THÈSE DE DOCTORAT EN INFORMATIQUE
–
CALCULABILITÉ, ALÉATOIRE ET THÉORIE ERGODIQUE SUR LES ESPACES MÉTRIQUES
–
RÉSUMÉ

MATHIEU HOYRUP

CONTEXTE

Entropies. Différentes notions appelées *entropies* ont été définies au cours du XX^{e}siècle. Si l'on pense à un espace mathématique comme espace des états ou des mesures d'un système (mathématique ou physique), une entropie a pour rôle de quantifier le degré d'incertitude que l'on a sur l'état du système. L'incertitude pouvant être exprimée comme un manque d'information, le terme *quantité d'information* est aussi utilisé, le cadre mathématique étant dénommé *théorie de l'information*. Suivant le point de vue adopté, qui se traduit par la structure dont on dote l'espace, cette entropie peut prendre différentes formes :

(1) Une possibilité est de mesurer la taille de l'espace, c'est-à-dire le nombre de points que l'on peut distinguer. C'est le point de vue topologique, adopté par Kolmogorov et Tikhomirov en 1959, définissant la ϵ-entropie d'un ensemble (voir [KT59]).

(2) Les probabilités offrent un cadre plus fin, permettant d'attribuer plus de poids à certaines régions de l'espace. Shannon en a tiré parti en définissant sa célèbre entropie en 1948. La non-uniformité de l'espace reflétée par les probabilités permet en outre d'associer à chaque point une quantité d'information, dont l'entropie est la moyenne. Dans le cadre topologique, l'espace étant en quelque sorte uniforme, tous les points possèdent la même quantité d'information.

(3) En 1965, Kolmogorov ([Kol65]) revient sur les deux notions d'entropies évoquées ci-dessus et, utilisant la théorie de la calculabilité, en propose une troisième. L'idée est simple : dans chaque contexte, topologique ou probabiliste, on peut interpréter la quantité d'information d'un point comme la longueur de la description du point (relativement à un certain système de description fixé), et l'entropie comme la longueur « moyenne »des descriptions des points. Or la théorie de la calculabilité offre un système de codage très général : la machine de Turing universelle permet de simuler tous les systèmes de décodage calculables. La *complexité de Kolmogorov* d'un point est alors la longueur minimale des codages du point par une machine universelle. Elle est aussi appelée quantité d'information algorithmique (voir [LV93] pour une introduction).

Entropies des systèmes dynamiques. La situation décrite ci-dessus est « statique » : on ne s'intéresse pas à l'évolution d'un système, on mesure seulement son état. L'évolution des

systèmes physiques est modélisée principalement par les systèmes dynamiques. Le déterminisme de ces modèles n'empêche pas de rendre compte de l'aléatoire, qui est alors compris comme imprédictibilité (en physique classique). Les idées de la théorie de l'information, qui on l'a vu cherchent à quantifier le degré d'incertitude, ont été appliquées aux systèmes dynamiques pour quantifier leur imprédictibilité. Chaque entropie statique possède alors sa version dynamique, qui est le taux de croissance de l'entropie statique au cours de l'évolution du système. Les versions dynamiques de la ϵ-entropie, de l'entropie de Shannon et de la complexité de Kolmogorov sont respectivement :

(1) L'entropie topologique d'une système (définie en 1965 par Adler, Konheim et McAndrew),

(2) L'entropie de Kolmogorov-Sinaï (1958, 1959),

(3) La complexité algorithmique des orbites d'un système (1983 par Brudno, [Bru83], améliorée en 2000 par Galatolo [Gal00]).

Comme dans le cas statique, des liens forts existent entre ces différentes notions. Remarquons aussi que l'approche algorithmique fournit une notion *individuelle* (attachée à chaque orbite) et en quelque sorte *intrinsèque* (définie indépendamment de la mesure).

Comme on vient de l'évoquer, la théorie des systèmes dynamiques étant un modèle du déterminisme, elle offre un cadre pour parler de l'aléatoire classique, compris comme imprédictibilité déterministe. Par ailleurs, le cadre naturel pour parler de l'aléatoire est la théorie des probabilités[1] (notons qu'elle ne se prononce pas sur la manière dont l'aléatoire est engendré). Il est donc très naturel de chercher à mélanger ces deux théories : c'est l'objet de la *théorie ergodique* qui traite les systèmes dynamiques d'un point de vue probabiliste, c'est-à-dire comme un type particulier de systèmes stochastiques (voir [Pet83]).

Aléatoire algorithmique. La théorie des probabilités offre un langage permettant d'exprimer et de prouver que telle propriété a probabilité 1, ou qu'elle est vraie pour *presque tout* point. Une idée qui trouve ses origines chez Laplace puis chez Von Mises est d'essayer d'identifier les propriétés qu'une suite binaire infinie doit vérifier pour dire qu'elle est « aléatoire ». Kolmogorov s'y est essayé dans les années 60, mais c'est Martin-Löf qui en 1966, utilisant la théorie de la calculabilité, a le premier proposé une définition convenable ([ML66]) : une suite binaire aléatoire est une suite qui satisfait, non pas toutes les propriétés de mesure 1 (de telles suites n'existent pas), mais seulement celles qui sont *algorithmiques* dans un certain sens. Une caractérisation de cette aléatoire en termes de complexité de Kolmogorov a plus tard été prouvée.

Ainsi l'on dispose d'un modèle plus fin que la théorie des probabilités, qui permet d'accoler à chaque suite symbolique infinie l'étiquette « aléatoire »ou « non-aléatoire ». La notion d'aléatoire algorithmique est alors destinée, entre autres choses, à être confrontée avec les théorèmes classiques de probabilités dont les énoncés prennent la forme :

$$\text{la propriété } P \text{ est vraie pour } \mu\text{-presque toute suite}$$

pour être convertis en théorèmes de la forme :

$$\text{la propriété } P \text{ est vraie pour toute suite } \mu\text{-aléatoire}$$

[1] dont l'axiomatisation est due à Kolmogorov en 1933

La théorie de l'aléatoire algorithmique était encore récemment limitée à l'espace des suites binaires (ou sur un alphabet fini) (voir [HW03] pour une première tentative d'extension à d'autres espaces, puis [Gác05]).
Apports de la thèse. L'objectif général de cette thèse est d'offrir un cadre de travail robuste pour traiter la théorie des probabilités et la théorie ergodique d'un point de vue algorithmique. Beaucoup de travail a déjà été effectué sur l'espace de Cantor, mais il est fondamental de pouvoir travailler sur des espaces plus généraux : se restreindre à l'espace de Cantor est en effet très limititatif en regard de la variété des espaces où les dynamiques physico-mathématiques sont étudiées. Nous nous plaçons dans les espaces métriques calculables, qui sont des espaces métriques avec des conditions raisonnables permettant d'y parler de points, fonctions, sous-ensembles calculables. Ceci nous a permis d'étendre l'aléatoire algorithmique à ces espaces. Ce travail est une amélioration d'un article récent ([Gác05]). Nous avons alors établi des liens forts entre l'aléatoire, les complexités algorithmiques des orbites d'un système dynamique et les entropies. Notre travail reprend en partie certains résultats précédemment établis sur l'espace des suites symboliques (notamment [V'y98]).

Ce qui suit est une présentation du contenu de la thèse, chapitre par chapitre.

1. Calculabilité

L'analyse calculable est aujourd'hui principalement pratiquée par deux écoles : l'une, tournée vers la théorie des langages fonctionnels, utilise le cadre de la théorie des domaines. L'école allemande utilise la théorie des représentations. La première approche (voir [Eda97] par exemple) permet d'obtenir d'intéressantes propriétés catégoriques, mais en contrepartie impose des restrictions aux espaces considérés. La deuxième approche ([Wei00]) imposant moins de structure est plus générale, mais perd la beauté de la première. En particulier, le langage utilisé ne se défait jamais du modèle de calcul qu'est la machine de Turing (en fait, une version à entrée infinie) et des questions de codage (et pour cause : les représentations de la théorie éponyme, ce sont les manières de coder les objets en suites symboliques).

En partant de la notion d'*enumerative lattice*, nous proposons une approche empruntant aux deux écoles, ayant la généralité de la seconde tout en cherchant à exprimer les choses plus en terme de structures que de machines. La structure relativement élémentaire d'*enumerative lattice* permet de définir une fois pour toutes quelques notions de constructivité, et toutes les notions habituellement considérées sur des espaces plus généraux s'y ramènent naturellement. Un des apports est donc langagier, permettant à ce qu'il nous semble de prendre de la hauteur par rapport aux questions de codage, et de manipuler les notions constructives de manière algébrique. Un autre apport de cette notion est qu'elle permet d'effectuer une fois pour toutes une construction récurrente en théorie de la calculabilité, à savoir l'énumération effective (le premier exemple que l'on rencontre dans un cursus d'informatique théorique est l'énumération effective des machines de Turing). Il semble que toutes les constructions particulières d'énumérations effectives peuvent s'y ramener, en identifiant la structure d'*enumerative lattice* appropriée. Nous utiliserons en particulier cette structure pour énumérer les tests d'aléatoirité.

2. Mesures de probabilité et calculabilité

L'espace des suites binaires, appelé espace de Cantor se prête facilement à la calculabilité, par sa structure symbolique. Par exemple, le fait qu'il est totalement déconnecté a de nombreuses conséquences : les cylindres sont des ensembles décidables, leur mesure est calculable, il peut être recouvert aussi finement que l'on veut par des ouverts disjoints deux à deux, etc. Dès que l'on passe à un espace métrique général, ces propriétés n'ont plus lieu, et posent de réelles difficultés lorsqu'on veut étendre des notions liées aux mesures.

Nous étudions la calculabilité sur les mesures de probabilité sur un espace métrique (qui doit bien sûr être calculable pour disposer de notions constructives). Nous suivons trois approches différentes d'une mesure de probabilité : (1) en mettant une métrique sur l'espace des mesures de probabilité, qui en fait un espace métrique calculable, (2) en envisageant les mesures de probabilité comme des valuations, c'est-à-dire des fonctions associant à chaque ouvert sa mesure, et (3) en identifiant une mesure à l'opérateur d'intégration associé. Ces trois approches sont montrés équivalentes. Ce travail prolonge et relie, en les reformulant, des travaux effectués précédemment.

Nous introduisons alors la version calculable des espaces de probabilités, et montrons une propriété remarquable :

Théorème 1. *Tout espace de probabilité calculable est isomorphe à l'espace de Cantor muni de la mesure uniforme.*

Ce résultat ouvre la voie au transfert des notions relatives aux mesures de probabilités et à la calculabilité (notamment l'aléatoire) de l'espace de Cantor vers n'importe quel espace de probabilités calculable. Gàcs avait proposé une construction qui donnait un isomorphisme avec Cantor dans le cas où l'espace est récursivement compact, mais sous l'hypothèse que l'on disposait d'une certaine famille de partitions calculables. Nous montrons donc qu'une telle construction est toujours possible, que l'espace soit récurisvement compact ou non. Mais la calculabilité de la mesure est un ingrédient essentiel pour faire cet isomorphisme : il existe des mesures pour lesquelles il n'y a pas d'isomorphisme avec l'espace de Cantor.

3. Aléatoire algorithmique

La section précédente fournit directement une notion d'aléatoire algorithmique sur un espace de probabilité calculable. Cependant, ceci ne s'applique que lorsque la mesure de probabilité considérée est calculable. Suivant l'idée de Levin reprise par Gács ([Gác05]), nous définissons l'aléatoire algorithmique vis-à-vis d'une mesure de probabilité quelconque (c'est-à-dire pas forcément calculable), et montrons qu'elle coincide avec la notion induite par l'isomorphisme avec l'espace de Cantor dans le cas où la mesure est calculable. L'idée est de faire jouer à la mesure le rôle d'un paramètre, ou argument supplémentaire, un test d'aléatoirité étant remplacé par un test *uniforme*, c'est-à-dire une fonction qui prend une mesure en argument, et renvoie un test pour cette mesure. Nous montrons deux résultats importants.

Théorème 2. *Soit X un espace métrique calculable.*

(1) *Si μ une mesure de probabilité sur X, tout test d'aléatoirité pour μ peut être étendu à un test uniforme.*

(2) *Il existe un test uniforme* universel.

Le premier point prend son importance lorsque l'on construit une fonction, mais que l'on utilise des propriétés particulières d'une certaine mesure (invariance, ergodicité, Bernoulli, etc) pour montrer que c'est un test pour cette mesure.

Le deuxième point étend le résultat classique de Martin-Löf sur l'espace de Cantor muni d'une mesure calculable. Nous renforçons ainsi le résultat de Gács qui n'avait lieu que sous une condition de calculabilité supplémentaire sur l'espace. La construction se fait de manière très synthétique grâce à l'utilisation du cadre des *enumerative lattices*.

4. Théorie ergodique algorithmique

La théorie ergodique est une source importante de théorèmes de forme probabiliste. En voici deux exemples classiques : étant donné un espace de probabilité (X, μ) et $T : X \to X$ un endomorphisme (c'est-à-dire qu'il préserve la mesure), (1) le théorème de récurrence de Poincaré dit que l'orbite de μ-presque tout point revient arbitrairement proche de la condition initiale (2) le théorème ergodique de Birkhoff dit que pour toute observable intégrable $f : X \to \mathbb{R}$ intégrable, la moyenne de f le long de μ-presque toute orbite converge. Nous dirons d'un point qu'il est *typique* s'il vérifie le théorème de Birkhoff pour toute observable continue bornée.

Théorème 3. *Soit X un espace métrique calculable muni d'une mesure de probabilité μ (pas forcément calculable). Alors tout point μ-aléatoire est récurrent et typique pour tout endomorphisme $T : X \to X$.*

Une partie de la preuve du théorème ergodique pour les points aléatoires est une extension d'un résultat existant sur l'espace de Cantor muni d'une mesure calculable ([V'y98]).

Nous menons ensuite une brève étude sur le problème suivant : lorsqu'un système dynamique résiste à l'analyse mathématique, l'ordinateur devient le seul recours pour avoir un aperçu du comportement du système. Une question se pose naturellement : étant donné qu'un ordinateur ne peut calculer que des objets... calculables, qui sont en quantité dénombrable, c'est-à-dire très peu dans l'espace d'états de la dynamique, les orbites affichées à l'écran sont-elles représentatives du système ? Le problème ici n'est pas lié aux erreurs arrondis : le modèle de calculabilité que l'on utilise permet de faire des calculs exacts, donc si la loi d'évolution est calculable, on peut calculer les orbites partant de points calculables, et ce sont d'authentiques orbites du système. Mais pourrait-il arriver que toutes les orbites calculables soient par exemple périodiques alors que la plupart des orbites ne le sont pas ? ou qu'aucune orbite calculable ne vérifie le théorème ergodique de Birkhoff ? Appelons orbites *pseudo-aléatoires* les orbites calculables typiques au sens du théorème ergodique : elles ne sont pas aléatoires au sens algorithmique, puisqu'elles sont calculables, mais elles ont les propriétés statistiques d'une orbite typique.

Nous apportons deux réponses positives.

Théorème 4. *Soit X un espace métrique calculable et $T : X \to X$ une transformation calculable. Si T possède une trajectoire dense, alors T possède une trajectoire dense* calculable.

Soit (X, μ) un espace de probabilité calculable. Tout endomorphisme $T : X \to X$ à mélange rapide *possède une orbite pseudo-aléatoire.*

Le mélange est une propriété plus forte que l'ergodicité : en gros, l'état du système à l'instant t est quasi-indépendant de l'état du système à un instant suffisamment lointain. Un corollaire immédiat de ce résultat est l'existence de réels calculables normaux dans toutes les bases, qui a déjà été établie en 2002, mais d'une manière très technique.

Une question reste ouverte : existe-t-il des dynamiques calculables pour lesquelles aucune orbite calculable n'est typique ?

5. Entropie et complexité des orbites

En 1998, V'yugin ([V'y98]) a montré une version du théorème de Shannon-McMillan-Breiman pour les points aléatoires, sur l'espace de Cantor muni d'une mesure calculable. Ce résultat se transfère facilement aux espaces métriques calculables via l'isomorphisme avec l'espace de Cantor, et permet de renforcer le résultat de Brudno : l'entropie algorithmique des orbites coïncide avec l'entropie globale pour les points aléatoires. Sans passer par les mesures, et par des arguments purement topologiques, nous montrons :

Théorème 5. *Soit X un espace métrique calculable compact et $T : X \to X$ une transformation calculable. La borne supérieure de la complexité algorithmique des orbites coïncide avec l'entropie topologique.*

Ces deux résultats sont comme une manifestation du principe varationnel.

Quelques références

[Bru83] A.A. Brudno. Entropy and the complexity of the trajectories of a dynamical system. *Trans. Moscow Math. Soc.*, 2 :127–151, 1983.

[Eda97] Abbas Edalat. Domains for computation in mathematics, physics and exact real arithmetic. *The Bulletin of Symbolic Logic*, 3(4) :401–452, 1997.

[Gác05] Peter Gács. Uniform test of algorithmic randomness over a general space. *Theoretical Computer Science*, 341 :91–137, 2005.

[Gal00] Stefano Galatolo. Orbit complexity by computable structures. *Nonlinearity*, 13 :1531–1546, September 2000.

[HW03] Peter Hertling and Klaus Weihrauch. Random elements in effective topological spaces with measure. *Information and Computation*, 181(1) :32–56, 2003.

[Kol65] A.N. Kolmogorov. Three approaches to the quantitive definition of information. *Problems of Information Transmission*, 1 :1–17, 1965.

[KT59] A.N. Kolmogorov and V.M. Tikhomirov. ϵ-entropy and ϵ-capacity. *Uspekhi Mat. Nauk*, 14 :3–86, 1959. English transl. AMS Translations, 17 :2 (1961), 277-364.

[LV93] Ming Li and Paul M. B. Vitanyi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer-Verlag, Berlin, 1993.

[ML66] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9(6) :602–619, 1966.

[Pet83] Karl Petersen. *Ergodic Theory*. Cambridge Univ. Press, 1983.

[V'y98] Vladimir V. V'yugin. Ergodic theorems for individual random sequences. *Theoretical Computer Science*, 207(4) :343–361, 1998.

[Wei00] Klaus Weihrauch. *Computable Analysis*. Springer, Berlin, 2000.