Complexité de Kolmogorov

Tony Ly

La notion de complexité est largement connue lorsqu'il s'agit d'espace ou de temps. Motivée par une caractérisation non probabiliste du « hasard », la complexité de Kolmogorov fournit de fait une mesure de la complexité de description de l'information.

1 Notions de base

Dans toute la suite les codages se feront sur l'alphabet $\{0,1\}$.

Définition 1.1 Soit $\mathcal{M}$ une machine de Turing (MdT) universelle. On définit la *complexité conditionnelle de Kolmogorov*[1] de x sachant w par :

$$\mathcal{K}_{\mathcal{M}}(x|w) = \min\{|\langle p\rangle| \mid p\ MdT,\ \mathcal{M}(\langle p\rangle w) = x\}^2.$$

La *complexité (inconditionnelle) de Kolmogorov* est alors :

$$\mathcal{K}_{\mathcal{M}}(x) = \min\{\mathcal{K}_{\mathcal{M}}(x|w) + |w| \mid w \in \{0,1\}^*\}.$$

Soit une machine $\mathcal{M}$ fixée ; $\mathcal{K}_{\mathcal{M}}$ sera noté simplement $\mathcal{K}$. En guise d'illustration, énonçons quelques propriétés évidentes :

Proposition 1.1
- $\exists c \forall x,\ \mathcal{K}(x) \leq |x| + c$
- $\exists c \forall x,\ \mathcal{K}(xx) \leq \mathcal{K}(x) + c$
- $\exists c \forall x, y,\ \mathcal{K}(xy) \leq 2\,\mathcal{K}(x) + \mathcal{K}(y) + c$

Idée de la preuve :
- Soit n une MdT qui ne fait rien : $\langle n\rangle x$ décrit x.
- Soit d une MdT universelle qui simule deux fois consécutivement. Si $\langle p\rangle w$ permet de décrire x, alors $\langle d\rangle\langle p\rangle w$ permet de décrire xx.

[1] parfois aussi appelée complexité de Kolmogorov-Chaitin (1965, 1975)

[2] remarquons la subtilité suivante : $\mathcal{M}$ doit pouvoir reconnaître la fin du codage de p et le début de w, ce qui doit être pris en compte dans $\langle p\rangle$(quelque chose de peu économe mais facile à la compréhension serait de doubler chaque caractère d'un $\langle p\rangle_{naïf}$, puis de dénoter sa fin par un délimiteur comme 01 ; une autre solution serait d'indiquer en début de code la longueur d'un $\langle p\rangle_{naïf}$ et de dédoubler seulement le code de cet indicateur). On dit que le codage de p se doit d'être auto-délimitant.

– Soit b une *MdT* qui coupe une description en deux, et exécute successivement une machine universelle sur chacun des deux morceaux, à condition de les reconnaître (via l'astuce indiquée en bas de page par exemple) : $\langle b\rangle\langle\langle p_x\rangle w_x\rangle\langle p_y\rangle w_y$ décrit xy.

Le résultat suivant nous permet de dire que le choix de $\mathcal{M}$ n'a aucune importance (ou de manière plus pragmatique que la complexité de Kolmogorov n'est pas réellement affectée par le langage de programmation choisi, modulo le codage d'un interpréteur) :

Théorème 1.1 (dit d'invariance)

$$\forall\mathcal{N}\,\forall\mathcal{M}\,\exists\alpha\,\forall x,\ \mathcal{K}_{\mathcal{N}}(x)\leq\mathcal{K}_{\mathcal{M}}(x)+\alpha$$

Idée de la preuve : Si $\langle\mathcal{M}\rangle_{\mathcal{N}}$ permet à $\mathcal{N}$ de simuler $\mathcal{M}$, et si $\langle p\rangle_{\mathcal{M}}w$ permet de décrire x pour $\mathcal{M}$, $\langle\mathcal{M}\rangle_{\mathcal{N}}\langle p\rangle_{\mathcal{M}}w$ décrit x pour $\mathcal{N}$.

La notion suivante traduit le fait que pour certaines suites, l'information transportée est aussi longue que la suite elle-même, c'est-à-dire en quelque sorte un grand caractère aléatoire.

Définition 1.2 Soit $c\in\mathbb{N}$. x est dit *c-incompressible* si $\mathcal{K}(x)\geq|x|-c$; si $c=0$, on parle simplement d'*incompressibilité*.

Cette définition a bien un sens : en effet, il y a $2^{n-c}-1$ suites de longueur inférieures à $n-c$, ce qui nous assure l'existence d'au moins $2^n-2^{n-c}+1>0$ suites c-incompressibles de longueur n.

2 De jolis résultats revisités

Continuons avec quelques applications anecdotiques mais si élégantes :

2.1 Distribution des nombres premiers

Notons $\pi(n)=|\{p\in\mathcal{P}\mid p\leq n\}|$. Les encadrements de Tchebycheff donnent que $\pi(n)=\Theta\big(\frac{n}{\ln(n)}\big)$; par la méthode d'incompressibilité, on obtient un résultat asymptotiquement moindre, mais sans réel effort :

Proposition 2.1

$$\frac{n}{\log_2(n)^2}=O_{n\to\infty}(\pi(n)).$$

Idée de la preuve : Notons p_k le k^e nombre premier. Soient n un entier non nul et $p_1^{e_1}\dots p_m^{e_m}$ (où $e_i\geq 1$) sa décomposition en facteurs premiers. Alors $\langle m\rangle\frac{n}{p_m}$ est une description de n (comme d'habitude, $\langle m\rangle$ doit être auto-délimitant) et $\mathcal{K}(n)\leq\log_2(m)+2\log_2(\log_2 m)+\mathcal{K}(\frac{n}{p_m})$. Maintenant, si n

est c-incompressible, $\log_2(n) - c \leq \log_2(m) + 2\log_2(\log_2 m) + \log_2(\frac{n}{p_m})$. Il vient $\log_2(p_m) \leq c + \log_2(m) + 2\log_2(\log_2 m)$ et $p_m = O(m\log_2(m)^2)$. Le résultat en découle.

2.2 Incomplétude de Gödel

Attaquons nous maintenant à un théorème qui a déçu et émerveillé des générations de mathématiciens. Un énoncé précis nécessiterait de se placer dans une théorie récursive consistante contenant PA_{faible}, mais il est seulement utile d'en garder l'esprit, et ainsi d'accepter tout axiome vraisemblable usuel...

Théorème 2.1 [3] *Il existe m_0 tel que pour tout x et tout $m \geq m_0$, si la proposition « $\mathcal{K}(x) \geq m$ »est vraie alors elle est non prouvable.*

Idée de la preuve : Soit d une MdT qui prend en entrée un entier m, passe en revue toutes les preuves possibles (parcours en largeur) et retourne x dès qu'une preuve démontre un théorème du type « $\mathcal{K}(x) \geq m$ ». $\langle d \rangle m$ est une description de x et si on note $c = |\langle d \rangle|$ (où le codage est auto-délimitant), on obtient $\mathcal{K}(x) \leq \log_2(m) + c$. D'autre part, x vérifie $\mathcal{K}(x) \geq m$ puisqu'on en a une preuve (et que la théorie est consistante). Dès lors, cela impose $m \leq \log_2(m) + c$; en prenant m_0 assez grand, le théorème est alors vérifié !

3 Alea et méthode d'incompressibilité

Revenons avec le but originel de l'introduction de la complexité de Kolmogorov : la mesure de l'aléatoire...

3.1 Suites infinies aléatoires

Soient $w \in \{0,1\}^\infty$ et $w_{a:b}$ son sous-mot fini constitué des lettres comprises entre les positions a et b ($a \leq b$). Il serait tentant de dire qu'une suite infinie w est aléatoire lorsqu'il existe $c > 0$ tel que pour tout n, $\mathcal{K}(w_{1:n}) \geq n - c$ (ie. tout préfixe est c-incompressible) ; mais le théorème suivant montre qu'une telle définition serait inutile :

Théorème 3.1 *Soient $f : \mathbb{N}^* \to \mathbb{N}$ une fonction récursive telle que $\sum 2^{-f(n)}$ diverge et $w \in \{0,1\}^\infty$. Alors, il existe une infinité de valeurs de n pour lesquelles $\mathcal{K}(w_{1:n}|n) \leq n - f(n)$.*

Idée de la preuve :
Dans l'idée de se débarrasser d'un $O(1)$ à la fin, on va introduire une fonction auxiliaire : $F(n) = \left\lfloor \log_2 \left(\sum\limits_{i=1..n} 2^{-f(i)} \right) \right\rfloor$ et $g(n) = f(n) + F(n)$. On note

[3]formulation due à Chaitin

que g est aussi récursive et que $\sum\limits_{F(n)=m} 2^{-f(n)} \geq 2^m - 1$. Dès lors, $\sum\limits_{n=1}^{\infty} 2^{-g(n)} = \sum\limits_{m=1}^{\infty} \sum\limits_{F(n)=m} 2^{-f(n)-m} \geq \sum\limits_{m=1}^{\infty} 2^{-m}(2^m - 1)$ diverge.

Plaçons-nous maintenant sur le cercle $\mathbb{R}/\mathbb{Z}$: notons $G_n = \sum\limits_{i=1}^{n} 2^{-g(i)}$, ainsi que $I_n = [G_{n-1}, G_n[$. Pour $x \in \{0,1\}^*$, prenons $\Gamma_x = \{w \in \{0,1\}^\infty \mid w_{1:|x|} = x\}$ et $A_n = \{x \in \{0,1\}^n \mid \Gamma_x \cap I_n \neq \varnothing\}$. Comme (G_n) diverge, pour chaque w, il existe un $N_w \subset \mathbb{N}$ infini contenant les n tels que $w_{1:n} \in A_n$. La construction précédente, suivie de la place de $w_{1:n}$ dans A_n, décrit complètement $w_{1:n}$ lorsque l'on a déjà accès à $n \in N_w$; ce qui donne :
$\mathcal{K}(w_{1:n}|n) \leq \log_2 |A_n| + O(1) \leq \log_2 \frac{G_n - G_{n-1}}{2^{-n}} + O(1)$, et pour conclure : $\mathcal{K}(w_{1:n}|n) \leq n - g(n) + O(1) \leq n - f(n)$.

Corollaire 3.1 *Dans les hypothèses du théorème précédent, si de plus on suppose que $\mathcal{K}(n|n - f(n)) = O(1)$, alors il existe une infinité de valeurs de n vérifiant $\mathcal{K}(w_{1:n}) \leq n - f(n)$.*

Remarque : En prenant $f(n) = \log_2(n)$, on remarque la naïveté de la définition de prime abord naturelle.
Idée de la preuve : Soient p une description de $w_{1:n}$ (sachant n) de longueur $\mathcal{K}(w_{1:n}|n)$ et q une MdT de longueur $O(1)$ retrouvant n à partir de $n - f(n)$. Aussi $|p| \leq n - g(n) + O(1)$ et $g(n) - f(n)$ croît vers ∞ : donc $\langle q \rangle p$ est une description de $w_{1:n}$ de longueur inférieure à $n - f(n)$ (lorsque n assez grand).

Une définition intéressante demande en fait un peu plus de travail :

Définition 3.1 Soit μ une mesure de probabilités récursive sur $\{0,1\}^\infty$. Une application $\delta : \{0,1\}^\infty \to \mathbb{N} \cup \{\infty\}$ est appelée *μ-test séquentiel de Martin-Löf* si

- $\delta(w) = \sup\limits_{n \in \mathbb{N}} \{\gamma(w_{1:n})\}$ où $\gamma : \mathbb{N} \to \mathbb{N}$ tel que $V_\gamma = \{(m, y) \mid \gamma(y) \geq m\}$ est récursivement énumérable ;
- $\forall m,\ \mu(\{w \mid \delta(w) \geq m\}) \leq 2^{-m}$.

Le résultat suivant va faciliter notre travail, puisqu'il va nous permettre de nous limiter à un μ-test particulier :

Théorème 3.2 *Soit μ une mesure de probabilités récursive sur $\{0,1\}^\infty$. Il existe un μ-test séquentiel de Martin-Löf δ_μ universel : pour tout μ-test δ, il existe un $c_\delta \geq 0$ tel que pour tout $w \in \{0,1\}^\infty$, $\delta_\mu(w) \geq \delta(w) + c_\delta$.*

Idée de la preuve :
Soit $\Phi_1, \Phi_2, \ldots, \Phi_n, \ldots$ une énumération des fonctions $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$ partielles récursives. Elle contient alors tous les V_γ associés aux μ-tests, mais il faut

éliminer ceux auxquels on ne peut pas associer de μ-test. Pour cela, prenons Ψ_i prenant les mêmes valeurs que Φ_i mais vérifiant de plus $\Psi_i(n) < \infty \Rightarrow \forall k < n,\ \Psi_i(k) < \infty$. Ensuite on construit de manière effective un $\delta_i : \{0,1\}^\infty \to \mathbb{N}$ associé par l'algorithme suivant :

1. Initialiser i à 0, ainsi que $\delta(w)$ et $\gamma(w_{1:n})$ (pour tout $w \in \{0,1\}^\infty,\ n$)
2. Faire $i \leftarrow i+1$ et calculer $(y,m) = \Psi(i)$
3. Si $\gamma(y) \geq m$, alors aller à l'étape (2) ; sinon faire $\gamma(y) \leftarrow m$
4. Si pour un certain $k \in [1,m]$, $\mu(\bigcup\limits_{\gamma(z)\geq k} \Gamma_z) > 2^{-k}$ (où $\Gamma_x = \{w \in \{0,1\}^\infty \mid w_{1:|x|} = x\}$) alors terminer ; sinon faire $\delta(w) \leftarrow \max\{m, \delta(w)\}$ pour tout $w \in \Gamma_y$ et aller à l'étape (2).

Notons que, peu importe que l'algorithme termine ou non, le δ obtenu (à chaque étape on en obtient une approximation qui se révèle être exacte sur un domaine grossissant) est un μ-test récursivement énumérable. Ainsi, on a une énumération $\delta_1, \delta_2, \ldots, \delta_n, \ldots$ des μ-tests.
En posant $\delta_\mu(w) = \sup\limits_{n\in\mathbb{N}}\{\delta_n(w) - n\}$, on a le μ-test universel voulu...

Il est maintenant temps d'introduire rigoureusement la notion de suite infinie aléatoire :

Définition 3.2 Soit μ une mesure de probabilités récursive sur $\{0,1\}^\infty$. Une suite infinie w est dite *μ-aléatoire au sens de Martin-Löf* si $\delta_\mu(w) < \infty$ (ou de manière équivalente, $\delta(w) < \infty$ pour tout μ-test δ)[4].

3.2 Caractérisation des suites aléatoires

Voyons que, comme le donne l'intuition, une suite infinie « prise au hasard »est aléatoire au sens de Martin-Löf :

Théorème 3.3 *Soit μ une mesure de probabilités récursive sur $\{0,1\}^\infty$. Alors l'ensemble des suites infinies aléatoires est de mesure $\mu = 1$.*

Idée de la preuve : Soient δ un μ-test et $V_m = \{w \in \{0,1\}^\infty \mid \delta(w) \geq m\}$ (pour tout $m \in \mathbb{N}$). Alors, $\delta(w) < \infty \Leftrightarrow w \notin \bigcap\limits_m V_m$; aussi $\mu(V_m) \leq 2^{-m}$ par hypothèse, et donc $\mu(\bigcap V_m) = 0$. En prenant l'union sur les δ_i ou en l'appliquant directement à δ_μ, on obtient le résultat.

A compter de maintenant, on munit $\{0,1\}^\infty$ de la mesure de probabilités uniforme λ. Les deux théorèmes suivants encadrent alors l'ensemble des suites aléatoires au sens de Martin-Löf par des ensembles bénéficiant d'une caractérisation plus sympathique :

[4]bien sûr, ce n'est pas la seule définition de l'aléatoire qui existe ; par exemple, pour les suites infinies aléatoires au sens de Solovay, on n'impose plus $\mu(V_m) \leq 2^{-m}$ mais seulement que $\sum \mu(V_m)$ converge...

Définition 3.3 La série $\sum 2^{-f(n)}$ est dite *récursivement convergente* s'il existe une suite récurrente (n_m) telle que pour tout $m \geq 0$,
$\sum\limits_{n=n_m}^{\infty} 2^{-f(n)} \leq 2^{-m}$.

Théorème 3.4 *Soit $f : \mathbb{N} \to \mathbb{N}$ une fonction récursive telle que $\sum 2^{-f(n)}$ soit récursivement convergente. Si w est une suite infinie λ-aléatoire au sens de Martin-Löf, alors $\mathcal{K}(w_{1:n}|n) \geq n - f(n)$ pour n assez grand.*

Remarque : Le théorème 3.3 nous assure que l'ensemble des tels w est de mesure $\lambda = 1$.
Idée de la preuve : Nous allons en démontrer la contraposée. Commençons par définir $V_m = \bigcup\limits_{n \geq n_m} \{w \in \{0,1\}^\infty \mid \mathcal{K}(w_{1:n}|n) < n - f(n)\}$. C'est en fait une réunion d'intervalles semi-ouverts de longueur 2^{-n}. Comme il y a au plus $2^{n-f(n)}$ suites y vérifiant $\mathcal{K}(y|n) < n - f(n)$, on obtient facilement que $\lambda(V_m) \leq \sum\limits_{n \geq n_m} 2^{n-f(n)} 2^{-n} \leq 2^{-m}$. D'où $\lambda(\bigcap V_m) = 0$. Aussi (V_m) est une suite décroissante d'ensembles récursivement énumérables. Autrement dit, on peut lui associer un λ-test ; et si $w \in \bigcap V_m$, il s'ensuit que w n'est pas aléatoire au sens de Martin-Löf.

Lemme 3.1 *Soit $w \in \{0,1\}^\infty$. Il existe $c \geq 0$ telle que $\{n \in \mathbb{N} \mid \mathcal{K}(w_{1:n}|n) \geq n - c\}$ est infini si et seulement si $\{n \in \mathbb{N} \mid \mathcal{K}(w_{1:n}) \geq n - c\}$ est infini.*

Idée de la preuve : $(\Rightarrow)$ est trivial. Pour $(\Leftarrow)$, il existe α tel que pour tous w et n, $\mathcal{K}(w_{1:n}) \leq \mathcal{K}(w_{1:n}|n) + 2|n - \mathcal{K}(w_{1:n}|n)| + \alpha$. D'où, pour une infinité de n, $n - \mathcal{K}(w_{1:n}|n) \leq c + \alpha + 2|n - \mathcal{K}(w_{1:n}|n)|$. Or, $|k| \leq \log_2(k) + 1$, et ainsi, il existe β tel que $n - \mathcal{K}(w_{1:n}|n) \leq \beta$ pour une infinité de n.

Lemme 3.2 $\begin{array}{cccc} \gamma_\lambda : & \{0,1\}^* & \mapsto & \mathbb{N} \cup \{\infty\} \\ & w & \to & |w| - \mathcal{K}(w||w|) - 1 \end{array}$ *est un λ-test universel de Martin-Löf.*

Remarque : Attention au fait que ce test est seulement défini sur l'ensemble des suites finies (il est dit *non-séquentiel*). Remarquons néanmoins que les propriétés demandées à la définition 3.1 restent valables.
Idée de la preuve : Il suffit, comme déjà fait précédemment de vérifier que les $V_m = \{w \in \{0,1\}^* \mid \gamma_\lambda(w) \geq m\}$ satisfont les bonnes propriétés, comme dans la preuve du théorème 3.4.

Théorème 3.5 *Soit w une suite infinie telle qu'il existe une constante $c \geq 0$ et une infinité de n telles que $\mathcal{K}(w_{1:n}) \geq n - c$. Alors w est λ-aléatoire au sens de Martin-Löf.*

Remarque : On pourrait prouver que l'ensemble des tels w est aussi de mesure $\lambda = 1$.
Idée de la preuve : On utilise le lemme 3.2 : un test séquentiel étant *a fortiori* un test non-séquentiel, on a une constante $c \geq 0$ telle que $\delta_\lambda(w_{1:n}) \leq \gamma_\lambda(w_{1:n}) + c$. Comme δ_λ est croissante, $\lim_{n\to\infty} \delta_\lambda(w_{1:n}) \leq \liminf_{n\to\infty}(n - \mathcal{K}(w_{1:n}|n)) + O(1)$. Par le lemme 3.1, cette quantité est majorée et w est alors λ-aléatoire.

Comme les ensembles de suites aléatoires considérées dans ces théorèmes sont clairement distincts, on est assuré que par rapport au caractère aléatoire au sens de Martin-Löf l'une au moins des inclusions est stricte.[5]

3.3 Graphes aléatoires

Dans toute la suite on ne considérera que des graphes étiquetés $\mathcal{G} = (V, E)$ où $V = \{1, \ldots, n\}$ représente les sommets et E les arêtes. Rappelons que le *diamètre* d'un graphe est la plus grande distance séparant deux sommets ; aussi une *clique de taille k* est un sous-graphe complet possédant k sommets.
Il est à noter qu'un graphe de taille n se code aisément sur $\{0,1\}^{\frac{n(n-1)}{2}}$ (et que réciproquement toute telle suite code un graphe) : en effet, si les arêtes sont ordonnées (par exemple lexicographiquement), il suffit d'inscrire 0 ou 1 sur le i^e chiffre de la suite selon que la i^e arête soit présente ou non dans le graphe considéré. Dorénavant, c'est effectivement ce codage que nous utiliserons, confondant alors le graphe et son codage.

Définition 3.4 Un graphe $\mathcal{G}$ de taille n est dit *$\delta(n)$-aléatoire* s'il vérifie $\mathcal{K}(\mathcal{G}|n, \delta) \geq \frac{n(n-1)}{2} - \delta(n)$.

Lemme 3.3 *La proportion des graphes de taille n qui sont $\delta(n)$-aléatoires est de $1 - \frac{1}{2^{\delta(n)}}$.*

Terminons par quelques exemples de résultats que l'on peut obtenir sur cette grande proportion de graphes :

Proposition 3.1 *Soient $\mathcal{G} = (V, E)$ un graphe $\delta(n)$-aléatoire et i un de ses sommets. Alors son degré d vérifie $|d - \frac{n-1}{2}| = O\left(\sqrt{(\delta(n) + \log_2 n)n}\right)$.*

Idée de la preuve : Supposons que le degré d de i soit à une distance strictement supérieure à k de $\frac{n-1}{2}$. Posons $V_i = \{j \in V \setminus \{i\} \mid (i,j) \in E\}$. Il y a $m = \sum_{|d-\frac{n-1}{2}|>k} C_{n-1}^d$ tels ensembles possibles. En utilisant les bornes de Chernoff (sur la loi binomiale de paramètre $\frac{1}{2}$), on majore : $m \leq 2^n e^{-\frac{2k^2}{3(n-1)}}$.

[5] en fait, on pourrait montrer qu'elles le sont toutes les deux...

Pour décrire tout cela, il faut coder une MdT pour reconstituer le code initial, les valeurs de i et d (tous deux en $\lceil \log_2 n \rceil$), puis V_i (en $\log_2 m$ si on note la position de l'ensemble dans une énumération de tous les tels ensembles possibles). Aussi, il devient inutile de coder les arêtes reliées à i. Alors $\mathcal{K}(\mathcal{G}|n) \leq \log_2 m + 2\lceil \log_2 n \rceil + \frac{n(n-1)}{2} - n + O(1)$. En exploitant le fait que $\mathcal{G}$ est $\delta(n)$-aléatoire, $\log_2 m \geq n - \delta(n) - 2\lceil \log_2 n \rceil + O(1)$. Mais on sait que $\log_2 m \leq n - \frac{2k^2}{3(n-1)} \log_2 e$. Reste à conclure : $k \leq \sqrt{\frac{3}{2} \frac{n-1}{\log_2 e} (\delta(n) + 2\lceil \log_2 n \rceil + O(1))}$.

Proposition 3.2 *Soit $\mathcal{G}$ un graphe $o(n)$-aléatoire. Alors son diamètre est égal à 2.*

Idée de la preuve : Les seuls graphes de diamètre 1 étant complets, leur description sachant n est de taille constante. Supposons $\mathcal{G}$ de diamètre au moins 3. Considérons alors $i < j$ deux sommets à distance ≥ 3. Pour décrire $\mathcal{G}$, on commence par coder i et j, puis une MdT pour retrouver le code initial. Par contre, à chaque fois que l'on a une arête (i, k) il devient inutile de préciser que (k, j) n'est pas présente (sinon la distance serait de 2) ; or, i est de degré $\frac{n}{2} + o(n)$ (par la proposition précédente), ce qui veut dire que l'on a économisé $\frac{n}{4} + o(n)$ arêtes dans notre description. Ainsi, $\mathcal{K}(\mathcal{G}|n) \leq \frac{n(n-1)}{2} - \frac{n}{4} + o(n)$, ce qui contredit l'hypothèse initiale.

Proposition 3.3 *Soit $\mathcal{G}$ un graphe $o(\log_2 n)$-aléatoire. Alors il ne contient pas de clique de taille supérieure à $2\log_2(n) + 1 + o(1)$.*

Idée de la preuve : Soit $\mathcal{H}$ une clique de taille maximale m de $\mathcal{G}$. Pour décrire $\mathcal{G}$, il suffit de coder les sommets de $\mathcal{H}$ (et ce de manière auto-délomitante), puis les arêtes de « $\mathcal{G} \setminus \mathcal{H}$ »(n'oublions pas non plus le code d'une MdT reconstituant le code initial). D'où $\mathcal{K}(\mathcal{G}|n) \leq \frac{n(n-1)}{2} + o(\log_2 n) - \frac{m(m-1)}{2} + m\lceil \log_2 n \rceil + O(1)$. Mais, par hypothèse, $\mathcal{K}(\mathcal{G}|n) \geq \frac{n(n-1)}{2} + o(\log_2 n)$: la majoration de m s'ensuit...

Références

[1] SIPSER Michael : *Introduction to the Theory of Computation*, PWS, 1997

[2] CHARPENTIER Eric & LESNE Annick & NIKOLSKI Nikolaï : *L'héritage de Kolmogorov en mathématiques*, Belin, 2004

[3] LI Ming & VITANYI Paul : *An Introduction to Kolmogorov complexity and its applications*, Springer, 1997