THE MONTE CARLO HYPOTHESIS

Aidan Rocke
aidanrocke@gmail.com

March 28, 2021

ABSTRACT

The Monte Carlo hypothesis implies a non-trivial interpretation of the prime number theorem. The Prime Number Theorem says how the prime numbers are distributed but not why. On the other hand, an information-theoretic analysis of the Prime Number Theorem indicates that they are distributed in this way because prime encodings are algorithmically random and the prime numbers have a maximum entropy distribution.

1 The Challenge

1. The challenge is to explore the hypothesis that prime-encodings are algorithmically random and that prime numbers have a maximum entropy distribution [1].

2. Conventional mathematical wisdom, at present, suggests that this hypothesis might be false. There might be 'conspiracies' among arbitrarily large subsequences of prime encodings, i.e. predictable behaviour, which may be exploited by a machine learning algorithm [2].

It is not possible to prove that a particular object is incompressible within algorithmic information theory so the best we can do is perform rigorous experimental analysis using machine learning methods. Hence this challenge.

2 The Monte Carlo Hypothesis

The scientific basis for this challenge may be expressed as follows. If $u_i = i$ defines an integer sequence then a prime encoding $X_N = \{x_i\}_{i=1}^{N}$ is given by $x_i = 1$ if u_i is prime and $x_i = 0$ otherwise. Under reasonable assumptions, an application of the Shannon source coding theorem allows us to derive the asymptotic incompressibility relation:

$$\mathbb{E}[K(X_N)] \sim \pi(N) \cdot \ln(N) \sim N \tag{1}$$

where $\pi(N)$ is the number of prime numbers less than N and $\mathbb{E}[K(X_N)]$ is the Expected Kolmogorov Complexity of X_N. More details are provided in [1] and a notion of asymptotic incompressibility that is consistent with Schnorr's martingale characterisation of algorithmic randomness is provided in the next section.

3 Algorithmic randomness and effective betting strategies

Given the binary sequence $X_N = \{x_i\}_{i=1}^{N}$, we say that X_N is asymptotically incompressible if given the subsequence X_k and $N >> k$ on average we would not profit by gambling on the $N - k$ terms in X_N based on the partial knowledge provided by X_k. If X_N satisfies these assumptions then we may state that for large N:

$$\mathbb{E}[K(X_N)] \sim N \tag{2}$$

Moreover, if there exists a deterministic function f such that $x_{n+1} = f \circ x_n$ an effective betting strategy is reducible to approximating f using a machine learning algorithm provided with the dataset X_k.

4 Challenges in Experimental Mathematics

The maximum entropy distribution of the prime numbers and the algorithmic randomness of prime encodings are naturally divided into two separate challenges. A couple getting started guides for each challenge are provided on Github: `https://github.com/AidanRocke/Monte-Carlo-Hypothesis`.

4.1 Probabilistic Primality Testing

The hypothesis that the prime numbers have a maximum entropy distribution implies that the accuracy of a prime formula with bounded algorithmic information content eventually converges to zero. Therefore, the true positive rate of a deterministic primality test with bounded algorithmic information content(ex. a machine learning model) would eventually converge to zero.

The appropriate data for such an experimental verification consists of binary encodings of the integers which for all prime numbers less than N requires $\log_2 N$ bits for representation.

4.2 Weighted Next-bit test

The hypothesis that prime encodings are algorithmically random implies that a sequence of prime encodings pass the *weighted* next bit test. To be precise, given a sequence of N consecutive prime encodings the probability that a machine learning model M accurately predicts the $N + 1$th term converges to $p \in [0, \frac{1}{2}]$. Therefore, the true-positive rate of M converges to $p \in [0, \frac{1}{2}]$.

Emphasis is made upon the fact that the next-bit test is weighted as the prime numbers are relatively sparse so the dataset is unbalanced and therefore the appropriate measure is the true-positive rate for predicting the location of the next prime number.

5 Scientific and Technological applications

This analysis of the distribution of the prime numbers was originally motivated by three questions:

1. What effective methods may be used to communicate with civilisations within the Turing limit that are advanced enough to do number theory? For concreteness, how might one determine whether a distant civilisation communicated partial knowledge concerning the distribution of prime numbers?

For this we would need signal processing algorithms that are invariant to all possible mathematical representations of the prime numbers.

2. Assuming that the Physical Church-Turing thesis is valid, what mathematical signatures might provide strong evidence for the simulation hypothesis?

It is worth noting that every Universal Turing Machine contains an encoding of arithmetic and therefore an encoding of the natural numbers. It follows that within a complete theory of algorithmic information civilisations within the Turing limit would have at best a probabilistic understanding of the distribution of prime numbers.

3. The general problem of rare-event modelling which has a variety of applications, may be formulated as a signal compression problem. Given that modelling the distribution of prime numbers has proven to be one of the hardest signal compression problems, it is invaluable to carefully assess the effectiveness of state-of-the-art machine learning models on this canonical problem.

6 Prizes

After careful deliberation, it was decided that no material prize could be offered for this challenge. The greatest possible recompense is an understanding of the platonic origins of the prime numbers.

Such an understanding contains the answer to questions which we do not yet know how to ask.

References

[1] Aidan Rocke (https://mathoverflow.net/users/56328/aidan-rocke), information-theoretic derivation of the prime number theorem, URL (version: 2021-02-20): https://mathoverflow.net/q/384109

[2] Terence Tao. Structure and randomness in the prime numbers: A small selection of results in number theory. Slides. 2007.

[3] Marcus Hutter. Algorithmic information theory. Scholarpedia. 2007.

[4] Schnorr, C. P. (1971). "A unified approach to the definition of a random sequence". Mathematical Systems Theory.

[5] Peter D. Grünwald. The Minimum Description Length Principle . MIT Press. 2007.

[6] David Deutsch. Quantum theory, the Church–Turing principle and the universal quantum computer. 1985.

[7] Don Zagier. Newman's short proof of the Prime Number Theorem. The American Mathematical Monthly, Vol. 104, No. 8 (Oct., 1997), pp. 705-708

[8] John A. Wheeler, 1990, "Information, physics, quantum: The search for links" in W. Zurek (ed.) Complexity, Entropy, and the Physics of Information. Redwood City, CA: Addison-Wesley.

[9] Guillermo Valle Pérez, Chico Camargo, Ard Louis. Deep Learning generalizes because the parameter-function map is biased towards simple functions. 2019.

[10] Aidan Rocke (https://cstheory.stackexchange.com/users/47594/aidan-rocke), Understanding the Physical Church-Turing thesis and its implications, URL (version: 2021-02-22): https://cstheory.stackexchange.com/q/48450

[11] Igor V. Volovich. Number Theory as the Ultimate Physical Theory. Pleiades Publishing. 2010.