à propos de la thèse de Martin-Löf sont évoqués pour terminer.

Le concept de suite aléatoire et la thèse de Church

J.-P. Delahaye

Plan.

Introduction.

(a) Une identification difficile.

Depuis quelques années les progrès importants qui ont été réalisés dans la compréhension mathématique du hasard ont conduit à la résolution du problème que von Mises posait : la formulation d'une définition purement mathématique de la notion de suite aléatoire. Ces progrès résultent de l'utilisation du concept de fonction récursive. Une fois percue la nécessité de faire intervenir la théorie de la récursivité on peut distinguer plusieurs époques dans ce développement.

La première voit la formulation et l'étude de définitions diverses des collectifs de von Mises dont on peut alors réaliser —grâce à une certaine catégorie de résultats mathématiques négatifs— qu'elles ne sont pas satisfaisantes; c'est l'échec de la tentative de von Mises.

La seconde qui commence en 1965 est ce qu'on peut appeler l'époque de la **première théorie algorithmique de l'information.** Elle est caractérisée par une multitude de définitions du hasard absolu ne coïncidant pas bien les unes avec les autres, mais sans qu'on sache établir pourquoi, ni quelle définition doit être retenue en définitive. Les travaux de Chaitin, Kolmogorov, Loveland, Martin-Löf, Levin, et Schnorr conduisent à la troisième époque.

Cette troisième époque est celle de la mise au point de ce qu'on peut appeler la **deuxième théorie algorithmique de l'information,** développée en parallèle par Chaitin, Levin et Schnorr et qui conduit alors à une entente sur le concept de suite infinie aléatoire et qui correspond à ce que j'appellerai thèse de Martin-Löf

(b) L'importance de la notion de fonction récursive.

L'identification mathématique de la notion intuitive d'algorithme a été un succès inattendu de la logique mathématique (1930-1936) et ce succès, en plus de ce qu'il apporte à la théorie des ordinateurs et du calcul (résultats d'indécidabilité, équivalence des capacités calculatoires de mécanismes divers, identification de l'importance des boucles à sorties conditionnelles) a permis l'élaboration d'un concept mathématiquement clair et délimité de **système formel**, essentiel en logique mathématique où il permet de saisir pleinement le sens des théorèmes d'incomplétude de Gödel. L'identification mathématique de la notion intuitive d'algorithme a finalement conduit, comme nous allons le voir, à une autre identification mathématique importante, celle de **suite infinie aléatoire**, tout cela au travers d'un processus d'élaboration complexe qui s'est déroulé sur plus de cinquante ans (de 1919 à 1975).

Un parallèle intéressant peut d'ailleurs être fait entre ce qui s'est passé avec la thèse de Church (aussi appelée thèse de Church-Turing) et ce qui s'est passé pour la notion de suite infinie aléatoire. L'histoire est moins simple car elle a été rendue confuse par de multiples facteurs. L'un d'eux est sans doute les prétentions peu justifiées de rigueur mathématique de von Mises, un autre est simplement la difficulté même des problèmes mathématiques sous-jacents qui a empêché la formulation rapide de la bonne "thèse" pour la notion de suite infinie aléatoire.

(c) Les thèses concurrentes.

Assez schématiquement on peut dire que l'histoire de cette identification tourne autour de plusieurs thèses (rarement présentées sous ce terme mais assez clairement énoncées par leurs défenseurs) dont certaines se sont révélées inadéquates alors que d'autres étaient démontrées équivalentes. C'est l'un des points qui distingue la théorie des fonctions récursives (où une seule thèse fut présentée, puis plus tard confirmée et enfin universellement admise) de celle des suites infinies aléatoires : l'intuition de plusieurs mathématiciens les fit défendre à certains moments des propositions que des arguments mathématiques forcèrent à abandonner plus tard. Ceci prouve d'ailleurs que, même en ces domaines non absolument mathématiques, tout n'est pas joué d'avance : une proposition d'identification entre un concept intuitif et un concept mathématique peut se révéler indéfendable en regard de progrès de nature purement mathématique.

Parmi les 'thèses' proposées de mathématisation de la notion de suite infinie aléatoire, on peut distinguer:

(remarque : nous nous limitons au cas des suites infinies de 0 et de 1 pour lesquelles 0 et 1 ont des rôles symétriques, et qui sont les suites associées à un lancé de pièces non biaisées; les cas plus généraux se déduisent assez facilement de celui-là)

- **la 'pseudo-thèse' de von Mises (1919) :** Les suites infinies aléatoires sont les collectifs de von Mises, c'est-à-dire les suites infinies de 0 et de 1 dont les fréquences limites de 0 et de 1 de toute suite extraite par un procédé 'raisonnable' sont égales à 1/2 . Autrement dit pour qu'une suite infinie soit considérée comme aléatoire on veut qu'à la limite il y ait autant de 0 que de 1, mais aussi qu'il y ait autant de 0 que de 1, si on ne considère qu'un terme sur trois de la suite ou si on considère une sous-suite quelconque 'raisonnablement' extraite de la suite de départ. C'est ce 'raisonnablement' que von Mises ne réussit pas à mathématiser d'une manière satisfaisante, quoi qu'il ait prétendu, et qui nous conduit à parler de 'pseudo-thèse' plutôt que de thèse.

- **la thèse de von Mises-Church (1940) :** Les suites infinies aléatoires sont les collectifs de von Mises en considérant que les sous-suites extraites à prendre en compte doivent être définies par un procédé de calcul effectif, c'est-à-dire par une fonction récursive (totale ou partielle). Il s'agit bien cette fois-ci d'une proposition d'identification mathématique, donc d'une thèse dans le sens où on parle de la thèse de Church. (Mise en garde : ne pas confondre la **thèse de Church**, appelée aussi thèse de Church-Turing, qui affirme que les fonctions calculables par algorithmes sont les fonctions récursives, et ce que nous appelons ici **thèse de von Mises-Church** qui concerne les suites infinies aléatoires)

- **la thèse de Kolmogorov-Chaitin (1965) :** Les suites infinies aléatoires sont les suites infinies dont la complexité de Kolmogorov-Chaitin est maximale. Une idée naturelle (mais conduisant à une notion vide) consiste à préciser cela de la manière suivante : les suites infinies aléatoires sont les suites $x = (x_0, x_1, .., x_n, ...)$ (notation $x^n = (x_0, x_1, .., x_n)$) vérifiant qu'il existe c tel que pour tout entier n : $K(x^n) > n - c$ (voir paragraphe 2 (a) pour $K(x^n)$ qui désigne la complexité de Kolmogorov-Chaitin de x^n).

- **la thèse de Martin-Löf (1966) :** Les suites infinies aléatoires sont les suites satisfaisant tous les "tests statistiques effectifs" (voir définition paragraphe 2 (c)) ou, ce qui revient au même, les suites non-aléatoires constituent un ensemble de mesure nulle constructif maximal.

- **La thèse de Chaitin-Levin (1975) :** Les suites infinies aléatoires sont les suites dont la complexité de Chaitin-Levin (par programmes auto-délimités) est maximale; de

manière plus précise, sont les suites vérifiant qu'il existe c tel que pour tout n : $H(x^n) > n - c$ (voir paragraphe 2 (d) pour $H(x^n)$ qui désigne la complexité de Chaitin-Levin de x^n).

Les thèses de von Mises-Church et de Kolmogorov-Chaitin se sont révélées insatisfaisantes (l'une trop tolérante, l'autre trop stricte puisque définissant une classe de suites vide), celles de Martin-Löf et Chaitin-Levin se sont révélées équivalentes entre elles et à d'autres encore. La thèse de Martin-Löf après une période un peu confuse de mise à l'épreuve et de mise en concurrence avec d'autres thèses, est donc acceptée, et cela vraisemblablement de manière définitive.

(d) Organisation de la présentation.

Pour raconter cette histoire assez étonnante de la difficile élaboration du versant formel et mathématique d'un concept philosophique et épistémologique délicat, nous commencerons par exposer le problème de départ qui est celui de l'analyse de la notion intuitive de suite statistiquement quelconque, ou suite infinie aléatoire (§ 1).

Nous parcourons alors rapidement les étapes principales de cette recherche du bon concept de suite aléatoire (§2). Nous commençons par la solution formulée par von Mises et de ses rapports avec la formalisation des probabilités par Kolmogorov, puis la façon dont elle fut rendue rigoureuse par Church et Loveland mais aussi éliminée par Ville. La première théorie algorithmique de l'information de Kolmogorov-Chaitin est alors présentée rapidement, ainsi que la solution de Martin-Löf, puis la mise au point de la théorie algorithmique de l'information de Chaitin-Levin, avec pour conséquence la reconnaissance sans doute définitive de la définition de Martin-Löf comme satisfaisante.

Un résumé de la situation actuelle est alors donnée (§3).

Une comparaison de la thèse de Martin-Löf avec celle de Church est détaillée (§4), puis nous évoquons les doutes qui persistent (§5).

1 L'analyse du concept de suite infinie aléatoire.

Le problème (philosophico-mathématique) de départ est assez simple à formuler: parmi toutes les suites infinies de 0 et de 1, quelles sont celles qu'on doit ou qu'on peut considérer comme aléatoires, c'est-à-dire comme pouvant provenir de la suite des lancers successifs d'une pièce de monnaie non biaisée? (répétons-le, dans tout le texte nous nous restreindrons à ce type de situations, tout en sachant que le cas des pièces biaisées, ou le cas des tirages donnant plus de deux résultats possibles se traite assez facilement dès que le cas d'une pièce symétrique est résolu).

Le problème paraît clair : il y a des suites dont on ne peut pas croire raisonnablement qu'elles sont le résultat d'un tirage au sort. La suite (0 1 0 1 0 1 0 1 0 1 0 1 ...) n'est certainement pas une suite infinie aléatoire. Il y a d'autres suites par contre dont on peut penser au moins momentanément qu'elles sont le résultat d'un tirage au hasard non truqué : par exemple la suite (0 1 0 1 1 0 1 0 0 1 1 0 0 1 0 0 1 0 1 1 1 0 0 1 0 0 0 0 1 0 1 1 0 1 0 1 1 0 1 0 1 1 ...). Et pourtant si on découvrait qu'une telle suite est le développement binaire du nombre π, on changerait alors son opinion, pour conclure que la suite n'est pas aléatoire.

Et donc, une première analyse rapide ne nous assure pas qu'il soit possible de définir la notion de suite infinie aléatoire : peut-être après tout, que n'importe quelle suite vue

d'une certaine façon présente des régularités telles qu'il est impossible de la considérer comme aléatoire. Peut-être donc, que la notion de suite infinie aléatoire n'est qu'une chimère et n'a aucune contrepartie mathématique, et aucune consistance véritable (comme on l'a cru un moment des 'infiniment petits' réhabilités par l'analyse non standard il y a moins de trente ans). La situation après ces considérations est loin d'être claire et l'issue paraît incertaine. Au moins une fois dans l'histoire de ces tentatives, il est d'ailleurs arrivé que tous les spécialistes intéressés doutent qu'une définition satisfaisante de la notion de suite infinie aléatoire puisse exister. Martin-Löf [Martin-Löf 1969 page 33] indique qu'en 1939 Ville, soutenu par Borel, Frechet et Lévy ne croyait plus en la possibilité d'une telle définition. Et pourtant, finalement, grâce à un sauvetage in-extremis par le concept de récursivité les tentatives aboutirent à un résultat satisfaisant.

Reprenons notre analyse.

(a) L'aléatoire comme absence de loi effective de production.

Bien sûr nous souhaiterions que notre définition permette de dire que la suite (0 1 0 1 0 1 0 1 0 ...) n'est pas aléatoire. De même que la suite (0 1 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 1 ...) ou que la suite partout nulle sauf pour les rangs premiers: (0 0 1 1 0 1 0 1 0 0 0 1 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 1 ...) ne sont pas aléatoires. Plus généralement on aimerait bien que les suites aléatoires ne vérifient aucune règle 'détectable' précise, c'est-à-dire qu'il n'y ait aucune loi déterminée qui en donne tous les termes un par un. Nous voudrions que les suites aléatoires soient absolument *sans forme identifiable*, qu'elles soient, autant que c'est possible, *atypiques* et *a-structurées*.

La difficulté pour pouvoir en tirer une définition mathématique est qu'on doit d'abord définir mathématiquement ce qu'est une "structure", une "loi pouvant en donner les termes un par un". On voit immédiatement que si on est trop tolérant avec la notion de loi – par exemple si on définit une loi comme une application qui à tout n associe 0 ou 1 – alors il n'y a aucune suite aléatoire (puisque toute suite sera définie par une loi) et donc la notion qu'on cherche à préciser est vide. A l'inverse si on essaie simplement d'exclure certains types de régularités élémentaires (périodicité, périodicité à partir d'un certain rang, présence des 1 fixée par polynômes etc.) on n'arrive pas non plus à une définition satisfaisante.

Une fois la notion de fonction récursive disponible et la thèse de Church acceptée, il devient naturel de tenter la définition :

> *est aléatoire toute suite infinie dont les digits ne sont donnés par aucune fonction récursive (totale),*
>
> *(variante : est aléatoire toute suite infinie dont l'ensemble des rangs des 1 n'est pas récursivement énumérable)*

On s'apercoit rapidement que les suites vérifiant de telles définitions (qu'on appelle "suites non récursives" et suites "non récursivement énumérables") ne répondent pas à notre attente. En effet de telles suites x ne vérifient pas toujours que la fréquence s_n/n des 1 parmi les n premiers éléments ($s_n=x_0+x_1+ ...+x_{n-1}$) tend vers 1/2 quand n tend vers l'infini, ce qui paraît être une condition nécessaire pour qu'une suite soit considérée comme aléatoire.

Les suites qui sont simplement "sans structure déterministe effective" (i.e. dont aucune loi récursive ou récursivement énumérable n'en produit les termes un par un) sont trop nombreuses et ne correspondent pas à ce que intuitivement nous concevons comme aléatoire. L'utilisation simple de la thèse de Church et de la notion de récursivité ne résout pas le problème.

Cette insuffisance de la définition de l'aléatoire comme simple absence de loi d'engendrement calculable est sans doute apparue immédiatement aux créateurs de la théorie des fonctions récursives (Gödel, Church, Kleene, Turing, Post, Markov), aussi est-ce pourquoi à ma connaissance, cette utilisation naïve de la récursivité ne fut jamais proposée.

On peut d'ailleurs remarquer que toute définition raisonnable de "loi de production effective" ne peut en considérer qu'une infinité dénombrable, et qu'en conséquence toute définition basée sur une certaine "effectivation" de la notion de loi de production ne pouvait exclure de l'ensemble des suites aléatoires qu'une infinité dénombrable de suites, ce qui est manifestement insuffisant, car par exemple l'ensemble de toutes les suites telles que $x_{2n} = x_{2n+1}$ pour tout n, est non dénombrable, quoique qu'aucune de ces suites ne puisse être considérée comme aléatoire.

Même si, ainsi que nous allons le voir plus loin, la solution du problème s'appuie d'une façon essentielle sur la théorie des fonctions récursives, elle n'en est pas une conséquence immédiate. Un long travail doit être fait.

(b) L'aléatoire défini à partir des fréquences limites.

Le problème de la limite des fréquences relatives de 1 et de 0 (qui est une des raisons de l'échec de la tentative (a)) introduit naturellement l'idée du second type de tentatives.

Puisqu'une suite ne peut être considérée comme aléatoire que si elle vérifie que: s_n/n converge vers 1/2, imposons cette condition. A elle seule bien sûr, elle est insuffisante car par exemple la suite (0 1 0 1 0 1 0 1 0 1 ...) satisfait cette condition mais n'est pas aléatoire. Tentons donc de la généraliser.

Plusieurs idées ont été proposées pour cela. Présentons-les (non par ordre chronologique, mais par ordre de force restrictive croissante).

(b1) Tentative de Popper (1935).

Une suite infinie aléatoire doit vérifier que:

> *quelle que soit la séquence finie* $a_1\ a_2 \ldots a_n$ *de* 0 *et de* 1, *la suite extraite à partir de la suite de départ en ne gardant que les digits qui suivent l'apparition de* $a_1\ a_2 \ldots a_n$ *doit avoir une suite de fréquences convergente vers* 1/2.

Nous appellerons suite indifférente [Ville 1939] une suite de 0 et de 1 ayant cette propriété. Une définition équivalente avait été formulée par Copeland [Copeland 1928, 1932].

Cette définition qui est une forme faible de celle considérée par von Mises (et rendue formelle par Church en 1940) fut révélée insatisfaisante par Ville ([Ville 1939]) En fait elle est réellement trop faible car elle permettrait de considérer comme aléatoires des suites que naturellement nous ne voulons pas accepter comme telles : avec cette définition certaines suites ayant toujours plus de 1 que de 0 parmi les n-premiers digits seraient considérées comme aléatoires.

(b2) Tentative par la normalité.

Une autre façon de généraliser la condition sur la limite de fréquence des 0 et des 1 est de l'étendre à toute combinaison de 0 et de 1, c'est-à-dire de considérer qu'une suite aléatoire doit vérifier que :

quelle que soit la séquence finie $a_1\ a_2 \ldots a_n$ *de* 0 *et de* 1, *la fréquence de son apparition converge et a pour limite* $1/2^n$ *(en particulier la fréquence des* 1 *doit être* 1/2, *la fréquence d'apparition de* 01 *doit être* 1/4, *la fréquence d'apparition de* 0100 *doit être* 1/16 *etc).*

Les nombres réels ayant pour développement binaire une suite satisfaisant cette condition sont appelés nombres normaux en base 2 ([Borel 1909]). Les suites satisfaisant cette propriété peuvent donc être appelées normales. Si on généralise la condition au cas d'une pièce à faces non équiprobables de probabilités (p, 1-p) en demandant que la fréquence d'apparition de $a_1\ a_2 \ldots a_n$ soit $p^m\ (1-p)^{n-m}$ où m est le nombre de 1 dans la séquence $a_1\ a_2 \ldots a_n$ on obtient la notion de suite de Bernouilli.

Bien que fondées sur des idées distinctes, les deux notions de suites indifférentes et de suites normales sont équivalentes (voir par exemple [Ville 1939]).

En envisageant des développements dans d'autres bases que 2 (i.e. en écrivant en base b le nombre réel dont le développement binaire est la suite considérée) on peut renforcer encore la condition. De toute manière cette définition se révèle insatisfaisante car des suites normales ayant une loi d'engendrement très simple (et donc non-aléatoires) furent proposées [Champernowne 1933]. La propriété de normalité ne peut pas convenir pour définir la notion de suite infinie aléatoire, elle est nécessaire mais n'est certainement pas suffisante (pour plus de détails sur ce point voir [Martin-Löf 1969]).

Remarque : Dans un article récent, M. O'Connor [O'Connor 1988] a montré que les suites normales en base 2 sont exactement les suites qui sont imprédictibles par des automates d'états finis, et que de plus cette classe de suites est stable par extraction effectuée par des automates d'états finis. Si on admettait (ce qui n'est pas le cas en général et est contraire à la thèse de Church) que la bonne notion de calculabilité est celle définie par les automates d'états finis alors tout serait parfaitement simple : la notion de suite infinie aléatoire serait celle de suite normale en base 2.

(b3) Tentative par extraction de von Mises.

Comme dans la tentative de Popper, l'idée de von Mises fut d'exiger que la propriété de convergence des fréquences pour les suites extraites soit préservée quel que soit le procédé d'extraction "raisonnable" ("procédé de sélection de places admissible"). Tout le problème est de définir ce que signifie "raisonnable". Et c'est ce problème que von Mises ne résolut jamais vraiment.

Si on est trop tolérant dans les procédés d'extraction, il n'y a aucune suite aléatoire : la notion de suite aléatoire est inconsistante. Si on est trop peu tolérant alors seront considérées comme aléatoires des suites qui ne le sont pas au sens intuitif.

L'inconsistance de la définition de von Mises a été depuis longtemps dénoncée. Ville [Ville 1939] par exemple s'exprime comme suit :

> Etant donnée une suite d'entiers croissants (n_i), le procédé qui d'une suite $x= x_1x_2...x_n...$ où $x_n = 0$ ou 1) extrait la suite partielle $y = x_{n1}\ x_{n2} \ldots x_{nm}\ldots$ est bien une sélection. Or, si nous considérons toutes les suites (n_i) possibles, quelle que soit x contenant une infinité de 1 et de 0, il existera une suite (n_i) telle que l'on ait identiquement $x_{ni}=1$ et une autre suite (m_i) telle que l'on ait identiquement $x_{mi}=0$. On en déduit qu'il n'existe aucune suite satisfaisant à la définition du collectif, telle que l'a présentée von Mises. J. Ville. Etude critique de la notion de collectif. Gauthier-Villars. Paris. 1939.

Les événements qui suivirent montrèrent qu'en fait toutes les tentatives pour définir les suites infinies aléatoires par une condition sur la convergence de la fréquence d'apparition des blocs de digits, étaient soit trop peu exigentes, soit trop exigentes. Même celle de Church [Church 1940] fondée sur la notion de suite extraite par un procédé d'extraction algorithmique s'est révélée trop faible. Church considère qu'un procédé d'extraction est raisonnable s'il existe une fonction récursive indiquant pour toute suite finie donnée (x_0, x_1, ...x_n) si on va garder x_{n+1} ou pas. D'une suite infinie fixée, les sous-suites que Church accepte d'extraire sont donc dénombrables. Les deux résultats suivant s'y appliquent.

Deux résultats importants doivent être retenus sur cette question, celui de Wald et celui de Ville.

Le résultat de Wald [Wald 1937] indique que si l'ensemble des règles d'extraction pris en compte est dénombrable alors il existe une infinité non dénombrable de suites satisfaisant les conditions de limite de fréquence de von Mises, autrement dit : à la condition de ne pas considérer trop de suites extraites, la notion de collectif de von Mises est consistante. En particulier la notion de suite aléatoire de von Mises-Church est consistante.

Le second résultat dû à Ville [Ville 1939] indique que si l'ensemble des règles d'extraction pris en compte est dénombrable alors il existe des suites vérifiant toutes les conditions de limite de fréquence mais qui sont telles que le nombre de 1 est supérieur au nombre de 0 dans tout début de la suite, ce qui implique en particulier que la loi du logarithme itéré n'est pas satisfaite et que donc la suite n'a pas toutes les propriétés qu'on s'attend à trouver pour une suite aléatoire. Ce résultat de Ville a été généralisé par van Lambalgen en 1987, qui a montré que "la plupart" des suites aléatoires au sens de von Mises-Church ne le sont pas au sens de Martin-Löf autrement dit la notion de von Mises-Church est bien plus faible que celle finalement retenue.

Là encore on le voit il ne suffit pas de faire intervenir la notion de fonction récursive pour obtenir une bonne définition. Et la voie empruntée par von Mises finalement s'est révélée sans issue. Rétrospectivement on ne comprend pas très bien pourquoi von Mises s'est accroché à cette idée d'extraction alors que d'autres voies comme celle des tests statistiques sont au moins aussi naturelles.

Une étude récente de van Lambalgen [1987] a tenté d'approfondir la notion de collectif "à la von Mises". La réflexion a porté sur la notion de "méthode de sélection raisonnable" et quelques résultats intéressants ont été obtenus. Il faut aussi noter que Kolmogorov lui-même dans un article de 1987 [Kolmogorov Uspenskii 1987] considérait comme peut-être envisageable une caractérisation satisfaisante des suites infinies aléatoires par une méthode "à la von Mises". L'article de [Shen 1989] a donné une réponse négative à la possibilité que Kolmogorov envisageait encore en 1987, mettant sans doute un terme définitif à tout espoir de trouver une définition satisfaisante du concept de suite aléatoire en terme d'extraction (sur ce sujet voir aussi [van Lambalgen 1990])

(c) L'aléatoire défini comme satisfaction à tous les tests statistiques effectifs.

Les statisticiens utilisent des tests pour décider si une suite de données est aléatoire ou non. Ces tests ne pourraient-ils pas être formalisés et ainsi conduire à une définition des suites infinies aléatoires comme *suites satisfaisant tous les tests statistiques envisageables par les statisticiens*. Là encore, sans la théorie des fonctions récursives, rien ne serait possible car la notion de test statistique telle qu'on peut la définir naïvement en théorie des ensembles est si large qu'aucune suite ne peut satisfaire à elle seule tout test abstrait envisageable, et au bout du compte donc conduit à une notion vide. Seule une version effective (donc fondée sur la thèse de Church et la théorie des fonctions récursives)

pouvait mener à une notion consistante de suites infinies satisfaisant tous les tests statistiques effectifs.

C'est la voie retenue par Martin-Löf et Schnorr. Elle conduisit Matin-Löf [Martin-Löf 1966] très rapidement à une définition qui bien que contestée finit par emporter la victoire. Nous la présentons en détail plus loin.

(d) L'aléatoire défini comme imprévisibilité absolue.

L'idée de von Mises était qu'en imposant son invariance des fréquences limites par extraction, il exprimait l'impossibilité d'un pari gagnant contre la suite (ce qui d'ailleurs n'est pas exact vu le résultat de Ville rappelé plus haut).

Mais l'idée même de fonder la notion de suite infinie aléatoire sur l'impossibilité d'un pari gagnant contre la suite peut être abordée d'une manière directe, sans passer par le biais de l'invariance des fréquences limites pour les suites extraites. Une formalisation de la notion de martingale fut proposée et étudiée par Ville en 1939 [Ville 1939], malheureusement il ne put aboutir à une proposition précise et consistante de suites infinies aléatoires par l'utilisation de ses concepts car il ne sut pas utiliser la thèse de Church et la théorie des fonctions récursives tout juste disponibles depuis 1936, et sans laquelle, là encore, il semble qu'il soit impossible d'aboutir. Il montra cependant qu'à toute martingale d'un certain type correspondait un ensemble de mesure nulle (dans l'ensemble de toutes les suites infinies de 0 et de 1) où le gain était infini, et que réciproquement à tout ensemble de mesure nulle on peut associer une certaine martingale (selon sa définition) qui fournit un gain infini pour toute suite prise dans cet ensemble. Il ne put tirer de sa définition de martingale une définition des suites aléatoires, car en définissant les suites aléatoires comme les suites pour lesquelles aucune martingale ne donne un gain infini il excluait toutes les suites possibles et obtenait donc une notion vide. Cependant sa méthode était intéressante et finalement en imposant certaines conditions de calculabilité aux martingales qu'on accepte de prendre en compte, on obtient différentes notions de suites aléatoires (exprimant qu'une suite aléatoire est une suite pour laquelle aucune martingale calculable ne permet de gagner une somme infinie) dont certaines correspondent exactement à celle de Martin-Löf. Il revient à Schnorr [Schnorr 1971a] [Schnorr 1971c] d'avoir étudié et résolu le problème. Les résultats qu'il obtient, d'une part montrent que l'idée de définir les suites aléatoires en se fondant sur le principe d'une 'ingagnabilité absolue' est une bonne idée, mais font aussi apparaître des difficultés : selon la force des conditions de calculabilité qu'on impose aux martingales on obtient diverses notions de suites aléatoires, entre lesquelles le choix est délicat. Il n'en reste pas moins que l'idée d' "ingagnabilité" de von Mises était praticable et que la solution de Martin-Löf y trouve là encore un argument en sa faveur.

Il faut noter que l'imprévisibilité peut être définie relativement à des classes plus abstraites d'algorithmes, et qu'alors elle ne conduit pas forcément à une notion associée acceptable de suite aléatoire. Un exemple de ce phénomène se trouve en théorie des transformations de suites [Delahaye 1988b]. La notion naturelle d'algorithme qu'on y définit ne fait pas intervenir de notions récursivistes mais impose à une transformation F donnant (t_n) à partir de (x_n) de ne pas faire dépendre t_n d'autres données que $x_0\ x_1\ \ldots\ x_n$. Dans une telle théorie une famille de suites peut être qualifiée d'imprévisible si, bien que ne contenant que des suites convergentes, aucune transformation algorithmique (au sens considéré) ne peut accélérer la convergence de chacune des suites de la famille. L'existence de familles accélérables relativement irrégulières établit que dans un tel cas l'imprévisibilité n'équivaut à aucune "aléatoirité" naturelle.

On peut voir dans ces résultats ainsi que dans les résultats analogues de O'Connor [O'Connor 1988] où une notion trop exigente d'algorithme conduit à une notion trop faible d'aléatoirité (trop de suites peuvent être considérées aléatoires) une preuve nouvelle indirecte de la thèse de Church : seule la bonne notion d'algorithme (celle considérée

dans la thèse de Church) peut conduire à une notion intuitivement satisfaisante de suites aléatoires.

(e) L'aléatoire défini comme incompressibilité.

L'idée d'une définition des suites infinies aléatoires comme suites ayant un contenu en information incompressible est apparue assez tardivement (en 1965), simplement parce qu'on n'imaginait pas avant cette date qu'une notion absolue de contenu en information puisse être formulée.

Une fois l'échec des voies (a) et (b) reconnu, ainsi que la possibilité d'utiliser les voies (c) et (d) mises en rapport l'une avec l'autre par Schnorr, la situation serait sans doute restée confuse sans l'intervention de la théorie algorithmique de l'information. Celle-ci en proposant une notion de contenu en information non basée sur une distribution de probabilité donnée a priori —comme c'est le cas dans la théorie de l'information de Shannon— ouvrit la porte à une autre tentative de définition de la notion de suite infinie aléatoire qui après une maturation de plus de dix ans (et cinq versions différentes au moins, dont deux principales) vint effectivement faire pencher la balance en faveur de la définition de Martin-Löf un moment contestée.

Que les voies (b) (c) et (d) puissent se retrouver et coïncider était inespéré et cela constitue sans doute le meilleur argument en faveur de la thèse qu'effectivement il existe une définition mathématique de la notion de suite infinie aléatoire, et que cette définition est celle formulée en 1966 par Martin-Löf. Nous reviendrons en détail sur ces points au paragraphe 4.

2 Des collectifs de von Mises à l'accord actuel.

Après la présentation de l'analyse conceptuelle de la notion de suite infinie aléatoire nous allons parcourrir rapidement le déroulement chronologiques des travaux sur ce thème.

(a) von Mises et l'axiomatisation de Kolmogorov.

Pendant près de cinquante ans, von Mises défendit l'idée que la théorie des probabilités ne pouvait trouver un fondement rigoureux et satisfaisant que dans une théorie mathématique des séries infinies d'objets d'un espace d'épreuves (fini ou infini) vérifiant certaines conditions de limite de fréquence. Ces séries infinies d'objets appelés par lui "collectifs", sont conçues comme des «suites infinies d'expériences ou d'observations, chacune fournissant un résultat précis sous la forme d'un nombre (ou d'un groupe de nombres dans le cas d'un collectif de plus d'une dimension). ... Dans un certain sens le collectif correspond à ce qui est appelé une population en statistique» ([von Mises 1941] p. 192).

Les idées de base qui servent à la définition mathématique des collectifs ont été présentées au paragraphe (1 b3) dans le cas d'un espace d'épreuves réduit à 0 et 1, et nous avons déjà mentionné le résultat de Ville qui d'une façon définitive montre l'impossibilité d'aboutir en suivant simplement les idées de von Mises. Plutôt que de formuler à nouveau une critique des idées de von Mises nous préférons simplement donner des extraits de textes divers qui expriment très bien le problème.

> We claim that our theory, which serves to describe observable facts, satisfies all reasonable requirements of logical consistency and is free from contradictions and obscurities of any kind. [...] I would even claim that the real meaning of the

Bernoulli theorem is inaccessible to any probability theory that does not start with the frequency definition of probability. [...] All axioms of Kolmogorov can be accepted within the framework of our theory as a part of it, but in no way as a substitute for the foregoing definition of probability. R. von Mises. On the foundations of Probability and Statistics. Ann. Math. Statist. 12. 1941. pp. 191-205.

The problem of giving an adequate mathematical definition of a random sequence was subjected to an intense discussion about thirty years ago. It was initiated by von Mises as early as 1919 and reached its climax in the thirties when it engaged most of the pioneers of probability theory of that time. [...] von Mises urged that a mathematical theory of probability should be based on a definition of randomness, the probability of an event then being introduced as the limit of the relative frequency as the number of trials tends to infinity. [...] It was objected that there is just as little need for a definition of random sequences and probabilities by means of them as there is need for a definition of point and straight lines in geometry. [...] The question was not whether the theory in spe should be axiomatised or not, but what objects should be taken as primitive and what axioms should be chosen to govern them. In the axiomatization of Kolmogorov 1933 the random sequences are left outside the theory. [...] [von Mises] wanted to define random sequences in an absolute sense, sequences that were to possess all conceivable properties of stochasticity. This program appears impossible to carry out within the measure theoretic framework of Kolmogorov 1933. [...] It seems as if it were this incapability of finding an adequate mathematical definition that brought the so rapid development in the thirties to an abrupt end. [...] A common feature of the experiments considered by von Mises is that they may be repeated any, or at least a very large, number of times. For the sequence of the successive outcomes x_1, x_2, ..., x_n, ... which is imagined to extend indefinitely, von Mises coined the term "Kollectiv". A Kollectiv has to satisfy two requirements. To formulate the first of these let n_A denote the frequency with which the event A has occurred in the first n trials, i.e. the number of points x_m $1 \leq m \leq n$, that belong to the subset A of the sample space. For every "angebbare Punktmenge" A the limit of the relative frequencies should exist, $\lim n_A/n = p(A)$. This limit is called the probability of the event A with respect to the given Kollectiv. [...] The second axiom is more intricate. It is to express the wellknown irregularity of a random sequence, the impossibility of characterizing the correspondence between the number of an experiment and its outcome by a mathematical law. In a gambler's terminology it may be called the axiom of the impossibility of a successful gambling system. Thus sequences like (0 1 0 1 0 1 0 1 0 1 0 1 0 ...), 0 denoting failure and 1 success, are excluded althought the limit frequency exists, since betting at every even trial would assure us constant success. The final form of the axiom is the following. If we select a subsequence of $x_1, x_2, x_3, ..., x_n, ...$ in such a way that the decision whether x_n should be selected or not does not depend on x_n, then the limiting relative frequency of the subsequence should exist and equal that of the original sequence. [...] The definition of a Kollectiv was criticised for being mathematically imprecise or even inconsistent. [...] The trouble was due to the fact that the concept of effectiveness was not a rigorous mathematical one at that time. P. Martin-Löf. The definition of Random Sequences. Information and Control. 9. 1966. pages 602-619.

The axiomatic construction of probability theory on the basis of mesure theory [Kolmogorov 1936] as a purely mathematical discipline is logically irreproachable and does not cast doubts in anybody's mind. However to be able to apply this theory rigorously in practice its physical interpretation has to be stated clearly. Until recently there was no satisfactory solution of this problem. Indeed, probability is usually interpreted by means of the following arguments: "If we perform many tests, then the ratio of the number of favourable outcomes to the number of tests performed will always give a number close to, and in the limit exactly equal to, the

probability (or measure) of the event in question. However to say "always" here would be untrue: strictly speaking, this does not always happen, but only with probability 1 (and for finite series of tests, with probability close to 1). In this way the concept of the probability of an arbitrary event is defined through the concept of an event that has probability close to (and in the limit equal) to 1, consequently cannot be defined in this manner without an obviously circular argument. In 1919 von Mises put forward the following way of eliminating these difficulties: according to von Mises there are random and non-random sequences. From the mathematical point of view, random sequences form a set of full mesure and all without exception satisfy all the laws of probability theory. It is physically possible to assume that as a result of an experiment only random sequences appear. However, the definition of random sequences proposed by von Mises and later defined more precisely by Wald [1937], Church [1940] and Kolmogorov [1963] turned out to be unsatisfactory. For example, the existence was proved of random sequences, according to von Mises (his so-called collectives) that do not satisfy the law of iterated logarithm [Ville 1939]. A.K. Zvonkin, L.A. Levin, The Complexity of finite object and the development of the concepts of information and randomness by means of the theory of algorithms, Russ. Math. Survey, 25,6,1970, pp 83-124.

Une des idées de base de von Mises fut que l'axiomatisation des probabilités par Kolmogorov ne les fondait pas véritablement, et ne résolvait pas tous les problèmes (mathématiques en particulier) qu'on pouvait poser à la notion de probabilité. Nous pensons qu'il faut admettre qu'il avait en cela parfaitement raison car tout en étant un formalisme utile et d'une efficacité fantastique pour développer les probabilités, la théorie des probabilités ramenée à la théorie de la mesure par l'axiomatisation de Kolmogorov n'est qu'un formalisme qui contourne le problème du sens profond des probabilités. D'une certaine façon l'axiomatisation de Kolmogorov consiste simplement à déplier des éléments formels présents dans des situations probabilistes en les supposant déjà analysés. En disant que pour un lancé de dés il y a une chance sur six d'avoir un, une chance sur six d'avoir deux, etc. on ne fait que décrire des éléments formels détachés du réel de la situation qu'on n'analyse pas plus. Les théorèmes de limite centrale, loi faible et forte des grands nombres etc. qu'on retrouve dans cette axiomatisation sont de grands succès pour cette axiomatisation, mais ils nous laissent sur notre faim quant à la question de la nature véritable des probabilités.

What are the contributions of axiomatic probability to an understanding of probability? What distinguishes characterizations of uncertainty and chance from characterization of other quantities? Where do probabilities come from? In view of the manyfold and extensive applications of axiomatic probability, it is remarkable how unsatisfactory the answers to these questions are. Stated baldly, axiomatic probability hardly enables us to understand what probability is about or what its subject matter is. The Kolmogorov axioms for example, provide neither a guide to the domain of applicability of probability nor a procedure for estimating insight into the nature of random phenomena. ... The calculus as it stands, can only transform given probabilities and not create them. T. Fine, Theories of Probability, Academic Press, New-York, 1973.(page 83)

Here I have in mind that there is no need whatsoever to change the established construction of the mathematical probability theory on the basis on the general theory of measure. I am not inclined to attribute the significance of necessary foundations of probability theory to the investigations [about the algorithmic information theory] I am now going to survey. But they are most interesting in themselves. ... We have a field for very interesting mathematical research. ... 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

> description 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 the 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 algorithmic side of mathematics as a whole. A. N. Kolmogorov. Combinatorial foundations of information theory and the calculus of probabilities. Russian Mathematical Surveys. Vol. 38.4. 1983. pp. 29-40.

Au-delà de l'axiomatisation de Kolmogorov deux pas supplémentaires au moins sont envisageables dans la compréhension du hasard et des probabilités :

- le premier, est celui de la définition des suites infinies aléatoires que nous essayons de détailler ici,

- le second est celui du sens physique et philosophique du hasard, qui est beaucoup plus difficile peut-être, mais doit tenir compte, nous pensons, de ce qui est mathématiquement résolu et donc en particulier de l'axiomatisation et de la formulation des notions de suites infinies aléatoires.

(b) La première théorie algorithmique de l'information.

Pendant que certains plus ou moins empêtrés dans les notions effectives de collectifs ne faisaient que redécouvrir et préciser l'échec de von Mises, presque simultanément Solomonoff, Kolmogorov et Chaitin introduisaient une notion qui allait dans la suite jouer un rôle décisif.

L'idée de départ (définir la complexité d'un objet par la taille du plus petit programme capable de l'engendrer) semble avoir été formulée la première fois dans une communication de Minsky [Minsky 1962] (voir [Chaitin 1977a] p. 38), mais peut-être était-elle implicitement présente dans bien des définitions de la simplicité, et en tout cas est devenue naturelle quand il fut établi (par Turing [1936]) que certains mécanismes de calcul avaient un pouvoir de calcul maximum : dès que la notion de machine universelle, fut disponible l'idée de l'utiliser comme outil de mesure universel de complexité devait sans doute voir le jour, et finalement, on doit plus s'étonner du fait qu'elle fut proposée si tardivement, que de sa découverte par Minsky, Solomonoff, Chaitin et Kolmogorov.

> We feel it is important to give a careful treatment of the genesis of the ideas in this area. Kolmogorov complexity originated with the discovery of the universal description, and recursively invariant approach to the concepts of complexity of description, randomness, and a priori probability. Historically, it is firmly rooted in von Mises' notion of random infinite sequences ... With the advent of electronic computer in the 1950's, a new emphasis on computer algorithms and a maturing general recursive function theory, ideas tantamount to Kolmogorov complexity came to many people's mind, because "when the time is ripe for certain things, these things appear in different places in the manner of violet coming to light in early spring" (Wolfgan Bolyai to his son Johann in urging him to claim the invention of non-Enclidean geometry without delay). Thus, with Kolmogorov complexity one can associate three inventors : R. J. Solomonoff in Cambridge, Massachusetts, close in time but far away in geography followed by A.N. Kolmogorov in Moscow, and once more by G. J. Chaitin in New York. R. J. Solomonoff had been a student of R. Carnap at the University of Chicago in the fifties. His objective was to formulate a completely general theory of inductive inference that would overcome shortcomings of previous methods like [R. Carnap. Logical Foundation of Probability. University of Chicago Press, Chicago, Ill. 1950.] Already in November 1960 Solomonoff had published a Zator Company technical report on the sublect of "Kolmogorov" complexity ([Solomonoff 1960]).

In March 1964 he published a long paper ([Solomonoff 1964] introducing an imprecise version of universal *a priori* probability, namely the forerunner of the Solomonoff-Levin distribution, through intermediate definition of what we have termed "Kolmogorov complexity", and proved the invariance Theorem. This paper received little attention until Kolmogorov started to refer to it from 1968 onward. It is interesting to note that Solomonoff also discusses informally the ideas of randomness of finite strings, noncomputability, and resource bounded Kolmogorov complexity. A paragraph refering to Solomonoff's work occurs in [Minski 1961]. To our knowledge, these are evidently the earliest documents outlining an algorithmic theory of descriptions. In 1933 the great Soviet Mathematician A. N. Kolmogorov supplied probability theory with a powerful mathematical foundation [Kolmogorov 1933]. Following a four decades long controversy on von Mises' concept of randomness, Kolmogorov finally introduced complexity of description of finite individual objects as a mesure of individual information content and randomness, and proved the invariance Theorem in his paper of spring 1965 [Kolmogorov 1965]. Kolmogorov invention of complexity of description was in no way a haphazard occurence, but on the contrary the invitable confluence of several of his major research threads : the foundation of probability and random sequences, information theory, and the theory of algorithms. Uspekhi Mat Nauk announced Kolmogorov's lectures on related subjects in 1961 and following years, and says Kolmogorov : "I came to similar conclusion [as Solomonoff] before, becoming aware of Solomonoff's work, in 1963-1964." [Kolmogorov 1968]. G. J. Chaitin had finished the Bronx High School of Science, and was an 18 year old undergraduate student at the City College of the City University of New York, when he submitted the original versions of [Chaitin 1966] [Chaitin 1969] for publication, in October and November 1965 respectively. Published in 1966, [Chaitin 1966] investigated "state/symbol" complexities relative to arbitrary algorithms. In this work Chaitin extended C.E. Shannon's earlier work on coding concepts [Shannon 1956], and did not introduce any invariant notion of complexity. However, at the end of his 1969 publication [Chaitin 1969], Chaitin apparently independently puts forward the proper notion of Kolmogorov complexity, prove the Invariance Theorem, and studies infinite random binary sequences (in the sense of having maximally random finite initial segments) and their complexity oscillations. According to Chaitin: "this definition [of Kolmogorov complexity] was independently proposed about 1965 by A. N. Kolmogorov and me ... Both Kolmogorov and I were unaware of related proposals made in 1960 by Ray Solomonoff" [Chaitin 1975b]. The Swedish mathematician P. Martin-Löf, visiting Kolmogorov in Moscow during 1964-1965, investigated complexity oscillations of infinite sequence and proposed a new definition of infinite random sequences which is based on constructive measure theory [Martin-Löf 1966] [Martin-Löf 1972]. L. A. Levin, then a student of Kolmogorov, found a proper definition of a priory probability (the Solomonoff-Levin Distribution) as a maximal semicomputable measure [Zvonkin Levin 1970], and introduced the quantity corresponding to self-delimiting variant of Kolmogorov complexity as the negative logarithm of a priori probability. In 1974 he more explicitly introduced Kolmogorov complexity based on self-delimiting program [Levin 1974]. This work relates to P. Gacs' result concerning the differences between symmetric and asymmetric expressions for information [Gacs 1974]. In 1975 Chaitin also discovered and investigated this Kolmogorov complexity based on self-delimiting programs [Chaitin 1975]. M. Li, P. M. B. Vitanyi. Kolmogorov Complexity and Its Applications. Handbook of Theoretical Computer Science. J. van Leeuwen Editor. North-Holland. 1989.

On doit donc attribuer à Solomonoff la première formulation technique (mentionnant la notion de machine de Turing universelle) de la théorie algorithmique de l'information. Les idées de Solomonoff furent exposées pour la première fois en 1962 par Minsky (voir [Minsky 1962] [Solomonoff 1964] [Chaitin 1977]). Les principes de base de la théorie furent retrouvés et développés indépendamment dans les travaux de Kolmogorov et

Martin-Löf ([Kolmogorov 1963, 1965 et 1968] [Martin-Löf 1966, 1969, 1970]), et de Chaitin ([Chaitin 1966,1969 et 1970]).

Conformément à un usage assez général nous appellerons cette première théorie de la complexité : **théorie de la complexité de Kolmogorov-Chaitin** oubliant ainsi celui qui en eut le premier l'idée mais eut tort sans doute de ne pas mener assez loin les développements mathématiques.

La notion de **complexité de Kolmogorov-Chaitin** se définit à l'aide de la notion de contenu en information des objets finis (pour nous des suites finies de 0 et de 1) notion elle-même formulée à l'aide de la notion de machine de Turing universelle. Une machine de Turing universelle est une machine de Turing capable de simuler toutes les machines de Turing [Turing 1936]. Le contenu en information de la suite finie s ou complexité algorithmique de Kolmogorov-Chaitin de s, noté K(s), est par définition la taille du plus petit programme (pour une machine de Turing universelle) capable de produire s. On montre que cette notion est indépendante, à une constante additive près, de la machine de Turing universelle utilisée dans la définition. Autrement dit si $K_U(s)$ désigne la complexité algorithmique de s obtenue en utilisant la machine de Turing universelle U et $K_V(s)$ celle obtenue avec V, il existe une constante c (ne dépendant que de U et V) telle que pour toute suite finie s : $|K_U(s)-K_V(s)| < c$. Ce résultat est essentiel puisqu'il justifie l'idée de prendre une machine universelle particulière. On l'appelle **Théorème d'Invariance**. Il fut énoncé pour la première fois par Solomonoff, puis indépendamment par Chaitin et Kolmogorov. On montre immédiatement qu'il existe une constante c' telle que pour toute suite finie s : K(s) < longueur(s) + c'. L'idée de la démonstration de ce résultat est très simple : pour toute suite s, il y a un programme produisant s dont le sens est simplement "copier s", et donc qui a pour longueur la longueur de s plus ce qu'il faut pour exprimer l'algorithme de copie, qui a une longueur c'.

Il faut bien noter car c'est très important pour la suite que ceci n'est possible que parce que la fin d'un programme n'a pas besoin d'être indiquée dans le programme lui-même. La fin du programme est atteinte quand on a lu tous les digits du programme. Dans la deuxième théorie de la complexité algorithmique (que nous appellerons **théorie de Chaitin-Levin**) on exigera que la fin des programmes soit précisée dans les programmes eux-mêmes (on parlera de programmes autodélimités) comme c'est le cas pour pratiquement tous les langages de programmation informatique actuels. Les conséquences de cette exigence seront (i) qu'aucun programme ne pourra être prolongé en un autre (et donc qu'on pourra attribuer un poids naturel à chaque programme), (ii) que l'inégalité précédente deviendra H(s) < longueur(s) + 2log(longueur(s)) + c' (le programme de copie d'une suite s devra contenir l'information de la longueur de s en plus simplement de la description digit par digit de s).

> On the basis of these considerations it may perhaps not appear entirely arbitrary to define a patternless or random finite binary sequence as a sequence which in order to be calculated requires, roughty speaking, at least as long a program as any other binary sequence of the same length. A patternless or random infinite binary sequence is then defined to be one whose initial segments are all random. G. J. Chaitin. On the Length of Programs for Computing Finite Binary Sequences. J. A.C.M. 13. 1966. pp.547-569. Repris dans [Chaitin 1987a].page 236.

La définition du contenu en information ou complexité telle qu'elle est formulée par Chaitin 1966, lui suggère une définition des suites infinies aléatoires comme suites dont le contenu en information est le plus grand possible à un logarithme près :

> The set C_∞ of patternless or random, infinite binary sequences is formally defined to be the set of all binary sequences S which satisfy the following inequality for all sufficiently large values of n : $L(S_n) > L(C_n) - f(n)$ where $f(n) = 3 \log_2 n$ and S_n is the sequence of the first n bits of S. [$L(C_n)$ désigne la complexité maximum d'une

suite de longueur n] This definition, unlike the first, is quite precise but also somewhat arbitrary. The failure to state the exact cut-off point at which L(S) become to small for S to be considered random or paternless gives to the first definition its informal character. But in the case of finite binary sequences, no gain in clarity is achieced by arbitrary settling on a cut-off point, while the opposite is true for infinite sequences. G. J. Chaitin. On the Length of Programs for Computing Finite Binary Sequences. Statistical Considerations. J.A.C.M. 16. 1969. pp.145-159. Repris dans [Chaitin 1987a]. page 240.

Dans cet article de 1969 (soumis en 1965) Chaitin démontre que sa définition de suite aléatoire implique la propriété de normalité quelle que soit la base considérée et qu'elle entraîne aussi les propriétés de limite de fréquence pour elle-même et ses suites extraites par procédés effectifs (il montre donc que ses suites infinies aléatoires sont des collectifs dans le sens de von Mises rendu précis par Church en 1940). Il montre enfin une certaine propriété d'imprédictibilité (page 242). Mais cette tentative ainsi que d'autres basées sur la première théorie algorithmique de l'information se révèla insatisfaisante. A cette époque donc aucune thèse définitive ne put donc être formulée sur la base de la nouvelle notion d'information.

Pour terminer ce paragraphe notons l'ironie de l'histoire : von Mises jugeait la théorie des probabilités de Kolmogorov insatisfaisante pour fonder les probabilités et était le promoteur d'une théorie (mathématiquement mal formalisée) contre Kolmogorov. Kolmogorov fournit lui-même les éléments de base pour fonder proprement la théorie que von Mises appelait de ses vœux. Mais von Mises même rétrospectivement n'a pas encore raison, car ce qui aurait pu être sa théorie n'invalide pas du tout la théorie probabiliste fondée sur la notion de mesure qui garde une puissance et a jusqu'à présent une fécondité plus grande que celle calculatoire. On peut donc dire que von Mises a perdu trois fois contre Kolmogorov : une première fois en défendant une théorie à laquelle il ne réussit pas à donner de la consistance; une seconde fois en laissant Kolmogorov et Martin-Löf formuler la théorie que lui aurait voulu formuler; une troisième fois en n'ayant pas vraiment raison contre la théorie probabiliste fondée sur la notion de mesure.

(c) La définition de Martin-Löf.

Alors que la théorie de la complexité algorithmique de Kolmogorov-Chaitin n'en n'était qu'à ses débuts, un travail étonnant fut réalisé par un jeune chercheur suédois qui avait fait un séjour auprès de Kolmogorov. Dans son article, il s'intéressait au problème de définir une notion d'objet aléatoire fini et de suite infinie aléatoire.

Détaillons la définition de Martin-Löf de suite infinie aléatoire.

On appelle intervalle binaire un intervalle de la forme [a, b] où a et b sont des nombres de la forme $p/2^q$. Un **test statistique effectif d'aléatoirité** est la donnée d'un ensemble récursivement énumérable d'intervalles I(m,n) tels que pour tout m fixé la longueur de tous les intervalles binaires I(m,n) est plus petite que $1/2^m$. La condition "plus petite que $1/2^m$" peut être remplacée par d'autres conditions, pour peu qu'elles soient effectives (par exemple "plus petite que 1/f(n) où f est une fonction récursive totale") Appliquer un test consiste à choisir un niveau m, puis à étudier si x est dans l'un des intervalles I(m,n) quand n varie. Si x est dans un tel intervalle x est éliminé, x n'est pas considérée comme aléatoire pour ce niveau de test. Les suites qui ne sont plus éliminées à partir d'un certain rang m sont, par définition, les suites qui satisfont le test. Par définition **une suite aléatoire au sens de Martin-Löf** est une suite qui réussit tous les tests statistiques effectifs d'aléatoirité.

Par hypothèse pour chaque test statistique effectif d'aléatoirité, seul un ensemble de mesure nulle de suites ne satisfait pas le test (c'est la condition sur les longueurs des

intervalles qui assure cela). En fait chaque test élimine une certaine propriété de régularité qu'on considère n'être pas compatible avec la propriété d'être aléatoire.

Tout test statistique d'aléatoirité un tant soit peu effectif et raisonnable (n'éliminant pas plus qu'un ensemble de mesure nulle de suites) est donc par hypothèse satisfait par les suites aléatoires au sens de Martin-Löf. Cette généralité de base de la définition de Martin-Löf est ce qui fait sa force, et est sans doute ce qui explique qu'elle finit par triompher des autres définitions, même si certaines furent un moment considérées comme aussi naturelles.

On peut construire un test statistique effectif par exemple pour la propriété "ne pas satisfaire la loi forte des grands nombres", donc, les suites aléatoires au sens de Martin-Löf satisfont toutes la loi forte des grands nombres. Pour la même raison les suites aléatoires au sens de Martin-Löf satisfont toutes à la loi du logarithme itéré. On comprend bien la force de la définition de Martin-Löf et son aspect naturel.

Une chose étonnante dans cette approche est qu'on peut établir l'existence d'un test maximum résumant tous les autres et qui, s'il est satisfait, assure à lui seul qu'une suite est aléatoire. Ce **test universel** est l'analogue de la machine universelle de Turing (qui est d'ailleurs utilisée dans la démonstration de l'existence de ce test universel)

Une question naturelle est de savoir si ce test universel est utilisable en pratique ? La réponse est sans doute non (bien qu'à ma connaissance nul n'ait essayé sérieusement de savoir ce qu'il pouvait donner) car si la notion de test que Martin-Löf s'est donné est effective d'une certaine façon, la plupart des tests vraiment utilisés le sont "encore plus". Par exemple les intervalles I(n,m) sont récursifs et pas seulement récursivement énumérables, et mieux encore pour chaque entier m les intervalles I(n,m) ne sont qu'en nombre fini ce qui permet vraiment d'envisager plusieurs niveaux de test. Or justement le test maximum lui n'a pas ces propriétés d'effectivité forte et donc ne peut être sérieusement envisagé en pratique. Remarquons aussi qu'un test peut éliminer à tort certaines suites aléatoires et ne cesse de les éliminer qu'à partir d'un certain niveau, le test universel sans doute repousse ce niveau assez loin et n'est donc pas intéressant : sa généralité interdit la rapidité de son effectivité. C. P. Schnorr a d'ailleurs démontré que tout test statistique universel avait un comportement qui d'une certaine façon était non calculable (Théorème 7 p. 385 [Schnorr 1973])

On peut se demander ce que vient faire la condition d'effectivité dans la définition de Martin-Löf. Que se passe-t-il si on la supprime ? Que se passe-t-il si on la remplace par une condition plus faible ou plus forte ? Si on la remplace par une condition plus forte on risque fort de perdre certaines propriétés attendues des suites infinies aléatoires (comme le fait qu'elles satisfont la loi forte des grands nombres). Il ne semble pas qu'on puisse raisonnablement renforcer le "récursivement énumérable" de la définition. Si on la supprime alors il n'y a plus de suites aléatoires (car l'ensemble des suites aléatoires devient la réunion de tous les ensembles de mesure nulle, c'est-à-dire tout). D'une manière générale on risque de ne plus avoir aucune suite aléatoire si on remplace le "récursivement énumérable" par une condition plus faible. Cependant l'existence de suites aléatoires provient essentiellement de la dénombrabilité des tests statistiques envisagés (et précisément du fait qu'une réunion dénombrable d'ensembles de mesure nulle l'est encore) on peut donc très bien remplacer le "récursivement énumérable" de la définition par une autre condition pourvu qu'elle garantisse que le nombre de tests acceptables reste dénombrable. Par exemple en remplaçant "récursivement énumérable" par "de la hiérarchie arithmétique" on définit une notion de suite aléatoire non vide. Finalement de la même façon qu'on peut imaginer des classes d'effectivité plus grandes (mais encore dénombrables) que la classe des fonctions récursives partielles correspondant à des capacités calculatoires d'êtres nous dépassant mais d'une manière mathématiquement raisonnable et concevable, on peut imaginer des notions plus fortes de suites aléatoires pour ces mêmes êtres au-delà du fini auquel nous pensons être confinés.

Concernant le problème de la définition de Martin-Löf on peut faire la remarque suivante. Lui-même insiste sur le fait que finalement, ce que nous souhaitons pour nos suites aléatoires, c'est qu'elles vérifient toutes les propriétés dont nous réussissons à montrer (par démonstrations véritablement produites) qu'elles sont satisfaites partout sauf sur un ensemble de mesure nulle (de l'intervalle [0,1]). C'est très naturel et sans doute même, cela l'est plus que ce que sa définition énonce. Est-il vrai alors que toute propriété (définissable par une formule précise, dans un système formel correct pour l'analyse) que l'on peut démontrer être vérifiée (dans ce système) sauf sur un ensemble de mesure nulle, est satisfaite par les suites aléatoires au sens de Martin-Löf ? Ce résultat métamathématique s'il pouvait être établi serait un nouvel argument extrêment fort en faveur de la définition de Martin-Löf.

(d) La seconde théorie algorithmique de l'information.

La première théorie algorithmique de l'information ne se révéla pas satisfaisante en particulier à cause de facteurs logarithmiques là où on aurait aimé ne rien avoir. La **deuxième théorie algorithmique de l'information** est basée sur les mêmes idées que la première : la complexité d'un objet, c'est la taille du plus petit programme permettant de le définir. Elle diffère sur un point très simple : les algorithmes qu'on considère doivent être **auto-délimités**, c'est-à-dire contenir en eux-même l'indication de leur fin.

Cette modification technique dans la définition de la complexité algorithmique apporte réellement une simplification. Voir [Chaitin 1975a] [Chaitin 1987b] [Delahaye 1989] [Gacs 1974] [Gacs 1983] [Kolmogorov Uspenskii 1987] [Levin 1973] [Levin 1976c] [Li Vitanyi 1989]. En particulier l'opposé du logarithme de la probabilité pour qu'une machine universelle à programmes autodélimités produise un objet donné est égal à la complexité de cet objet : $H(s) = - \log (P(s)) + O(1)$.

Surtout, la nouvelle théorie conduit à la formulation de ce que nous avons appelé la **thèse de Chaitin-Levin** (voir paragraphe 1) dont on démontre qu'elle coïncide avec la **thèse de Martin-Löf**. Cette convergence de deux définitions construites sur des principes totalement différents est l'argument majeur en faveur de la thèse de Martin-Löf.

3 La situation actuelle.

On peut résumer la situation actuelle de la manière suivante :

Nombre aléatoire: *Les quatre définitions suivantes sont équivalentes* ([Chatin 1987b])

le nombre réel x *de l'intervalle* [0,1] *(assimilé à la suite binaire des digits de son développement binaire propre) est* ***aléatoire*** *si et seulement si :*

Définition 1 (*aléatoire au sens de* **Martin-Löf**, 1966)

> *Pour toute suite récursivement énumérable infinie* A_i *d'ensembles d'intervalles, chaque* A_i *de mesure majorée par* 2^{-i} $(\mu(A_i) < 2^{-i})$: x *n'appartient pas à au moins un des* A_i.

Définition 2 (*aléatoire au sens de* **Solovay**, 1975)

Pour toute suite récursivement énumérable infinie d'ensembles d'intervalles A_i *dont la somme des mesures est bornée* $(\sum\mu(A_i) < \infty)$: x *est au plus dans un nombre fini de* A_i.

Définition 3 (*aléatoire au sens de* **Chaitin-Levin**)

La complexité $H(x^n)$ *des* n *premiers digits de* x *vérifie :*

$$\exists\, c\ \forall n : H(x^n) \geq n - c$$

Définition 4 (*fortement aléatoire au sens de* **Chaitin** 1987)

La complexité $H(x^n)$ *des* n *premiers digits de* x *vérifie :*

$$\forall k\ \exists\, N\ \forall n \geq N :\ H(x^n) \geq n + k$$

Commentaires.

Une **régularité statistique exceptionnelle effectivement testable** (et on veut que les suites aléatoires n'en vérifient aucune!) pour une suite de 0 et de 1 est une propriété qu'on peut vérifier effectivement avec une précision plus ou moins grande (c'est-à-dire avec un niveau d'exigence plus ou moins grand) et qui n'est vérifiée que par un ensemble de mesure nulle de suites infinies. Dès qu'une suite vérifie une telle régularité à tous les niveaux, elle ne peut plus être considérée comme aléatoire. Les suites aléatoires au sens de Martin-Löf sont celles qui pour toute régularité exceptionnelle effectivement testable ne satisfont pas le test au moins à un certain niveau d'exigence. A chaque régularité statistique exceptionnelle correspond ce que nous avons appelé plus haut un test statistique effectif d'aléatoirité.

On définit une régularité statistique exceptionnelle effectivement testable par une suite A_i récursivement énumérable de réunions d'intervalles, chaque A_i ayant une longueur inférieure à 2^{-i}. Cette définition de régularité statistique peut sembler assez arbitraire, mais on peut l'exprimer en faisant intervenir uniquement les digits des suites qu'on considère (voir [Chaitin 1987b]). Pour bien comprendre en quoi elle est satisfaisante donnons un exemple.

Considérons la propriété statistique exceptionnelle effectivement testable "pour tout i les digits 2i et 2i+1 sont toujours 01 ou 10". Les suites vérifiant cette propriété sont nombreuses (elles forment un ensemble ayant la puissance du continu, mais de mesure nulle) et clairement, au sens intuitif, elles ne sont pas aléatoires (à cause du fait que le digit de rang 2N+1 est toujours différent de celui de rang 2N). Montrons que selon le critère de Martin-Löf elles ne sont pas aléatoires. On prend pour A_0 l'ensemble des nombres dont le développement binaire propre commence par 01 ou par 10, cet ensemble a exactement comme mesure 1/2. On prend pour A_1 l'ensemble des nombres dont le développement binaire propre commence par 0101 ou 0110 ou 1010 ou 1001, cet ensemble a pour mesure 1/4 (4 fois 1/16). ... On prend pour A_i l'ensemble des nombres dont le développement binaire propre commence par n doublets $d_1\ d_2\ ...d_n$ chacun égal soit à 01 soit à 10, cet ensemble a pour mesure $1/2^n$ (2^n fois $1/4^n$). La propriété statistique exceptionnelle envisagée correspond au test statistique d'aléatoirité défini en prenant les intervalles composant les A_i.

Autres exemples de propriétés statistiques exceptionnelles effectivement testables (et que les suites aléatoires au sens de Martin-Löf ne vérifient donc pas) : (i) Avoir toujours

plus de zéro que de 1 dans les n premiers digits. (ii) Avoir toujours la différence (nombre de 1 - nombre de 0) inférieure à m (fixé). (iii) Ne pas vérifier la loi du logarithme itéré. [Lim sup (resp. inf) $(2s_n-n) / (2n \log \log n)^{1/2} = +1$ (resp -1) où s_n désigne le nombre de 1 parmi les n premiers digits]

Formulée d'une manière moins technique que précédemment l'analyse actuelle donne donc que :

Etre aléatoire c'est vérifier l'une des propriétés équivalentes suivantes :

- ne satisfaire aucune régularité exceptionnelle effectivement testable (i.e. passer tous les tests statistiques effectifs d'aléatoirité [Martin-Löf 1966]),

- posséder un contenu en information incompressible (i.e. une complexité algorithmique maximum [Levin 1974] [Chaitin 1975a]),

- être imprévisible ou "ingagnable" (qui parie en utilisant une stratégie effective sur les digits d'un nombre aléatoire ne peut pas avoir de gain infini [Schnorr 1971a] [Schnorr 1977]).

Etre aléatoire implique :

- ne possèder aucune forme algorithmique (avoir une forme algorithmique signifiant pouvoir être décrit par un moyen fini effectif, comme les digits de Pi qui n'est donc pas aléatoire);

- satisfaire la propriété de convergence des fréquences pour toute suite extraite de manière effective (propriété de von Mises-Church);

- être normale (dans toutes base);

- être sans post-effet [Popper 1935] ("free from aftereffect") forme faible de la propriété de von Mises-Church : la limite des fréquences de succès ne change pas pour les suites extraites par le procédé : une suite de 0 et de 1 étant fixée $x_1, x_2, ..., x_n$ extraire la sous-suite en prenant ce qui vient après chaque occurrence de $x_1, x_2, ..., x_n$.

4 Comparaison entre la thèse de Martin-Löf et la thèse de Church.

Il semble intéressant de comparer les situations respectives de la thèse de Church (à propos des fonctions calculables par algorithmes) et de la thèse Martin-Löf (à propos des suites aléatoires infinies de zéro et de un) .

D'abord précisons que nous ne voulons parler que de "la thèse de Church standard" qui identifie le concept métamathématique (nous dirons aussi intuitif ou philosophique) de 'fonction définissable par algorithme' avec le concept mathématique de 'fonction définissable par machine de Turing' (c'est-à-dire de 'fonction récursive partielle'). Les variantes de la thèse de Church mettant en cause les concepts de 'fonctions calculables par une machine du monde physique', ou le concept de 'fonction calculable par un être vivant ou par esprit doué d'intelligence' ne nous concernent pas ici. Les arguments qu'on a parfois avancés contre la thèse de Church en fait s'adressent à la thèse de Church physique ou biologique, mais concernent rarement la thèse de Church standard. Cette dernière est en effet maintenant universellement admise, et mieux, utilisée quotidiennement en théorie de la récursivité.

Nous insistons sur le fait que certaines incompréhensions à propos de la thèse de Martin-Löf proviennent d'une confusion analogue entre hasard physique et hasard au sens métamathématique.

Dans chacune des deux thèses, il y a une proposition d'identification entre un concept métamathématique (intuitif, philosophique), et un concept mathématique. Le concept métamathématique visé par la définition mathématique n'est peut-être pas très précis au départ, et peut-être même n'acquiert-il un sens définitif que grâce à la thèse (de Church ou de Martin-Löf) cependant dans la mesure où ce concept préexistait, et où de toute évidence les thèses (de Church ou de Martin-Löf) étaient falsifiables (et le restent) il serait inexact de les considérer seulement comme des définitions.

La thèse de Church comme la thèse de Martin-Löf sont donc bien de la même nature: on ne peut espérer les démontrer, on peut seulement soit les falsifier, soit argumenter pour les défendre ou s'y opposer.

Les arguments utilisés pour les défendre ou s'y opposer peuvent être classés en différentes catégories que nous allons examiner les unes après les autres, en faisant à chaque fois apparaître ce qui diffère d'une thèse à l'autre.

Puisque nous ne nous intéressons qu'aux versions standard de ces thèses les problèmes du mécanisme, de l'identité corps-esprit, du 'hasard vrai' en mécanique quantique ou en physique classique ne sont pas concernés, pas plus d'ailleurs que les problèmes que posent ces thèses dans le cadre d'une philosophie constructiviste des mathématiques.

(a) Arguments par les exemples.

(a1) Défense de la thèse de Church.

Les fonctions usuelles comme les fonctions `n -> 2n; n -> n!; n -> [n-ième nombre premier]` et bien d'autres, sont des fonctions clairement calculables par algorithme, on montre facilement qu'elles sont récursives, donc la thèse de Church n'est pas trop restrictive.

(a2) Défense de la thèse de Martin-Löf.

Les suites (0 0 0 0 ...) (0 1 0 1 0 1 0 1 ...) et bien d'autres, sont des suites non-aléatoires au sens intuitif, et elles sont aussi non ML-aléatoires, la thèse de Martin-Löf n'est donc pas trop tolérante.

(a3) Comparaison et remarques.

Les arguments par les exemples dans le premier cas donnent la conclusion que la thèse n'est pas trop restrictive, dans l'autre qu'elle n'est pas trop tolérante. A un passage au complémentaire près (par exemple: considérer que la thèse de Martin-Löf définit la notion de suite infinie non-aléatoire) cela revient exactement au même. Ces arguments peuvent être vus comme des arguments d'**adéquation minimum** des définitions mathématiques utilisées dans les deux thèses.

Les exemples utilisés dans le cas de la thèse de Church sont parfois proposés par familles entières (par exemple toutes les fonctions polynômes) mais toujours par familles dénombrables, et il ne peut en être autrement puisque l'ensemble des fonctions récursives est dénombrable. Par contre pour la thèse de Martin-Löf des familles non dénombrables d'exemples (de suites non-aléatoires) peuvent être proposées, comme par exemple la famille des suites telles que $x_{2n}=x_{2n+1}$ pour tout n. Sur ce point donc la thèse de Martin-Löf possède donc un avantage : elle est soutenue par plus d'exemples.

Ces arguments d'adéquation minimum sont insuffisants bien sûr, mais ils sont quand même très importants et loin d'être évidents pour certains d'entre eux. Par exemple il a fallu du temps pour comprendre que la thèse de von Mises-Church n'entraînait pas que les suites aléatoires (selon cette thèse) vérifiaient la loi du logarithme itéré.

(b) Arguments par les contre-exemples.

(b1) Défense de la thèse de Church.

Par diagonalisation on arrive à construire des fonctions qui ne sont pas récursives, cela montre que la thèse de Church n'est pas trop tolérante, mais comme de plus on ne trouve pour ces fonctions contre-exemples aucun algorithme qui les calcule, tout est bien cohérent.

(b2) Défense de la thèse de Martin-Löf.

Par des arguments non-constructifs de théorie de la mesure, mais aussi par des arguments plus directs (définition du nombre oméga [Bennett 1979] [Chaitin 87b] [Gardner 1980] [Delahaye 1988a 1989b]) on réussit à établir qu'il existe des suites aléatoires au sens de Martin-Löf. La thèse de Martin-Löf n'est donc pas trop restrictive. Ces suites aléatoires autant qu'on peut les étudier semblent bien ne jamais posséder de propriété contraire à l'idée intuitive qu'on peut avoir. Là-aussi, tout est cohérent.

(b3) Comparaison et remarques.

A un passage au complémentaire près, comme précédemment, les situations paraissent équivalentes. Mais en fait, il est beaucoup plus facile de construire des fonctions non-récursives que des suites infinies aléatoires au sens de Martin-Löf. Et même si à chaque fois on doit faire appel à des méthodes non effectives (diagonalisation) il semble bien qu'on puisse dire que la 'non-effectivité' des suites aléatoires est plus grave que celle des fonctions non récursives dont certaines sont presque concrètes ! Cette non-effectivité peut être mesurée grâce à la hiérarchie arithmétique, ce qui confirme le sentiment que les arguments par les contre-exemples sont moins convaincants pour la thèse de Martin-Löf que pour la thèse de Church.

(c) Arguments d'analyse conceptuelle.

(c1) Défense de la thèse de Church.

Les multiples façons de définir la notion mathématique de fonction récursive sont basées :

- soit sur la modélisation mathématique de mécanismes physiques (exemples : les machines de Turing, les automates à registres)

- soit sur la formalisation d'opérations intuitivement considérées comme calculables (par exemple définition à partir de fonctions élémentaires et clôture par compositions, schémas récursifs, et minimisations bornées)

- soit sur la définition de langages de programmation.

Chacune de ces définitions se fonde sur une analyse conceptuelle, chacune de ces définitions peut d'ailleurs être vue comme une tentative directe de formaliser le concept intuitif de fonctions calculables par algorithme. Et donc, que toutes correspondent permet de cumuler en faveur de la thèse de Church tous les arguments qu'on fait valoir pour chacun des modèles de calcul représentés par ces définitions. C'est ce faisceau

d'arguments qui constitue la meilleure argumentation en faveur de la thèse de Church, surtout que des dizaines et des dizaines de modèles ont été proposés.

(c2) Défense de la thèse de Martin-Löf.

Les définitions des suites aléatoires de Martin-Löf par imprévisibilité absolue, par satisfaction à tous les tests effectifs, par incompressibilité du contenu en information, fournissent chacune des arguments d'analyse conceptuelle convaincants (et convergents). Là encore c'est sans doute les plus forts de tous les types d'arguments que nous allons envisager.

(c3) Comparaison et remarques.

Si aujourd'hui les arguments d'analyse conceptuelle en faveur de la thèse de Church semblent extrêmement puissants, on ne peut s'empêcher de remarquer qu'en 1934-36 ils ne l'étaient pas tant que cela puisque Gödel, par exemple, ne les jugeaient pas suffisants (c'est seulement à la suite de l'article de 1936 de Turing qu'il accepta la thèse de Church). De toutes façons, si ces arguments indiquent incontestablement qu'on a mis dans la définition mathématique beaucoup de choses qui doivent y être, ils n'indiquent jamais qu'on les a toutes mises. Il reste toujours envisageable que quelqu'un vienne proposer un exemple de fonction qu'on reconnaîtra calculable au sens intuitif et qui ne le sera pas par machine de Turing.

Concernant la thèse de Martin-Löf, on a encore plus de raisons d'être prudent, puisqu'on a cru un certain nombre de fois détenir des définitions satisfaisantes et naturelles des suites infinies aléatoires qui, plus tard ont dû être revues. Il y a des arguments d'analyse conceptuelle en faveur de la thèse de Martin-Löf, mais il y en avait aussi en faveur de la thèse de Popper, ou de la thèse de von Mises-Church, et même si aujourd'hui ils semblent plus nombreux et mieux organisés, on doit garder à l'esprit que l'intuition qu'on a eu, s'est révélée plusieurs fois insuffisante et insatisfaisante pour le problème de suites infinies aléatoires.

La phase de propositions erronées ne s'est pas produite pour la notion de fonction calculable par algorithme, et les choses semblent moins délicates à son sujet. Là encore donc la thèse de Church est mieux étayée que la thèse de Martin-Löf.

(d) Argument de la confluence des définitions.

(d1) Défense de la thèse de Church.

Chacune des définitions données de la famille des fonctions récursives partielles renforce l'idée que le concept intuitif d'algorithme est bien représenté par le concept mathématique. Le fait que des dizaines de définitions mathématiques différentes donnent le même concept mathématique (autrement dit les démonstrations d'équivalence entre toutes ces définitions mathématiques) est tout à fait étonnant et constitue un argument supplémentaire extrêmement puissant en faveur de la thèse de Church. En effet on pouvait penser a priori qu'il y avait plusieurs concepts mathématiques différents chacun pouvant peut-être donner une traduction mathématique du concept, au départ vague, de fonction calculable par algorithme. Découvrir que les différentes formalisations naturelles fournissent le même concept mathématique conduit à penser qu'on tient véritablement la bonne notion : si l'une des définitions avait laissé échapper une idée (i.e. une classe de fonctions intuitivement calculables par algorithme) il est peu vraisemblable qu'une autre fondée sur une méthode de formalisation totalement différente ait laissé échapper exactement les mêmes fonctions. Au bout du compte il est inimaginable que les dizaines de définitions proposées se trompent toutes, et se trompent toutes de la même façon.

La confluence des définitions est véritablement un argument de premier ordre en faveur de la thèse de Church. C'est aussi un argument qui plaît bien aux mathématiciens car il nécessite des démonstrations parfois difficiles et correspond donc à un authentique "travail mathématique".

(d2) Défense de la thèse de Martin-Löf.

Pour la thèse de Martin-Löf il y a aussi confluence des définitions. Des définitions par imprévisibilité, par tests statistiques effectifs et par incompressibilité coïncident et c'est cette coïncidence qu'invoque Kolmogorov et Uspenskii [Kolmogorov Uspenskii 1987] pour conclure en faveur de la thèse de Martin-Löf. Il y a cependant beaucoup moins de définitions confluentes pour la notion de suite infinie aléatoire que pour celle de fonction calculable par algorithme. Et si on peut espérer que d'autres définitions voient le jour dans un proche avenir, aujourd'hui l'argument de confluence est moins fort pour la thèse de Martin-Löf que pour la thèse de Church.

(d3) Comparaison et remarques.

Il ne faut pas croire malgré tout que la confluence seule pourrait établir les thèses que nous étudions. Il y a en effet des dizaines de façons de définir la classe des fonctions récursives primitives, ou la classe des suites infinies récursives alors qu'aucune de ces deux notions ne peut être considérée comme une formalisation satisfaisante des notions intuitives visées. Les arguments de confluence n'ont véritablement de force qu'associés aux arguments d'analyse conceptuelle s'appuyant sur l'aspect naturel des définitions mathématiques, et donc là encore la thèse de Church est mieux étayée que celle de Martin-Löf.

(e) Arguments de robustesse.

(e1) Défense de la thèse de Church.

Le fait que la classe des fonctions récursives partielles ne se laisse pas diagonaliser est ce qui convainquit Kleene en 1934 que la thèse que Church venait juste de formuler était correcte. Cette robustesse de la classe des fonctions récursives partielles se manifeste de bien d'autres façons encore. Par exemple toutes les variantes de la notion de machine de Turing (machine de Turing à plusieurs rubans, machine de Turing sur un alphabet étendu, etc), conduisent toujours à la classe des fonctions récursives partielles. Cette insensibilité de la notion de fonctions récursives partielles aux variations des éléments qui entrent dans ses diverses définitions en démontre l'aspect 'intrinsèque'; pour sortir de cette classe de fonctions (contrairement par exemple à la classe des fonctions récursives primitives), il faut y aller fort.

(e2) Défense de la thèse de Martin-Löf.

La classe des suites infinies aléatoires au sens de Martin-Löf, elle aussi, présente cette résistance aux modifications dans la formulation des définitions. L'une des plus remarquables invariances est celle que Chaitin a mis en avant récemment : exiger que $H(x_n) > n - c$ est équivalent à exiger que $H(x_n) - n$ tende vers l'infini. Mais il y en a bien d'autres et on peut par exemple, faire varier de nombreuses façons différentes les éléments entrant dans la définition des tests effectifs de Martin-Löf.

(e3) Comparaison et remarques.

La confluence et la robustesse montrent que les notions mathématiques de fonctions récursives partielles et de suites aléatoires au sens de Martin-Löf, de la même façon que le nombre π ou le corps des nombres complexes, sont des objets mathématiques doués

d'ubiquité; on les rencontre à propos de nombreux problèmes de nature différente, ce sont des notions profondes et intrinsèques. Là encore cela renforce la conviction qu'on a exhumé des notions mathématiques importantes, qui ne peuvent être que les formes mathématiques des concepts intuitifs, qu'on visait.

Concernant la propriété de robustesse, remarquons quand même que, bien sûr, elle n'est pas absolue, et qu'en particulier vis à vis des problèmes de finitude tout ne se passe pas très bien. Même si une fonction f est telle que toutes ses approximations finies (coïncidant avec f pour tous les entiers inférieurs à m et non définie ailleurs) sont récursives, il se peut que f ne soit pas récursive. Des problèmes analogues concernant les suites aléatoires se produisent aussi (voir [Daley 1975])

(f) Arguments de résistance au temps et aux propositions concurrentes.

(f1) Défense de la thèse de Church.

Depuis qu'elle a été énoncée la thèse de Church (standard) n'a pas été vraiment contestée, et aucune proposition concurrente sérieuse n'a été formulée. Mieux, d'année en année elle a été admise par tous les mathématiciens qui s'y sont intéressés, et s'est vue renforcée par de nouveaux arguments. La mise au point des ordinateurs constitue même une sorte de confirmation concrète indirecte de la thèse de Church : en effet aucune des architectures de machine, et aucun des langages élaborés en plus de quarante ans d'informatique n'est venu infirmer la thèse de Church.

(f2) Défense de la thèse de Martin-Löf.

Pour la thèse de Martin-Löf les choses sont bien moins claires, d'abord elle est plus récente, et ensuite elle a été contestée, et ne s'est imposée face aux propositions concurrentes qu'après une période incertaine qui vient juste de s'achever : la définition de la famille des suites aléatoires au sens de Martin-Löf en terme de complexité algorithmique —grâce à la seconde théorie algorithmique de l'information— est relativement récente. La thèse de Martin-Löf est acceptée par la majorité des spécialistes, en particulier par Chaitin, par Kolmogorov [1987] et Uspenskii, par Gac, par Schnorr (avec une légère restriction voir plus loin), et Levin.

(f3) Comparaison et remarques.

Incontestablement la thèse de Martin-Löf ne peut se prévaloir ni d'un accord aussi unanime et ancien que celui qui règne à propos de la thèse de Church, ni d'une résistance aussi avérée face aux propositions concurrentes.

Il semble totalement impensable qu'on revienne en arrière à propos de la thèse de Church —au moins dans un cadre de philosophie classique des mathématiques— car cela signifierait que depuis cinquante ans, tous ceux qui se sont intéressés à la question ont commis un grave oubli ou une grave erreur d'appréciation. Un retour en arrière semble impensable pour la thèse de Martin-Löf, mais moins de gens y ont réfléchi, moins de conséquences en ont été tirées, et donc la conviction qu'on peut avoir en sa justesse n'est pas aussi fortement enracinée. Dans 30 ou 50 ans peut-être se seront-elles rejointes, aujourd'hui l'une est majeure et l'autre juste sortie de l'enfance.

(g) Arguments d'effectivité, de fécondité et d'utilité.

(g1) Défense de la thèse de Church.

On ne peut espérer démontrer que quelque chose est vrai (ou correct) parce que fécond, utile ou effectif. Mais l'indifférence qu'on éprouve vis à vis d'une thèse sans utilité, sans fécondité et impraticable (car par exemple faisant intervenir des éléments

mathématiquement non constructifs) pourrait conduire à ne pas découvrir son incorrection. Et donc la fécondité, l'utilité et l'effectivité d'une thèse en en garantissant une mise à l'épreuve minimale, sont indirectement des arguments positifs en sa faveur.

Là encore une différence indiscutable existe entre les deux thèses qui nous occupent. La thèse de Church est mathématiquement féconde : nous avons déjà dit qu'on l'utilise quotidiennement en théorie de la récursivité, laquelle d'une certaine façon n'aurait pas pu se développer, et n'aurait aucun sens sans elle. La thèse de Church est concrètement utilisée en informatique lorsqu'on juge qu'un langage (par exemple Prolog) est algorithmiquement complet dès qu'on a établi que toute fonction récursive peut s'y programmer. La thèse de Church est effective à l'opposé de la thèse de Martin-Löf, car elle concerne des objets concrètement appréhendables.

(g2) Défense de la thèse de Martin-Löf.

La thèse de Martin-Löf est certes féconde (voir par exemple: [Chaitin Schwartz 1978] [Li Vitanyi 1989]) mais elle est tellement peu effective par nature (à cause des versions fortes des théorèmes de Gödel qui y sont liés [Chaitin 1974]) qu'on a tenté de contourner les problèmes qu'elle pose en définissant d'autres notions plus effectives de hasard algorithmique (voir [Goldreich 1988] et paragraphe 5 (b)).

(g3) Comparaison et remarques.

Difficilement utilisable en pratique, peu confrontée aux faits concrets —en particulier de l'informatique— la thèse de Martin-Löf se révèle finalement décevante. Il se peut qu'elle puisse avoir son importance en physique ([Bennett 1988]), pourquoi pas en biologie ([Chaitin 1979]), en philosophie (Levin l'utilise à propos de l'intuitionnisme [Levin 1976a] [Levin 1984]) mais aujourd'hui elle semble bien faible en regard de sa grande sœur, la thèse de Church.

Finalement bien qu'une comparaison point par point des deux thèses puisse être menée —il s'agit bien de deux thèses philosophiquement de la même nature— on pressent que jamais la seconde ne jouera un rôle comparable à la première, ni même ne réussira à créer une unanimité aussi puissante en sa faveur. Notons encore un dernier élément faussant la symétrie entre les deux thèses : la thèse de Martin-Löf présuppose la thèse de Church et donc, bien sûr, sera toujours plus incertaine.

Dans le paragraphe suivant, nous montrons d'ailleurs qu'au tableau peut-être un peu simplificateur que nous avons proposé, il faut ajouter des éléments qui encore une fois, font douter de l'analogie de statut à terme des deux thèses.

5 Trois doutes.

Telle que nous l'avons présentée jusqu'à présent, l'histoire parait simple et belle : un concept petit à petit s'élabore, se précise et triomphe des propositions concurrentes, et après quelques péripéties inévitables tout le monde tombe d'accord, un nouveau pas a été franchi dans l'histoire des idées, la brume s'est dissipée, la thèse de Martin-Löf (aux incompréhensions près) comme la thèse de Church devient une pierre d'appui sur laquelle l'histoire peut continuer de se construire.

Malheureusement ce n'est pas si simple. Et nous avons simplifié le tableau en ne mentionnant que deux versions de la théorie de la complexité algorithmique et une seule définition de la notion de suite aléatoire par les tests statistiques, mais aussi parce que aujourd'hui il n'est pas si clair que cela qu'un accord soit établi sur la thèse de Martin-Löf.

La comparaison que nous avons faite entre la thèse de Church et celle de Martin-Löf a bien mis en évidence que cette dernière dispose en sa faveur de moins bons arguments que l'autre. Trois points méritent d'être précisés.

(a) Les relativismes.

L'identification proposée par Martin-Löf semble aujourd'hui acceptée, mais il faut mentionner une identification concurrente qui a été très sérieusement argumentée par Schnorr.

> The deficiency residing in the previous concepts of randomness is, in our opinion, that properties of random sequences are postulated which are of no significance to statistics. Many insufficient approaches have been made until a definition of random sequences was proposed by Martin-Löf which for the first time included all standard statistical properties of randomness. However, the inverse postulate now seems to have been violated. The acceptable definition of random sequences cannot be any formulation of recursive function theory which contains all relevant statistical properties of randomness, but it has to be precisely a characterization of all those properties of randomness that have a physical meaning. These are intuitively those properties that can be satisfied by statistical experience. This means that a sequence fails to be random in this sense if and only if there is an effective process in which this failure becomes evident. C.P. Schnorr, A unified approach to the definition of random sequence, Math. Systems Theory, 5, 1971, pp. 246-258.

La "thèse de Schnorr" (que mathématiquement on peut définir en disant qu'elle consiste à exiger dans la définition 1 donnée au paragraphe 3 que la fonction $n \to \mu(A_i)$ soit récursive) a été très bien défendue, et mis à part qu'elle est un peu moins simple, elle peut se prévaloir d'un assez grand nombre d'arguments bien souvent parallèles à ceux retenus pour la thèse de Martin-Löf. Finalement il y aurait au moins deux notions candidates de suite aléatoire, l'une un peu plus effective que l'autre, l'une un peu plus simple.

Un certain nombre d'articles récents semble pencher vers ce point de vue [Schnorr Fuchs 1977] [Gaifman Snir 1982] [van Lambalgen 1987]. L'avis de Schnorr est plus compliqué :

> I do not think that there is a simple natural concept of random sequence. The situation is comparable with the relationship between partial and total recursive functions. You cannot say that one of these concepts captures the natural meaning of "computable function" and the other does not. Martin-Löf definition corresponds to partial recursive functions and mine to recursive functions. I agree that Martin-Löf concept is more basic and more simple. The universal test is a very nice concept as the universal Turing Machine. However if it comes to applications the universality is very problematic. The universal Turing machine does not tell you how build an efficient computer; nor does the universal test tell you how to test randomness efficiently. So far of the work on randomness is purely theoretical with no applications in sight. C.P. Schnorr 1989, lettre personelle.

L'analogie proposée ne me semble pas très convaincante. En effet on peut dire que la notion de fonction récursive partielle est la bonne notion effective correspondant au concept de fonction partielle, et que la notion de fonction récursive totale (définie partout) est la bonne notion effective correspondant au concept de fonction totale : une seule thèse —celle de Church !— donne les deux identifications qui ne s'opposent nullement. Je pense en définitive qu'il faut interpréter le point de vue de Schnorr d'aujourd'hui comme prudent, il ne défend pas vraiment sa thèse contre celle de Martin-Löf, et reconnait même que la notion de Martin-Löf est plus simple.

Et à y regarder de plus près, en allant encore vers plus d'effectivité, ou au contraire en allant vers un peu moins d'effectivité ([Martin-Löf 1970], [Gaifman Snir 1982]), on peut trouver d'autres notions qui, elles aussi, semblent de bonnes candidates pour définir la notion d'aléatoire. Et ainsi au-delà de la proposition de Schnorr et de l'hésitation entre deux possibilités, on peut aller encore vers plus de relativisme. Dans un article récent un point de vue relativiste extrême (tempéré par des remarques que nous ne reproduisons pas) a été développé.

> Random sequences are those which exhibit phenomena whose probability is one. ... In all proposal randomness is defined as the satisfaction of a certain class of properties that have probability 1. The differences arise out of choosing different classes. H. Gaifman, M. Snir. Probabilities over rich languages, randomness and testing. Journal of Symbolic Logic. Vol. 47. 1982. pp. 495-548.

Si on refuse de trancher et en acceptant d'aller dans le sens proposé par Schnorr et poussé à l'extrême par Gailfman et Snir, le paysage final serait qu'il y a une famille (discrète) de propositions raisonnables pour définir la notion de suites infinies aléatoires, et que celle qui peut avancer la meilleure et la plus simple collection d'arguments est la proposition de Martin-Löf, mais que d'autres propositions ne sont pas absurdes : la proposition de Schnorr par exemple arrivant un peu derrière mais pas très loin.

En admettant ce point de vue on peut alors remarquer que ce type de situations n'est pas totalement nouveau, et que bien souvent l'utilisation de la théorie de la récursivité conduit à plusieurs "effectivisations" possibles, des notions classiques. L'exemple proposé par Schnorr, de l'hésitation entre fonctions partielles et fonctions totales nous paraît mal choisi, car si on veut "effectiviser" la notion de *fonction totale* on doit prendre la notion de *fonction récursive totale*, et si on veut "effectiviser" la notion de *fonction partielle* on doit prendre la notion de *fonction récursive partielle*. Par contre un parallèle est sans doute possible avec les difficultés qu'on rencontre quand on veut "effectiviser" la notion de *suite convergente de nombres réels*. On peut effectiviser cette notion par *suite convergente de nombres rationnels*, par *suite convergente de nombres réels calculables* ou par *suite calculable convergente de nombres réels calculables*, ou par *suite calculable effectivement convergente de nombres réels calculables*. Bien que plus compliquée, la quatrième définition est dans bien des cas préférable aux autres, quoi qu'aucune ne soit définitivement meilleure.

Malgré tout, le relativisme dans le cas de la notion de suite infinie aléatoire différerait au moins par le fait que son histoire a été bien plus compliquée. Plus aucune définition formulée avant la théorie de la calculabilité ne tient encore aujourd'hui. Celle de Wald [Wald 1938] (voir aussi [Martin-Löf 1970] consistant à définir l'ensemble de suites aléatoires comme l'intersection de tous les ensembles de mesure 1 définissables dans la théorie des Principia aurait été séduisante sans le théorème de Gödel qui en démontre l'insuffisance et qui conduit à voir toute proposition analogue comme 'trop faible'. La théorie de la calculabilité a ouvert la porte à une foule de propositions possibles (celle de Church étant la première) chacune s'appuyant sur la thèse de Church, mais d'une façon différente. Enormément de progrès et de résultats mathématiques ont pu être obtenus, qui ont nettoyé le tableau des propositions possibles, l'introduction de la notion de complexité algorithmique a mis un peu de relief dans le paysage et finalement désigne la définition de Martin-Löf comme le meilleur compromis possible.

La situation d'arrivée n'est pas totalement analogue à celle obtenue pour la thèse de Church, tous les arguments se présentant sous des formes affaiblies, mais elle est aussi très différente de celle qu'on trouve quand on cherche à effectiviser la notion de suite de réels convergente. Nous ne revenons donc pas en arrière sur la conclusion des paragraphes précédents.

(b) Une approche moins ontologique.

La grave ineffectivité de la notion de suite infinie aléatoire de Martin-Löf et les développements de la théorie de la complexité dynamique des programmes ont conduit à des développements particulièrement intéressants, qui se fondent sur la distinction entre 'effectif' (existence d'algorithmes) et 'pratiquable' (existence d'algorithmes polynomiaux). Cette approche toute récente ne devrait pas remettre en cause les conclusions précédentes (pas plus que la notion d'algorithme praticable ne remet en cause la notion d'algorithme), mais il ne fait aucun doute qu'elle va jouer un rôle important et que du point de vue pratique elle risque d'avoir beaucoup plus d'impact que la notion absolue de Martin-Löf.

> In this approach a probability distribution is considered "pseudorandom" if no "efficient procedure" can distinguish it from the uniform probability distribution. Remarkably, pseudorandomness so defined is expandable in the sense that (assuming the existence of 1-1 and onto one-way functions) short pseudorandom sequences can be deterministically and efficiently expended into much longer pseudorandom sequences. ... A Kolmogorov-random string is [...] a string which does not have a substantially simpler (i.e. shorter) explanation than itself. Considering the simplest explanation of a phenomenon is certainly an ontological approach. In contrast, considering the effect of phenomena on certain objects, as underlying the definition of pseudorandomness, is a behavioristic approach. [...] We conclude that the randomness of an event is relative to the information and computing resources at our disposal. Pseudorandom ensembles are unpredictable by probabilistic polynomial-time machines (associated with feasible computations), but may be predictable by infinitely powerful machines (not at our disposal!). [...] Another interesting property of the above approach to randomness is that pseudorandomness is effective in the following two senses : First, one may construct an efficient (universal) test that distinguishes pseudorandom distributions from ones which are not pseudorandom. In constrast, the problem of determining whether a string is Kolmogorov-random is undecidable. Second, assuming the existence of one-way 1-1 and onto functions, long pseudorandom strings can be efficiently and deterministically generated from much shorter pseudorandom strings. [...] The existence of pseudorandom generator has applications to the construction of efficient probabilistic algorithms (Turing machines). Such algorithms maintains the same performance when substituting their internal coin tosses by pseudorandom sequences. O. Goldreich, Randomness, Interactive Proofs, and Zero-Knowledge, In "The Universal Turing Machine: A Half-Century Survey", Edited by R. Herken, Oxford University Press, 1988, pp. 376-405.

Sur la possibilité de définir un concept de suite aléatoire prenant en compte l'efficacité des algorithmes et en particulier l'efficacité des tests statistiques, voir [Blum Micali 1984] [Ko 1986] [Goldreich Goldwasser Micali 1986] [Chor Goldreich 1988].

(c) Une définition non classique qui réhabiliterait von Mises.

Une autre voie semble encore envisageable, elle est suggérée par van Lambalgen et part de la remarque que pour montrer l'inconsistance de la notion de collectif, on utilise un argument existentiel non constructif (appelé parfois argument de Kamke) et qui est : si x est un prétendu collectif alors parmi tous les procédés de sélection possibles il y a celui qui consiste à ne garder que les 1 (et ce procédé préexiste à x) et il y a celui qui consiste à ne garder que les 0. Ces deux procédés de sélection donnent des sous-suites qui, bien sûr, ne peuvent pas avoir les mêmes fréquences limites de 0 et de 1.

> The author is convinced that a satisfactory treatment of random sequences is possible only in set theory lacking the power set axiom, in which random

> sequences "are not already there". [...] Even if we uncritically accept classical mathematics, Kamke's argument is somewhat beside the mark in that it fails to appreciate the purpose of the von Mises' axiomatization. It refers to what *could* happen, whereas von Mises' axioms are rooted in experience and refer to what *does* happen. An analogy, which turns out to have heuristic value, may be helpful here. In various places von Mises likens condition (2) [about extracted sequence] to the first law of thermodynamics. Both are statements of impossibility : condition (2) is the "principle of the excluded gambling strategy", while the first law (conservation of energy) is equivalent to the impossibility of perpetuum mobile of the first kind. It may be more appropriate to compare condition (2) to the second law of thermodynamics, the law of increase of entropy or the impossibility of a perpetumm mobile of the second kind, especially in view of Kamke's criticism. Indeed Kamke's objection is reminiscent of Maxwell's celebrated demon, that "very observant and neat-fingered being", invented to show that entropy decreasing evolution may occur. Maxwell's argument of course in no way detracts from validity of the second law, but serves to highlight the fact that statistical mechanics cannot provide an absolute foundation for entropy increase, since it does not talk about what actually happens. [...] Summarising we can say that the intent of von Mises' axioms is not affected by Kamke's criticism; the question how to develop an adequate formalisation is still open. M. van Lambalgen, Von Mises' definition of random sequences reconsidered, Journal of Symbolic Logic, Vol 52, 1987, pp. 725-755 (page 728)

Une réponse partielle a depuis été donnée par van Lambalgen lui-même [van Lambalgen 1990].

Conclusion.

La thèse de Martin-Löf est aujourd'hui bien établie et même si de nouvelles idées doivent encore être explorées (voir paragraphe 5), on peut penser qu'elles viendront enrichir et compléter le paysage aujourd'hui fixé, mais sans le bouleverser. Un progrès analogue à celui résultant de la mise au point des concepts récursivistes (et fondés par la thèse de Church) vient de se produire : nous avons maintenant un concept bien justifié de **suite infinie aléatoire** (fondé par la thèse de Martin-Löf). Il faudra sans doute de nombreuses années avant qu'il ne soit assimilé et qu'on en mesure toute l'importance mathématique, physique et philosophique.

Bibliographie.

[Barzdin' 1968] Y. M. Barzdin'. The Complexity of Programs to Determine Whether Natural Numbers not Greater than n, Belong to a Recursively Enumerable Set. Soviet Math. Dokl. 9. 1968. pp. 1251-1254.

[Bennett 1979] C. H. Bennett. On Random and Hard-to-Describe Numbers. IBM Research Report RC 7483 1-16-79, IBM Watson Research Center. 1979.

[Bennett 1988] C. H. Bennett. Logical Depth and Physical Complexity. In "The Universal Turing Machine : A Half-Century Survey". Edited by R. Herken. Oxford University Press. 1988. pp. 227-257.

[Blum Micali 1984] M. Blum, S. Micali. How to Generate Cryptographically Strong Sequences of Pseudo-Random bits. SIAM J. Comp. 13. 1984. pp 850-864.

[Borel 1909] E. Borel. Presque tous les nombres réels sont normaux. Rend. Circ. Mat. Palermo. 27. 1909. pp. 247-271.

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

[Chaitin 1969a] G. J. Chaitin. On the Length of Programs for Computing Finite Binary Sequences, Statistical Considerations. J.A.C.M. 16. 1969. pp.145-159. Repris dans [Chaitin 1987a].

[Chaitin 1974] G. L. Chaitin. Information Theoretic Limitations of Formal Systems. J.A.C.M. 1974. pp. 403-424. Repris dans [Chatin 87a].

[Chaitin 1975a] G. L. Chaitin. A theory of program size formally identical to information theory. J.A.C.M. 22. 1975. pp. 329-340. Repris dans [Chaitin 87a].

[Chaitin 1975b] G. J. Chaitin. Randomness and Mathematical Proof. Scientific American. 232. May 1975. pp. 47-52.

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

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

[Chaitin 1987b] G. J. Chaitin. Algorithmic information theory. Cambridge Tracts in Theoretical Computer Science 1. Cambridge University Press, New York, 1987.

[Chaitin 1987c] G. J. Chaitin. Incompleteness Theorems for Random Reals. Advances in Applied Mathematics. 8. 1987. pp. 119-146. Repris dans [Chaitin 1987a]

[Chaitin 1988] G. J. Chaitin. Randomness in Arithmetic. Scientific American. July 1988. pp. 80-85.

[Chaitin Schwartz 1978] G. J. Chaitin and J.T. Schwartz. A note on Monte Carlo Primality Tests and Algorithmic Information Theory. Communication on Pure and Applied Mathematics. 31. 1978. pp. 521-527. Repris dans [Chaitin 1987a].

[Champernowne 1933] D. G. Champernowne. The construction of decimal normal in the scale of ten. J. London Math. Soc. 8. 1933. pp. 254-260.

[Chor Goldreich 1988] B. Chor, O. Goldreich. Unbiased bits Sources of Weak randomness and Probabilistic Communication Complexity. SIAM J. Comp. 17, 2. 1988. pp. 230-261.

[Church 1940] A. Church. On the concept of a random sequence. Bulletin Amer. Math. Soc. 46. 1940. pp. 130-135.

[Cover Gacs Gray 1989] T. M. Cover, P. Gacs, R. M. Gray. Kolmogorov's contributions to information theory and algorithmic complexity. The Annals of Probability, Vol 17, n°3, 1989, pp. 840-865.

[Daley 1973] R. P. Daley. Minimal-Program Complexity of Sequence with Restricted Resources. Information and Control. 23. 1973. pp. 301-312.

[Daley 1975] R.P. Daley. Minimal-Program Complexity of Pseudo-Recursive and Pseudo-Random Sequences. Mathematical System Theory, vol 9, n° 1, 1975. pp. 83-94.

[Daley 1976] R.P. Daley. Noncomplex Sequences : Characterizations and Examples. J. Symbol. Logic. 41. 1976. pp. 626

[Delahaye 1988a] J.P. Delahaye. Une Extension Spectaculaire du Théorème de Gödel. La Recherche n°200, juin 1988. pp. 860-862.

[Delahaye 1988b] J.P. Delahaye. Sequence transformations. Springer-Verlag, Series in Computational Mathematics. Berlin. 1988.

[Delahaye 1989a] 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 1989b] J.-P. Delahaye. Chaitin's Equation: An Extension of Gödel's Theorem. Notices of The American Mathematical Society. October 1989. pp. 984-987.

[Delahaye 1990] J.-P. Delahaye. Randomness, Unpredictability and Absence of Order. 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. A paraître : Kluwer Academic Publ.

[Dies 1976 1978] J. E. Dies. Information et Complexité. Annales de L'institut Henry Poincaré, Section B, Calcul des Probabilités et Statistiques. Nouvelle Série Vol 12. 1976. pp.365-390 et Vol 14. 1978. pp. 113-118.

[Fine 1973] T. Fine. Theories of Probability. Academic Press. New-York. 1973.

[Gacs 1974] P. Gacs. On the symmetry of algorithmic information. Dokl. Akad. Nauk SSSR. 218. 1974. pp. 1477-1480.

[Gacs 1980] P. Gacs. Exact Expressions for Some Randomness Tests. Z. Math. Logik. Grundl. Math. 26. 1980.

[Gacs 1983] P. Gacs. On the relation between descriptional complexity and probability. Theo. Comp. 22. 1983. pp.71-93.

[Gacs 1986] P. Gacs. Every Sequence Is Reducible to a Random One. Information and Control. 70. 1986. pp. 186-192.

[Gaifman Snir 1982] H. Gaifman, M. Snir. Probabilities over rich languages, randomness and testing. Journal of Symbolic Logic. Vol. 47. 1982. pp. 495-548.

[Gardner 1980] M. Gardner. Le nombre aléatoire oméga semble bien recéler les mystères de l'univers. Pour La Science. 1980.

[Goldreich 1988] O. Goldreich. Randomness, Interactive Proofs, and Zero-Knowledge. In "The Universal Turing Machine : A Half-Century Survey". Edited by R. Herken. Oxford University Press. 1988. pp. 376-405.

[Goldreich Goldwasser Micali 1986] O. Goldreich, S. Goldwasser, S. Micali. How to construct random functions. J. A. C. M. 33, 4. 1986. pp. 792-807.

[Ko 1986] K. Ko. On the notion of Infinite pseudo-random sequences. Theoretical Computer Science. 48. 1986. pp. 9-33

[Kolmogorov 1933] A. N. Kolmogorov. Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer Verlag, Berlin, 1933. (2nd Russian Edition 1974, 'Osnovnye Poniatija Teorii Verojatnostej' Nauka, Moscow)

[Kolmogorov 1936] A. N. Kolmogorov. Osnovye ponyatiya teorii veroyatnostei. ONTI. Moscow, 1936. Traduction anglaise: Foundations of the theory of probability. Chelsea, New-York. 1950.

[Kolmogorov 1963] A. N. Kolmogorov. On table of random numbers. Sankhya The Indian Journal of Statistics. A25. 369. 1963. pp. 369-376.

[Kolmogorov 1965] A. N. Kolmogorov. Three approaches for defining the concept of information quantity. Information Transmission. V. 1. 1965. pp. 3-11.

[Kolmogorov 1968] A. N. Kolmogorov. Logical basis for Information Theory and Probability Theory. IEEE Transaction on Information Theory. Vol. IT14, n°5. 1968. pp. 662-664.

[Kolmogorov 1968] A. N. Kolmogorov. Some Theorems on algorithmic entropy and the algorithmic quantity of information. Uspeki Mat. Nauk, Vol. 23:2. 1968. pp. 201.

[Kolmogorov 1983] A. N. Kolmogorov. Combinatorial foundations of information theory and the calculus of probabilities. Russian Mathematical Surveys. Vol. 38.4. 1983. pp. 29-40.

[Kolmogorov 1983] A. N. Kolmogorov. On logical foundations of probability theory. In "Probability Theory and Mathematical Statistics. Lecture Notes in Mathematics." Ed. K. Ito and J. V. Prokhorov. Vol. 1021. Springer-Verlag., Berlin. 1983. pp. 1-5.

[Kolmogorov Uspenskii 1987] A.N. Kolmogorov and V. A. Uspenskii. Algorithms and Randomness. SIAM Theory Probab. Appl. Vol. 32. 1987. pp. 389-412.

[Levin 1973] L. A. Levin. On the notion of random sequence. Dokl. Akad. Nauk SSSR. 212, 3. 1973.

[Levin 1974] L. A. Levin. Laws of information conservation (non-growth) and aspects of the foundation of probability theory. Problems Inform. Transmission. 10 n°3. 1974. pp. 206-210.

[Levin 1976a] L. A. Levin. On the principle of conservation of information in intuitionistic mathematics. Dokl. Akad. Nauk. SSSR. Tom 227 n°6. 1976. Soviet Math. Dokl. 17 n°2. 1976. pp. 601-605.

[Levin 1976b] L. A. Levin. Various measures of complexity for finite objects (axiomatic descriptions). Soviet Math. Dokl. 17. n°2. 1976. pp. 552-526.

[Levin 1976c] L. A. Levin. Uniform tests of randomness. Soviet Math. Dokl. 17, n°2. 1976. pp. 337-340.

[Levin 1984] L. A. Levin. Randomness conservative inequalities: Information and independence in mathematical theories. Inf. Contr. 61. 1984. pp. 15-37.

[Levin 1985] L. A. Levin. One-way function and pseudorandom generators. In "Proceedings of the 17th ACM Symposium on Theory of Computing". 1985. pp. 363-365.

[Levin 1989] L. A. Levin. Lettre à l'auteur au sujet des aspects historiques de la théorie algorithmique de l'information. Novembre 1989.

[Levin V'Yugin 1977] L. A. Levin and V. V. V'Yugin. Invariant Properties of Informational Bulks. Lecture Notes in Computer Science n°53. Springer, Berlin. 1977. pp. 359-364.

[Li Vitanyi 1989a] M. Li, P. M. B. Vitanyi. A New Approach to Formal Language Theory by Kolmogorov Complexity. Proc 16th International Colloquium on Automata Languages and Programming. 1989.

[Li Vitanyi 1989b] M. Li, P. M. B. Vitanyi. Inductive Reasoning and Kolmogorov Complexity. Proc. 4th Annual IEEE Structure in Complexity Theory Conference. 1989.

[Li Vitanyi 1990] M. Li, P. M. B. Vitanyi. Kolmogorov Complexity and Its Applications. Handbook of Theoretical Computer Science. J. van Leeuwen Editor. North-Holland. 1990. pp. 187-254

[Li Vitanyi 1990] M. Li, P. M. B. Vitanyi. Introduction to Kolmogorov Complexity and Its Applications. Addison-Wesley, Reading, Mass. To appear.

[Loveland 1964] D. W. Loveland. Recursively Random Sequences. Ph. D. Thesis. N.Y.U. June 1964.

[Loveland 1966a] D. W. Loveland. A new interpretation of the von Mises' Concept of random sequence. Zeitschr. F. Math. Logik und Grundlagen d. Math. Bd12. 1966. pp. 279-294.

[Loveland 1966b] D. W. Loveland. The Kleene Hierarchy Classification of Recursively Random Sequences. Trans. Amer. Math. Soc. 125. 1966. pp. 487-510.

[Loveland 1968] D. W. Loveland. Minimal Program Complexity Measure. Conference Record ACM Symposium on Theory of Computing. May 1968. pp. 61-65.

[Loveland 1969] D. W. Loveland. A variant of the Kolmogorov concept of complexity. Information and Control. 15. 1969. pp. 510-526.

[Martin-Löf 1966] P. Martin-Löf. On the Concept of a Random Sequence. Theory Probability Appl.. Vol. 11 1966. pp. 177-179

[Martin-Löf 1966] P. Martin-Löf. The Definition of Random Sequences. Information and Control. 9. 1966. pp. 602-619.

[Martin-Löf 1969a] P. Martin-Löf. Algorithms and Randomness. Intl. Stat. Rev. 37, 265. 1969. pp. 265-272.

[Martin-Löf 1969b] P. Martin-Löf. The Literature on von Mises' Kollektivs Revisited. Theoria, XXXV. 1969. pp. 12-37.

[Martin-Löf 1970] P. Martin-Löf. On the notion of Randomness. in "Intuitionism and Proof Theory". A. Kino, J. Myhill and R.E. Vesley, eds. North-Holland Publishing Co. Amsterdam. 1970, pp.73-78.

[Martin-Löf 1971] P. Martin-Löf. Complexity Oscillations in Infinite Binary Sequences. Zeitschrift fur Wahrscheinlichkeitstheory und Vervandte Gebiete. 19. 1971. pp.225-230.

[Martin-Löf 1974] P. Martin-Löf. The notion of redundancy and its use as a quantitative measure of the discrepancy between statistical hypothesis and a set of observational data. Scand. J. Stat. Vol 1. 1974. pp. 3-18.

[Martin-Löf 1975] P. Martin-Löf. Reply to Sverdrup's polemical article "Tests without power". Scand. J. Stat. Vol. 2. 1975. pp. 161-165.

[Minsky 1961] M. L. Minsky. Steps towards Artificial Intelligence. Proceedings I.R.E. 1961. pp. 8-30.

[Minsky 1962] M. L. Minsky. Problems of Formulation for Artificial Intelligence. Mathematical Problems in Biological Sciences. Proceedings of Symposia in Applied Mathematics XIV. R.E. Bellmann ed. American Math. Soc. Providence. RI. 1962.

[O'Connor 1988] M. O'Connor. An Unpredictibility Approach to Finite-State Randomness. Journal of Computer and System Sciences. 37. 1988 pp. 324-336.

[Popper 1935] K. R. Popper. Logik der Forschung. Springer. 1935. Traduction Française: La Logique de la Découverte Scientifique. Payot, Paris. 1978.

[Schnorr 1971a] C. P. Schnorr. A unified approach to the definition of random sequence. Math. Systems Theory. 5. 1971. pp. 246-258.

[Schnorr 1971b] C. P. Schnorr. Optimal Gödel numberings. Proc. IFIP Congres 1971. Ljubljana, Yugoslavia. TA-2. pp. 12-24.

[Schnorr 1971c] C. P. Schnorr. Zufälligkeit und Wahrscheinlichkeit. Lecture Notes in Mathemathics. Vol 218. Berlin-Heidelberg-New York. Springer, 1971.

[Schnorr 1972] C. P. Schnorr. The process complexity and effective random tests. Proc. ACM Conf. on Theory of Computing. 1972. pp. 168-176.

[Schnorr 1973] C. P. Schnorr. Process complexity and effective random tests. J. Comput. Syst. Sci. 7. 1973. pp 376-388.

[Schnorr 1977] C. P. Schnorr. A survey of the theory of random sequences. In "Basic Problems in Mathodology and Linguistics." Ed. R. E. Butts, J. Hintikka. D. Reidel, Dordrecht. 1977. pp. 193-210.

[Schnorr 1989] C. P. Schnorr. Lettre à l'auteur au sujet des aspects historiques de la théorie algorithmique de l'information.. Septembre 1989.

[Shannon Weaver 1949] C.E. Shannon and W. Weaver. The Mathematical Theory of Communication. Univ. of Illinois Press, Urbana, 1949.

[Shen' 1983] A. Kh. Shen'. The concept of (α,β)-stochasticity in the Kolmogorov sense, and its properties. Soviet Math. Dokl. Vol. 28. 1983. pp. 295-299.

[Shen' 1989] A. Kh. Shen'. On relation between different algorithmic definitions of randomness. Soviet Math. Dokl. Vol. 38. 1989. pp. 316-319.

[Solomonoff 1960] R. J. Solomonoff. A preliminary report on a general theory of inductive inference. Tech. Rep. ZTB-138. Zator Company. Cambridge, Mass. November 1960.

[Solomonoff 1964] R.J. Solomonoff. A formal theory of inductive Inference. Information and Control. 7. 1964. pp. 1-22.

[Solomonoff 1978] R. J. Solomonoff. Complexity-based induction systems: comparisons and convergence theorems. IEEE Transaction on Information Theory. IT-24. 1978. pp. 422-432.

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

[van Lambalgen 1987a] M. van Lambalgen. Random sequences. Ph. D. Thesis Department of Mathematics, University of Amsterdam, 1987.

[van Lambalgen 1987b] M. van Lambalgen. Von Mises' definition of random sequences reconsidered. The Journal of Symbolic Logic. 52. 1987. pp. 725-755.

[van Lambalgen 1989] M. van Lambalgen. Algorithmic Information Theory. The Journal of Symbolic Logic. 54 n°4, 1989. pp. 1389-1400.

[van Lambalgen 1990] M. van Lambalgen. The axiomatization of randomness. The Journal of Symbolic Logic. 55. 1990. pp.1143-1167.

[Ville 1936] J. Ville. Sur la notion de collectif. C.R. Acad. Scien. Paris. 203. 1936. pp. 26-27. Sur les suites indifférentes. C.R. Acad. Scien. Paris. 202 . 1936. p.1393

[Ville 1939] J. Ville. Etude critique de la notion de collectif. Gauthier-Villars. Paris. 1939.

[von Mises 1919] R. von Mises. Grundlagen der Wahrscheinlichkietsrechnung. Math. Z. 5. 100. 1919.

[von Mises 1941] R. von Mises. On the foundation of probability and statistics. Am. Math. Statist. 12. 1941. pp.191-205.

[von Mises 1964] R. von Mises. Selected papers of Richard von Mises. Providence, Rhode Island, Amer. Math. Soc. 1964.

[von Mises Geiringer 1964] R. von Mises, H. Geiringer. Mathematical Theory of Probability and Statistics. Academic Press, New-York and London. 1964.

[Wald 1936] A. Wald. Sur la notion de collectif dans le calcul des probabilités. C. R. Acad. Sc. Paris. 202. 1936. pp. 180-183.

[Wald 1937] A. Wald. Die Widerspruchsfreiheit des Kollektivbegriffes der Wahrscheinlichkeitsrechnung. Ergebnisse eines Mathetischen Kolloquiums. 8. 1937. pp. 38-72.

[Wald 1938] A. Wald. Die Widerspruchsfreiheit des Kollektivgriffes. Actualités Sci. Indust. 735. 1938. pp. 79-99.

[Zvonkin Levin 1970] A.K. Zvonkin, L.A. Levin. The Complexity of finite object and the development of the concepts of information and randomness by means of the theory of algorithms. Russ. Math. Survey. 25, 6. 1970. pp 83-124.