Annals of Pure and Applied Logic 165 (2014) 1380–1402

Contents lists available at ScienceDirect

Annals of Pure and Applied Logic

www.elsevier.com/locate/apal

The axiomatic power of Kolmogorov complexity

Laurent Bienvenu [a,*], Andrei Romashchenko [b], Alexander Shen [b], Antoine Taveneaux [c], Stijn Vermeeren [d]

[a] *Laboratoire J.-V. Poncelet, CNRS, Moscow, Russian Federation*
[b] *LIRMM, CNRS & Université Montpellier 2, France*
[c] *LIAFA, CNRS & Université Paris 7, France*
[d] *University of Leeds, United Kingdom*

ARTICLE INFO

Article history:
Available online 13 May 2014

MSC:
primary 03D32, 03F99, 68Q30
secondary 03F20

Keywords:
Probabilistic proofs
Kolmogorov complexity
Martin-Löf randomness

ABSTRACT

The famous Gödel incompleteness theorem states that for every consistent, recursive, and sufficiently rich formal theory T there exist true statements that are unprovable in T. Such statements would be natural candidates for being added as axioms, but how can we obtain them? One classical (and well studied) approach is to add to some theory T an axiom that claims the consistency of T. In this paper we discuss another approach motivated by Chaitin's version of Gödel's theorem where axioms claiming the randomness (or incompressibility) of some strings are probabilistically added, and show that it is not really useful, in the sense that this does not help us prove new interesting theorems. This result answers a question recently asked by Lipton. The situation changes if we take into account the size of the proofs: randomly chosen axioms may help making proofs much shorter (unless NP = PSPACE).

We then study the axiomatic power of the statements of type "the Kolmogorov complexity of x exceeds n" (where x is some string, and n is some integer) in general. They are Π_1 (universally quantified) statements of Peano arithmetic. We show that by adding all true statements of this type, we obtain a theory that proves all true Π_1-statements, and also provide a more detailed classification. In particular, as Theorem 7 shows, to derive all true Π_1-statements it is enough to add one statement of this type for each n (or even for infinitely many n) if strings are chosen in a special way. On the other hand, one may add statements of this type for most x of length n (for every n) and still obtain a weak theory. We also study other logical questions related to "random axioms".

Finally, we consider a theory that claims Martin-Löf randomness of a given *infinite* binary sequence. This claim can be formalized in different ways. We show that different formalizations are closely related but not equivalent, and study their properties.

[*] Corresponding author.

http://dx.doi.org/10.1016/j.apal.2014.04.009

1. Introduction

We assume that the reader is familiar with the notion of Kolmogorov complexity and Martin-Löf randomness (see [14,16,7] for background information about Kolmogorov complexity and related topics), but since for our purposes this notion needs to be expressed in formal arithmetic, we recall some basic definitions. The Kolmogorov complexity $\mathrm{C}(x)$ of a binary string x is defined as the minimal length of a program (without input) that outputs x and terminates. This definition depends on a programming language, and one should choose one that makes complexity minimal up to $\mathrm{O}(1)$ additive term. Technically, there exist different versions of Kolmogorov complexity. Prefix complexity $\mathrm{K}(x)$ assumes that programs are self-delimiting. We consider plain complexity $\mathrm{C}(x)$ where no such assumptions are made; any partial function D can be used as an "interpreter" of a programming language, so $D(p)$ is considered as an output of program p, and $\mathrm{C}_D(x)$ is defined as the minimal length of p such that $D(p) = x$. Then some optimal D is fixed (such that C_D is minimal up to $\mathrm{O}(1)$ additive term), and $\mathrm{C}_D(x)$ is called (plain Kolmogorov) complexity of x and denoted $\mathrm{C}(x)$. Most strings of length n have complexity close to n. More precisely, the fraction of n-bit strings that have complexity less than $n - c$, is at most 2^{-c}. In particular, there exist strings of arbitrary high complexity.

However, as G. Chaitin pointed out in [5], the situation changes if we look for strings of *provably* high complexity. More precisely, we are looking for strings x and numbers n such that the statement "$\mathrm{C}(x) > n$" (properly formalized in arithmetic) is provable in Peano arithmetic PA. Chaitin showed that there is a constant c such that *no* statement "$\mathrm{C}(x) > n$" is provable in PA for $n > c$. Chaitin's argument is a version of Berry's paradox: Assume that for every integer k we can find some string x such that "$\mathrm{C}(x) > k$" is provable; let x_k be the first string with this property in the order of enumeration of all proofs; this definition provides a program of size $\mathrm{O}(\log k)$ that generates x_k, which is impossible for large k since "$\mathrm{C}(x_k) > k$" is provable in PA and therefore true (in the standard model).[1]

This leads to a natural idea. Toss a coin n times to obtain a string x of length n, and consider the statement "$\mathrm{C}(x) \geq n - 1000$". This statement is true unless we are extremely unlucky. The probability of being unlucky is less than 2^{-1000}. In natural sciences we are accustomed to identify this with impossibility. So we can add this statement and be almost sure that it is true; if n is large enough, we get a true non-provable statement and could use it as a new axiom. We can even repeat this procedure several times: if the number of iterations m is not astronomically large, 2^{-1000} m is still astronomically small.

Now the question: *Can we obtain a richer theory in this way and get some interesting consequences, still being practically sure that they are true*? This is in substance what Lipton asked in [13] (Shen [17] also discussed similar issues). In the next section, we will answer this question and show:

- yes, this is a safe way of enriching PA (Theorem 1);
- yes, we can get a stronger theory this way (Chaitin's theorem), but
- no, we cannot prove anything interesting this way (Theorem 3).

So the answer to the main question is negative; however, as we show in Section 2.3 (Theorem 4), these "random axioms" do give some advantages: while they cannot help us to prove new interesting statements, they can significantly shorten some proofs (unless PSPACE = NP).

[1] Another proof of the same result shows that Kolmogorov complexity is actually not very essential here. By a standard fixed-point argument one can construct a program p (without input) such that for every program q (without input) the assumption "q is equivalent to p" (i.e., q produces the same output as p if p terminates, and q does not terminate if p does not terminate) is consistent with PA. If p has length k, for every x we may assume without contradiction that p produces x, so one cannot prove that $\mathrm{C}(x)$ exceeds k.