La complexité de Kolmogorov

Adrien Nolhier

École Polytechnique, Master Foundation of Computer Science

30 décembre 2024

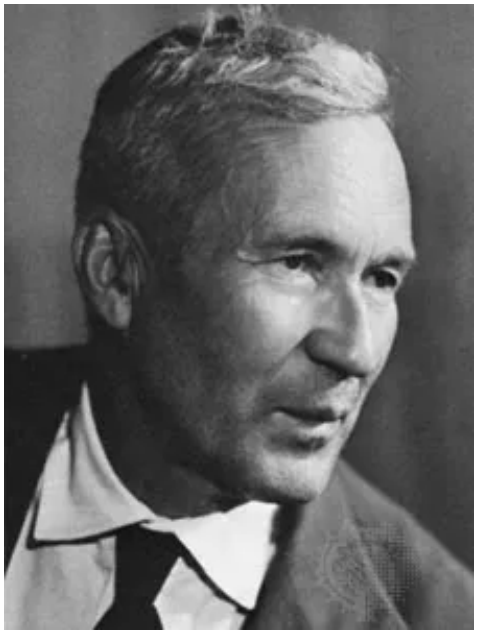
cours de Théorie de l'Information donné par Thomas Debris-Alazard, INRIA Paris-Saclay.

Qu'est-ce que l'aléatoire ?

"La complexité de Kolmogorov est une mesure de la quantité d'information nécessaire pour calculer cet objet"

Qu'est-ce que l'aléatoire ?

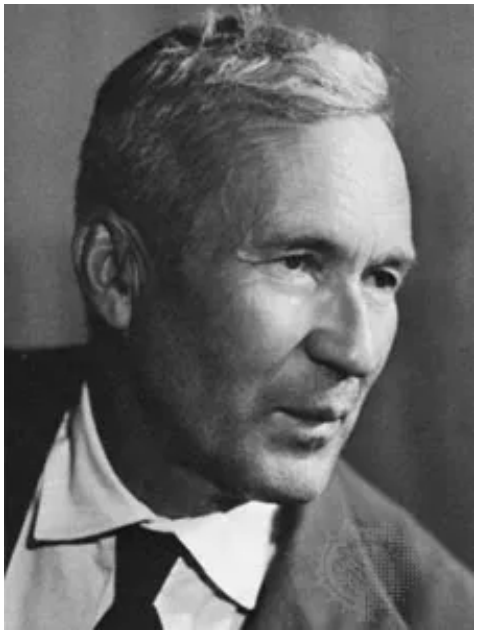
"La complexité de Kolmogorov est une mesure de la quantité d'information nécessaire pour calculer cet objet"



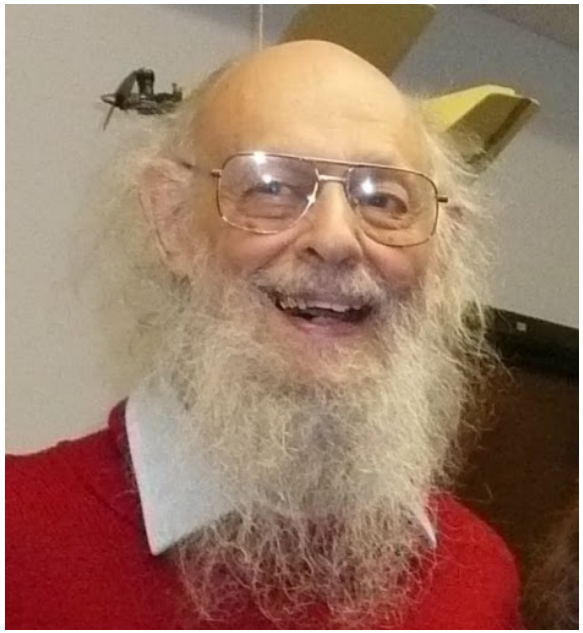
Andreï Kolmogorov (1903-1987)

Qu'est-ce que l'aléatoire ?

"La complexité de Kolmogorov est une mesure de la quantité d'information nécessaire pour calculer cet objet"



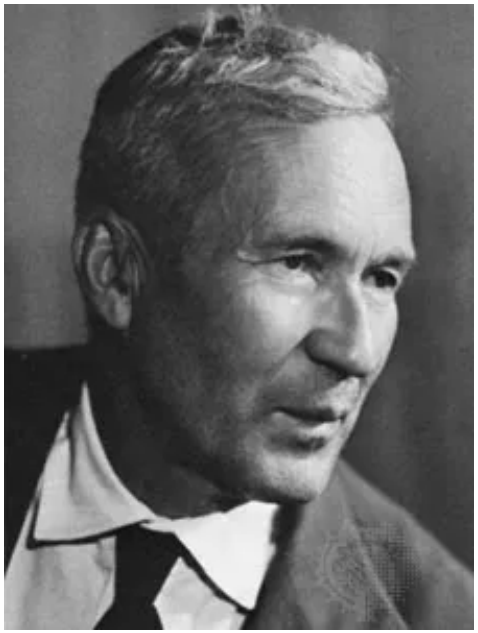
Andreï Kolmogorov (1903-1987)



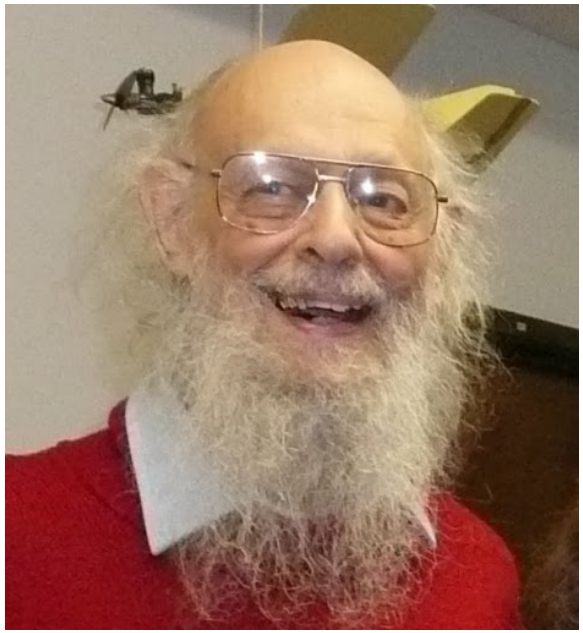
Ray Solomonoff (1926-2009)

Qu'est-ce que l'aléatoire ?

"La complexité de Kolmogorov est une mesure de la quantité d'information nécessaire pour calculer cet objet"



Andreï Kolmogorov (1903-1987)



Ray Solomonoff (1926-2009)

- Coïncide avec la définition de l'entropie
- Les objets considérés seront des chaînes sur un alphabet fini Σ
- Objets plus généraux : utiliser des représentations

Table des matières

- Un objet se défini lui-même
- On peut imaginer des machines qui reproduiraient l'objet à partir d'une description plus courte que l'objet tout entier

- Un objet se défini lui-même
- On peut imaginer des machines qui reproduiraient l'objet à partir d'une description plus courte que l'objet tout entier

Définition (Machine)

On appelle machine les fonctions **partielles** *calculables* $\Sigma^{<\mathbb{N}} \to \Sigma^{<\mathbb{N}}$.

Elles se représentent avec un entier.

- Un objet se défini lui-même
- On peut imaginer des machines qui reproduiraient l'objet à partir d'une description plus courte que l'objet tout entier

Définition (Machine)

On appelle machine les fonctions **partielles** *calculables* $\Sigma^{<\mathbb{N}} \to \Sigma^{<\mathbb{N}}$.

Elles se représentent avec un entier.

Définition (Complexité de Kolmogorov)

La complexité de Kolmogorov d'une chaîne finie $x \in \Sigma^{<\mathbb{N}}$ *sur une machine* M *est définie par :*

$$\mathrm{K}_M(x) = \min\{|p| : \mathrm{M}(p) = x\}$$

- Un objet se défini lui-même
- On peut imaginer des machines qui reproduiraient l'objet à partir d'une description plus courte que l'objet tout entier

Définition (Machine)

On appelle machine les fonctions **partielles** *calculables* $\Sigma^{<\mathbb{N}} \to \Sigma^{<\mathbb{N}}$.

Elles se représentent avec un entier.

Définition (Complexité de Kolmogorov)

La complexité de Kolmogorov d'une chaîne finie $x \in \Sigma^{<\mathbb{N}}$ *sur une machine* M *est définie par :*

$$\mathrm{K}_M(x) = \min\{|p| : \mathrm{M}(p) = x\}$$

Dans cette définition, la complexité dépend donc de la machine.

Peut-on décrire une machine, dans l'absolu ?

- Pour faire fonctionner une machine, l'on doit savoir la décrire
- Cela dépend du sens considéré comme commun

Peut-on décrire une machine, dans l'absolu ?

- Pour faire fonctionner une machine, l'on doit savoir la décrire
- Cela dépend du sens considéré comme commun

$\Rightarrow$ On supposera que la donnée d'un code informatique , d'un diagramme ou même d'un simple entier est suffisante pour décrire parfaitement une machine.

Quelle est la meilleure machine ?

Définition (Machine universelle)

Une machine U est dite universelle si pour toute autre machine M, la complexité de toute chaîne finie x ne diffère que d'une constante c_M :

$$\exists c_M \in \mathbb{N},\ \forall x \in \Sigma^{<\mathbb{N}},\ K_U(x) \leq K_M(x) + c_M$$

Quelle est la meilleure machine ?

Définition (Machine universelle)

Une machine U est dite universelle si pour toute autre machine M, la complexité de toute chaîne finie x ne diffère que d'une constante c_M :

$$\exists c_M \in \mathbb{N},\ \forall x \in \Sigma^{<\mathbb{N}},\ \mathrm{K}_U(x) \leq K_M(x) + c_M$$

$\Rightarrow$ La constante c_M est indépendante de x !

Théorème (Existence d'une UTM - Solomonoff, Kolmogorov)

Il existe une machine de Turing universelle.

Théorème (Existence d'une UTM - Solomonoff, Kolmogorov)

Il existe une machine de Turing universelle.

Démonstration.

On considère une énumération $(M_e)_{e\in\mathbb{N}}$ des machines, où M_e est la machine de code e. La machine suivante est universelle :

$$\mathsf{U} := \begin{cases} \Sigma^{<\mathbb{N}} \to \Sigma^{<\mathbb{N}} \\ 0^e 1x \mapsto \mathsf{M}_e(x) \end{cases}$$

- On peut vérifier que chaque mot (non vide) est de la forme $0^e 1x$ pour un et un seul code e.
- Pour toute machine M_e, on a bien

$$\mathsf{K}_U(x) \leq \mathsf{K}_{M_e}(x) + e + 1$$

Définition

Une machine de Turing est

- un ensemble fini d'états lisant dans un alphabet Σ une entrée finie inscrite sur une bande dite **d'entrée**
- travaillant sur un nombre fini de bandes dites **de travail**
- et inscrivant de gauche à droite sans retours une sortie finie sur une bande réservée dite **de sortie**.

La thèse de Turing-Church affirme qu'une telle machine est au moins aussi puissante que toute autre :

- Ce qui est calculable par M est calculable par une TM
- Une TM peut "simuler" toute machine M

La thèse de Turing-Church affirme qu'une telle machine est au moins aussi puissante que toute autre :

- Ce qui est calculable par M est calculable par une TM
- Une TM peut "simuler" toute machine M

Par la preuve précédente, on comprend qu'une telle machine est universelle au sens de la complexité de Kolmogorov.

$$x \text{ est aléatoire} \simeq K(x) = |x|$$

- chaîne fortement aléatoire : 01011000100101
- chaîne peu aléatoire : 101010101010101010

$$x \text{ est aléatoire} \simeq K(x) = |x|$$

- chaîne fortement aléatoire : 01011000100101
- chaîne peu aléatoire : 10101010101010101010

⚠ Peu de sens pour une chaîne finie fixée.

Proposition (Existence de chaînes aléatoires)

Pour toute taille n, il existe une chaîne finie x de taille n peu compressible, c'est à dire telle que $K(x) \geq n$.

Démonstration.

Par un argument de comptage.

- Il y a 2^n chaînes de tailles n.
- Avec jusqu'à $n-1$ bits, on peut représenter $\Sigma_{i=0}^{n} 2^i = 2^n - 1$ chaînes différentes

$\Rightarrow$ Une chaîne ne peux pas être représentée en moins de n bits. □

On peut supposer qu'une machine lit son entrée de gauche à droite, sans retours arrières.

Définition (Langage préfixe - ou "sans préfixe")

Un ensemble de mots finis L est dit sans préfixe si tout préfixe d'un mot du langage n'appartient pas au langage :

$$x \in L \implies \sigma \prec x \implies \sigma \notin L$$

On peut supposer qu'une machine lit son entrée de gauche à droite, sans retours arrières.

Définition (Langage préfixe - ou "sans préfixe")

Un ensemble de mots finis L est dit sans préfixe si tout préfixe d'un mot du langage n'appartient pas au langage :

$$x \in L \implies \sigma \prec x \implies \sigma \notin L$$

Définition (Machine sans préfixe)

Si $\sigma \prec x$ *et* $\mathrm{M}(x)\downarrow$, *alors* $\mathrm{M}(\sigma) \neq \mathrm{M}(x)$.

Inégalité de Kraft - cas particulier des codes préfixes

Théorème

Pour tout arbre binaire avec F l'ensemble de ses feuilles et $d(f)$ la profondeur de la feuille f,

$$\Sigma_{f \in F} 2^{-d(f)} \leq 1$$

Proposition

Pour toute machine U, l'inégalité suivante est vérifiée :

$$\Sigma_{\{x:U(x)\downarrow\}} 2^{-|x|} \leq 1$$

Démonstration.

Puisqu'on a supposé que les machines sont un code préfixe, on peut appliquer dans ce cas particulier l'inégalité de Kraft. □

$\Rightarrow$ La théorie sur les chaînes binaire se transpose aux entiers.

$\Rightarrow$ La théorie sur les chaînes binaire se transpose aux entiers.

Théorème (Infinité de nombres aléatoires)

Il y a une infinité d'entiers n tels que $K(n) > \log n$.

$\Rightarrow$ La théorie sur les chaînes binaire se transpose aux entiers.

Théorème (Infinité de nombres aléatoires)

Il y a une infinité d'entiers n tels que $K(n) > \log n$.

Démonstration.

Pour M une machine qui termine sur toute entrée :

$$\Sigma_n 2^{-K(n)} \leq 1$$

Or

$$\Sigma_n 2^{-\log n} = \Sigma_n \frac{1}{n} = \infty$$

S'il y avait un nombre fini de n tels que $K(n) > \log n$, il existe n_0 tel que :

$$\Sigma_{n=n_0}^{\infty} 2^{-K(n)} > \Sigma_{n=n_0}^{\infty} 2^{-\log n} = \infty$$

Complexité de Kolmogorov conditionnelle

Définition

Pour x une chaîne binaire (ou un entier), on peut définir sa complexité sachant sa longueur $|x|$:

$$K(x : |x|) = \min_{\{p \ : \ U(p,|x|)=x\}} |p|$$

Théorème (Relation entre H et K)

Pour un processus stochastique $\{X_i\}$ tiré de manière indépendante et identiquement distribué,

$$\mathbb{E}(\frac{K(X^n : n)}{n}) \rightarrow_{n\rightarrow\infty} \mathrm{H}(X)$$

Où H est l'entropie de Shannon : $\mathrm{H}(X) = -\Sigma_{i=1}^{n} p(x_i) \log p(x_i)$

Définition

Une suite $x_1, \ldots, x_n$ *est dite algorithmiquement aléatoire si*

$$K(x_1 \ldots x_n : n) \geq n$$

Définition

Une suite $x_1, \ldots, x_n$ est dite algorithmiquement aléatoire si

$$K(x_1 \ldots x_n : n) \geq n$$

Définition

Une suite infinie x est dite incompressible si :

$$\lim_{n \to \infty} \frac{K(x_1 \ldots x_n : n)}{n} = 1$$

Théorème (Loi des grands nombres pour les chaînes incompressibles)

Si une chaîne infinie x est incompressible, alors

$$\frac{1}{n}\Sigma_{i=1}^{n} x_i \rightarrow_{n\rightarrow\infty} \frac{1}{2}$$

Théorème (Loi des grands nombres pour les chaînes incompressibles)

Si une chaîne infinie x est incompressible, alors

$$\frac{1}{n}\sum_{i=1}^{n} x_i \rightarrow_{n\rightarrow\infty} \frac{1}{2}$$

"dans une suite incompressible, il y a autant de 0 que de 1"

Définition (Ω de Chaitin)

Le nombre Oméga de Chaitin est défini par :

$$\Omega = \Sigma_{\{x:U(x)\downarrow\}} 2^{-|x|}$$

- $0 \leq \Omega \leq 1$
- $\Omega = Pr(U(p)\downarrow)$, pour p tiré suivant un processus de Bernouilli de paramètre $\frac{1}{2}$.

Gregory Chaitin (1947-)

Définition (Ω de Chaitin)

Le nombre Oméga de Chaitin est défini par :

$$\Omega = \Sigma_{\{x:U(x)\downarrow\}} 2^{-|x|}$$

- $0 \leq \Omega \leq 1$
- $\Omega = Pr(U(p)\downarrow)$, pour p tiré suivant un processus de Bernouilli de paramètre $\frac{1}{2}$.

Gregory Chaitin (1947-)

On ne sait pas énumérer dans l'ordre les codes qui terminent, donc Ω est "incalculable".

Ω est dans le même degré Turing que l'arrêt

Preuve dûe à Benoît Monin et Ludovic Patey

Définition (Approchable par la gauche)

Un ensemble $X \in 2^{\mathbb{N}}$ est approchable par la gauche s'il existe une suite calculable $(X_s)_{s\in\mathbb{N}}$ telle que X_s est lexicographiquement plus petit que X_{s+1} pour tout $s \in \mathbb{N}$ et telle que X est la limite de cette suite :

$$X = \lim_{s\to\infty} X_s$$

Ω est dans le même degré Turing que l'arrêt

Preuve dûe à Benoît Monin et Ludovic Patey

Définition (Approchable par la gauche)

Un ensemble $X \in 2^{\mathbb{N}}$ est approchable par la gauche s'il existe une suite calculable $(X_s)_{s\in\mathbb{N}}$ telle que X_s est lexicographiquement plus petit que X_{s+1} pour tout $s \in \mathbb{N}$ et telle que X est la limite de cette suite :

$$X = \lim_{s\to\infty} X_s$$

Proposition

Ω est approchable par la gauche. (Donc $\emptyset'$-calculable)

Démonstration.

prendre $\Omega_s = \Sigma_{\{x:|x|\leq s \wedge U(x)[s]\downarrow\}} 2^{-|x|}$. □

Théorème (Chaitin)

$$\Omega \equiv_T \emptyset'$$

Démonstration.

Montrons que Ω est la représentation d'une machine universelle. □

Théorème

Il existe une constante c telle que Ω ne puisse pas être compressé de plus de c :

$$K(\omega_1 \dots \omega_n) \geq n - c, \ \forall n \in \mathbb{N}$$

- La complexité algorithmique est un concepte précurseur à celui de l'entropie
- On peut construire des nombres aléatoires mais c'est difficile

Résumé

Ouverture

- Considérer plus généralement "Incalculable $\Longrightarrow$ aléatoire", et l'aléatoire de choses encore plus incalculables

- Étude des fonctions réelles aléatoires, avec les machines de Turing de type 2

Bibliographie

- "Elements of Information Theory", Thomas Cover et Joy Thomas

- "Calculabilité", Benoît Monin et Ludovic Patey

Merci de votre attention.