The Kolmogorov-random numbers in synthetic computability theory

Nils Lauermann

Advisor: Fabian Kunze

Programming Systems Lab
Saarland University

August 26, 2021

Content

Synthetic Computability Theory[1]

Constructive Type Theory: all functions $\mathbb{N} \to \mathbb{N}$ are computable

$\Rightarrow$ No external model of computation necessary

Instead we use a universal function ϕ:[2]

$$\phi : \underbrace{\mathbb{N}}_{\text{code}} \to \underbrace{\mathbb{N}}_{\text{input}} \to \underbrace{\mathbb{N}}_{\text{steps}} \to \underbrace{\mathbb{O}\mathbb{N}}_{\text{output}}$$

ϕ is a partial function: Either ϕ always returns Some x after some step count or diverges

[1] Richman 1983; Bridges and Richman 1987; Bauer 2006.
[2] Forster 2021.

Church's Thesis

All Coq functions are computable, so ϕ is universal for all (Coq) functions $\mathbb{N} \to \mathbb{N}$:

Church's Thesis[3]

$$\mathsf{CT} := \forall f : \mathbb{N} \to \mathbb{N}.\, \exists c : \mathbb{N}.\, \forall x : \mathbb{N}.\, \exists s : \mathbb{N}.\, \phi_c^s\, x = \mathsf{Some}\,(f\ x)$$

There also exists a version of CT for partial (step-indexed) functions $f : \mathbb{N} \to \mathbb{N} \to \mathbb{O}\mathbb{N}$

[3] Forster 2021.

Bijective Binary Encoding

To determine the size of a number we will use a bijective binary encoding:

- $\lceil \cdot \rceil : \mathbb{N} \to \mathbb{LB}$
- $\lfloor \cdot \rfloor : \mathbb{LB} \to \mathbb{N}$

with

- $\forall l : \mathbb{LB}.\ \lceil \lfloor l \rfloor \rceil = l$
- $\forall n : \mathbb{N}.\ \lfloor \lceil n \rceil \rfloor = n$

For simplicity we assume this encoding.

Smullyan defines the *2-adic* representation[4]:

n	0	1	2	3	4	5	6	7	8	$\cdots$
$\lceil n \rceil$	ϵ	0	1	00	01	10	11	000	001	$\cdots$

[4] Smullyan 2016.

Kolmogorov Complexity[5]

$$\mathrm{KC} \quad : \quad \underbrace{\mathbb{N}}_{\text{code}} \to \underbrace{\mathbb{N}}_{\text{number}} \to \underbrace{\mathbb{N}}_{\text{KC}} \to \mathbb{P}$$

$$\mathrm{KC}_c\ x\ k \quad :\Leftrightarrow \quad \mathrm{least}\,(\lambda k.\,\exists is.\,|\lceil i \rceil| = k \land \phi^s_c\, i = \mathrm{Some}\, x)\; k$$

Notation: $\mathrm{KC}_c\ x\ k \to p(k) \quad \sim \quad p(\mathrm{KC}_c\, x)$

Reminder: Kolmogorov complexity is uncomputable

[5] Solomonoff 1960; Kolmogorov 1965.

Universal Codes

Universal codes simulate any other code with linear overhead to the input size!

We have proven the existence of a universal code with CT for partial functions.

In the following c will be a universal code.

The Random Numbers

The Random Numbers[6]

More intuitive: incompressible numbers

Definition: random numbers

$$R_c\,(x:\mathbb{N}):\mathbb{P} := \forall is.\ \phi_n^s\ i = \mathsf{Some}\, x \rightarrow |\lceil i\rceil| \geq |\lceil x\rceil|$$

In the literature: $R_c\, x := \mathsf{KC}_c\, x \geq |\lceil x\rceil|$

These definitions are classically equivalent:

Decide termination of ϕ with excluded middle.

[6]Kolmogorov 1965.

Properties of the Non-Random Numbers

The non-random numbers $\overline{R}_c$ are

- undecidable[7] ✓
- enumerable[7] ✓
- many-one incomplete[7] ✓
- truth-table complete[8]

[7] Zvonkin and Levin 1970.
[8] Kummer 1996.

Simple Predicates[9]

Definition: simple predicate[10]

A predicate p is simple if

- p is enumerable
- $\overline{p}$ is infinite
- there is no infinite, enumerable sub-predicate of $\overline{p}$

Simple predicates are **undecidable** and **many-one incomplete**.

[9] Post 1944.
[10] Forster, Jahn, and Smolka 2021.

Non-Random Numbers: Enumerable

Definition: enumerable predicate[11]

A predicate $p : X \to \mathbb{P}$ is enumerable if $\exists f : \mathbb{N} \to \mathbb{O}X.\, \forall x.\, px \leftrightarrow \exists n.\, fn = \mathsf{Some}\, x$

Enumerator for $\overline{R}_c$:

$$\begin{array}{lll} \lambda\langle i, s\rangle. & \text{if } \phi_c^s\, i & \text{is Some } o \\ & \text{then} & \text{if } i <_{\mathbb{B}} o \text{ then Some } o \text{ else None} \\ & \text{else} & \text{None} \end{array}$$

[11] Forster, Kirst, and Smolka 2019.

Random Numbers: Infinite

Definition: infinite predicates[12]

A predicate $p : X \to \mathbb{P}$ is infinite if $\neg\exists l : \mathbb{L}X.\, \forall x : X.\, px \to x \in l$

The random numbers are unbounded:

$\forall k.\, \neg\neg\exists x.\, |\lceil x \rceil| = k \wedge R_c\, x$

There are $2^k - 1$ numbers i with $|\lceil i \rceil| < k$ and 2^k numbers o with $|\lceil o \rceil| = k$.

$\Rightarrow$ There can be at most $2^k - 1$ non-random numbers of length k.

Pigeonhole Principle: There exists an x with $|\lceil x \rceil| = k$ that must be random.

[12] Forster, Jahn, and Smolka 2021.

Random Numbers: Infinite

Definition: infinite predicates[12]

A predicate $p : X \to \mathbb{P}$ is infinite if $\neg \exists l : \mathbb{L}X.\, \forall x : X.\, px \to x \in l$

The random numbers are infinite:

Given a list l that contains all random numbers.

By the unboundedness there exists a random number x with $|\lceil x \rceil| = \max_{y \in l}(|\lceil y \rceil| + 1)$.

Contradiction!

$\Rightarrow$ The random numbers must be infinite!

[12] Forster, Jahn, and Smolka 2021.

Random Numbers: No infinite, enumerable sub-predicate

Reminder: Uncomputability of Kolmogorov complexity

Berry Paradox[13]: The smallest number x with $KC_c(x) > m$

Almost identical proof:

The smallest number x that satisfies the sub-predicate and $|\lceil x \rceil| > m$.

Remark: Similarly to the uncomputability proof, Markov's principle is used.

Assuming Markov's principle, the non-random numbers are simple and hence undecidable und many-one incomplete!

[13] Russell 1908.

Lower Bound for Random Numbers

A lower bound for the count of random numbers[14]

Let c be universal:
There exists a constant d so that at least $\frac{1}{d}$ of the numbers of every length k are random!

- Similar core idea as in Kummer's truth-table completeness proof
- Currently uses excluded middle

[14] Kummer 1996.

Conclusion

Conclusion

Working in synthetic computability is extremely natural and convenient!

Contributions

- To the best of our knowledge, the first formalization of Kolmogorov complexity
  - in Coq
  - in synthetic computability theory
- Undecidability of Kolmogorov complexity
- Simpleness of the non-random numbers
- Lower bound for the count of random numbers
- First steps towards a truth-table completeness proof of the non-random numbers in Coq

Related Work

Catt and Norrish formalized KC in HOL4:

- Classical logic
- With λ-*calculus* and *general recursive functions* as model of computation
- Focus on inequalities involving Kolmogorov complexity

Future Work

- Uncomputability/Simpleness: Investigate an elimination of Markov's principle
- truth-table completeness of the non-random numbers

Thank you!

References i

Bauer, Andrej (2006). "First Steps in Synthetic Computability Theory". In: *Electronic Notes in Theoretical Computer Science* 155, pp. 5–31. DOI: 10.1016/j.entcs.2005.11.049.

Bridges, Douglas and Fred Richman (1987). *Varieties of constructive mathematics*. Vol. 97. London Mathematical Society Lecture Note Series. Cambridge University Press. DOI: 10.1017/CBO9780511565663.

Catt, Elliot and Michael Norrish (2021). "On the formalisation of Kolmogorov complexity". In: *Proceedings of the 10th ACM SIGPLAN International Conference on Certified Programs and Proofs*. ACM, pp. 291–299. DOI: 10.1145/3437992.3439921.

- Forster, Yannick (2021). "Computability in Constructive Type Theory". PhD thesis. Saarland University. URL: https://ps.uni-saarland.de/~forster/thesis.php.
- Forster, Yannick, Felix Jahn, and Gert Smolka (2021). "A Constructive and Synthetic Theory of Reducibility". Unpublished draft.
- Forster, Yannick, Dominik Kirst, and Gert Smolka (2019). "On synthetic undecidability in Coq, with an application to the Entscheidungsproblem". In: *Proceedings of the 8th ACM SIGPLAN International Conference on Certified Programs and Proofs*. ACM, pp. 38–51. DOI: 10.1145/3293880.3294091.
- Kolmogorov, Andrei N. (1965). "Three approaches to the quantitative definition of information". In: *Problems of information transmission* 1.1, pp. 3–11.

References iii

- Kummer, Martin (1996). "On the complexity of random strings". In: *Annual Symposium on Theoretical Aspects of Computer Science*. Springer, pp. 25–36. DOI: 10.1007/3-540-60922-9_3.
- Post, Emil L. (1944). "Recursively enumerable sets of positive integers and their decision problems". In: *bulletin of the American Mathematical Society* 50.5, pp. 284–316. DOI: 10.1090/S0002-9904-1944-08111-1.
- Richman, Fred (1983). "Church's thesis without tears". In: *The Journal of symbolic logic* 48.3, pp. 797–803. DOI: 10.2307/2273473.
- Russell, Bertrand (1908). "Mathematical Logic as Based on the Theory of Types". In: *American Journal of Mathematics* 30.3, pp. 222–262. DOI: 10.2307/2369948.
- Smullyan, Raymond M. (2016). *Theory of Formal Systems. (AM-47)*. Vol. 47. Princeton University Press. DOI: 10.1515/9781400882007.

- Solomonoff, Ray J. (1960). "A preliminary report on a general theory of inductive inference". In: Citeseer.
- Zvonkin, Alexander K. and Leonid A. Levin (1970). "The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms". In: *Russian Mathematical Surveys* 25.6, pp. 83–124. DOI: 10.1070/rm1970v025n06abeh001269.

LOC

Content	Spec	Proof
Preliminaries	34	167
List Facts	78	557
Binary Encoding	65	453
Kolmogorov Complexity and Facts for KC	21	157
The Uncomputability of KC	14	162
Simpleness of the Non-Random Numbers	47	365
Lower Bound for the Random Numbers	64	838
Total	323	2699

Lower Bound for Random Numbers

Making outputs non-random

Reminder: Invariance Theorem[15]

$$\mathrm{univ}\, c \to \forall c'.\, \exists d.\, \forall x.\, \mathrm{KC}_c\, x \le \mathrm{KC}_{c'}\, x + d$$

Goal: Make a number x, with $|\lceil x \rceil| = k$, non-random with regard to c:

Idea: Construct c' with $\mathrm{KC}_{c'}\, x < k - d$

Problem: We cannot know d during the definition of c'

Solution: Incorporate d into input for c'.

[15] Kolmogorov 1965.

The Lower Bound

There exists a function $f : \mathbb{N} \to \mathbb{N}$ so that we can ensure the non-randomness of $2^{n-f(d)}$ numbers of length n.

Which numbers will we force non-random?

For all $x < 2^{n-f(d)}$: Try to enumerate $2^n - x$ non-random numbers and make a number that was not enumerated random!

There must be at least $2^{n-f(d)}$ random numbers of length n

Proof by Contradiction: Assume there are less than $2^{n-f(d)}$ random numbers. Then there are more than $2^n - 2^{n-f(d)}$ non-random numbers.

Some x will enumerate all non-random numbers. Hence the number that is made non-random, is random. Contradiction!