Integer Complexity: Experimental and Analytical Results II

Juris Čerņenoks[1] **Jānis Iraids**[1] Mārtiņš Opmanis[2]
Rihards Opmanis[2] Kārlis Podnieks[1]

[1]University of Latvia

[2]Institute of Mathematics and Computer Science, University of Latvia

DCFS 2015, University of Waterloo, Friday 26th June, 2015

Definition of Integer Complexity

Integer complexity

Integer complexity of a positive integer n, denoted by $\|n\|$, is the least amount of 1's in an arithmetic expression for n consisting of 1's, $+$, $\cdot$ and brackets.

For example,

$$\|1\| = 1$$
$$\|2\| = 2; 2 = 1+1$$
$$\|3\| = 3; 3 = 1+1+1$$
$$\|6\| = 5; 6 = (1+1)\cdot(1+1+1)$$
$$\|8\| = 6; 8 = (1+1)\cdot(1+1)\cdot(1+1)$$
$$\|11\| = 8; 11 = (1+1+1)\cdot(1+1+1)+1+1$$

http://oeis.org/A005245

Lower and Upper Bounds

Theorem

$$\|n\| \in \Theta(\log n)$$

Sketch of proof.

1. $\|n\| \leq 3\log_2 n$ – Horner's rule
   Expand n in binary: $n = \overline{a_k a_{k-1} \cdots a_1 a_0}$.
   Express as
   $$a_0 + (1+1)\cdot(a_1 + (1+1)\cdot\ldots(a_{k-1} + (1+1)\cdot a_k)\ldots).$$
2. $\|n\| \geq 3\log_3 n$
   Idea: denote by $E(k)$ the **largest** number having complexity k.

Lower and Upper Bounds

Theorem

$$\|n\| \in \Theta(\log n)$$

Sketch of proof.

We will show that

$$E(3k+2) = 2 \cdot 3^k;$$
$$E(3k+3) = 3 \cdot 3^k;$$
$$E(3k+4) = 4 \cdot 3^k.$$

□

Largest Number of Complexity k

Theorem

For all $k \geq 0$:

$$E(3k+2) = 2 \cdot 3^k;\ E(3k+3) = 3 \cdot 3^k;\ E(3k+4) = 4 \cdot 3^k.$$

Proof (by H. Altman).

The value of an expression **does not decrease** if we:

- Replace all $x \cdot 1$ by $x + 1$;
- Replace all $x \cdot y + 1$ by $x \cdot (y+1)$;
- Replace all $x \cdot y + u \cdot v$ by $x \cdot y \cdot u \cdot v$;
- If $x = 1 + 1 + \ldots + 1 > 3$, split it into product of $(1+1)$'s and $(1+1+1)$'s;
- Replace all $(1+1) \cdot (1+1) \cdot (1+1)$ by $(1+1+1) \cdot (1+1+1)$.

□

Complexity of Powers

From $E(3k) = 3^k$ we arrive at

$$\left\|3^k\right\| = 3k.$$

What about powers of other numbers $\left\|n^k\right\|$?
There exist n with $\left\|n^k\right\| < k \cdot \|n\|$:

$$\begin{aligned} \|5\| &= 5 \\ \left\|5^2\right\| &= 10 \\ &\dots \\ \left\|5^5\right\| &= 25 \\ \left\|5^6\right\| &= 29; 5^6 = (3^3 \cdot 2^3 + 1) \cdot 3^2 \cdot 2^3 + 1 \end{aligned}$$

Complexity of 2^a

Richard K. Guy, "Unsolved Problems in Number Theory", problem F26

Is $\|2^a3^b\| = 2a + 3b$ for all $(a, b) \neq (0, 0)$?
In particular, is $\|2^a\| = 2a$ for all a? [Attributed to Selfridge, Hypothesis H1]

- Having computed $\|n\|$ for n up to 10^{12} hypothesis H1 holds for all $a \leq 39$ [2010].
- Recently Harry Altman showed H1 holds for all (a, b) with $a \leq 48$ (See the PhD thesis of Altman "Integer Complexity, Addition Chains, and Well-Ordering" for excellent introduction to integer complexity.)

Logarithmic Complexity

Let the **logarithmic complexity** of n be denoted by $\|n\|_{\log} = \frac{\|n\|}{\log_3 n}$.

- $3 \leq \|n\|_{\log} \leq 3\log_2 3 \approx 4.755$

Richard K. Guy, "Unsolved Problems in Number Theory", problem F26

As $n \to \infty$ does $\|n\|_{\log} \to 3$? [Hypothesis H2]

- For all n up to 10^{12}:

$$\|n\|_{\log} \leq \|1439\|_{\log} \approx 3.928.$$

- In 2014 Arias de Reyna and van de Lune showed that for most n:

$$\|n\|_{\log} < 3.635.$$

Distribution of Logarithmic Complexity [1]

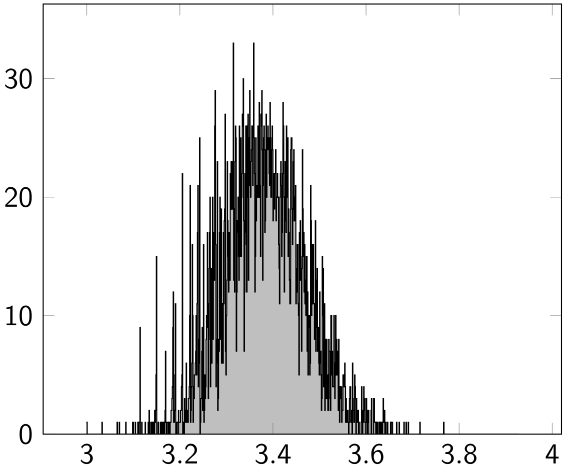

Figure : Distribution of logarithmic complexity of numbers with $\|n\| = 30$

Distribution of Logarithmic Complexity [2]

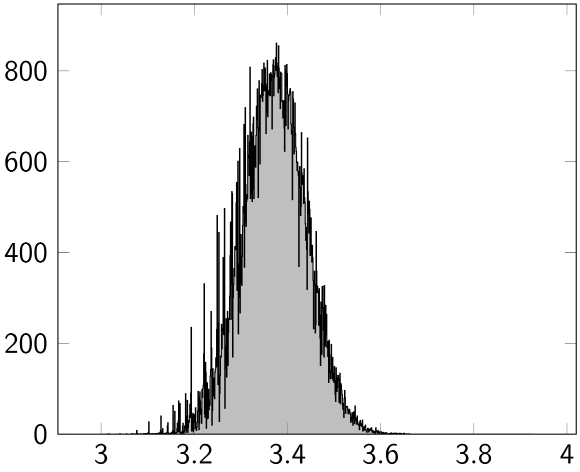

Figure : Distribution of logarithmic complexity of numbers with $\|n\| = 40$

Distribution of Logarithmic Complexity [3]

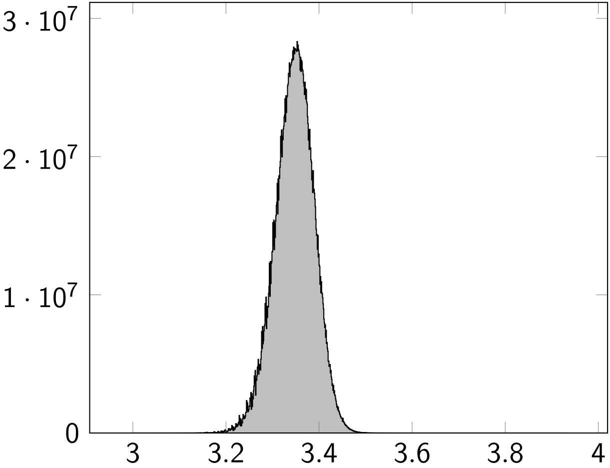

Figure : Distribution of logarithmic complexity of numbers with $\|n\| = 70$

Relation of the Open Problems

Richard K. Guy, "Unsolved Problems in Number Theory", problem F26

- Is $\|2^a\| = 2a$ for all a? [Hypothesis H1]
- As $n \to \infty$ does $\|n\|_{\log} \to 3$? [Hypothesis H2]

- $H1 \implies \neg H2$, because

$$\|2^a\|_{\log} = \frac{2a}{\log_3 2^a} \approx 3.170;$$

  hence H2 *should* be easier to settle.
- We have not succeeded to prove or disprove either of them.

Base-3 Representations of 2^n

Observation

$$\|8\| = 6; 8 = (1+1)(1+1)(1+1)$$
$$\|9\| = 6; 9 = (1+1+1)(1+1+1)$$

Base-3 representation of powers of 2:

$$(2)_3 = 2$$
$$(2^2)_3 = 11$$
$$(2^3)_3 = 22$$
$$(2^{10})_3 = 1101221$$
$$(2^{30})_3 = 2202211102201212201$$
$$(2^{50})_3 = 12110122110222110100112122112211$$

The digits seem "random, uniformly distributed".

Pseudorandomness of Powers

Let $S_q(p^n)$ denote the **sum of digits** of p^n in base q. If the digits were to be **independent, uniformly distributed** random variables then the pseudo expectation would be:

$$E_n \approx n \log_q p \cdot \frac{q-1}{2}$$

and pseudo variance

$$V_n \approx n \log_q p \cdot \frac{q^2-1}{12};$$

and the corresponding normed and centered variable $s_q(p^n)$ should behave as the **standard normal distribution**.
We can try to *verify* this experimentally...

Distribution of Normalized Digit Sums

The results for n up to 10^5:

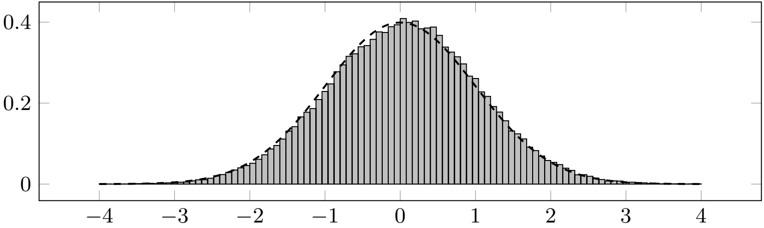

Figure : Histogram of the centered and normed variable $s_3(2^n)$

Related Theoretical Results

Conjecture by Paul Erdős

For $n > 8$, the base-3 representation of 2^n contains digit "2".

Corollary of a theorem by C. L. Stewart

There exists a constant $C_{p,q} > 0$ such that:

$$S_q(p^n) > \frac{\log n}{\log \log n + C_{p,q}} - 1.$$

Our result

If H1 holds, i.e., if indeed $\|2^n\| = 2n$, then

$$S_3(2^n) > 0.107n.$$

Does this mean proving H1 is very difficult?

H1 Implies Linear Sum of Digits

Theorem (Čerņenoks et al.)

If, for a prime p, $\exists \epsilon > 0 \forall n > 0 : \|p^n\|_{\log} \geq 3 + \epsilon$, *then*

$$S_3(p^n) \geq \epsilon n \log_3 p.$$

Proof.

Write p^n in base q: $a_m a_{m-1} \cdots a_0$.
Using Horner's rule we obtain an arithmetic expression for p^n:

$$\|p^n\| \leq qm + S_q(p^n).$$

Since $m \leq \log_q p^n$,

$$\|p^n\| \leq q \log_q p^n + S_q(p^n).$$

H1 Implies Linear Sum of Digits

Theorem (Čerņenoks et al.)

If, for a prime p, $\exists \epsilon > 0 \forall n > 0 : \|p^n\|_{\log} \geq 3 + \epsilon$, then

$$S_3(p^n) \geq \epsilon n \log_3 p.$$

Proof.

$$\|p^n\| \leq q \log_q p^n + S_q(p^n).$$

When $q = 3$:

$$\begin{aligned} S_3(p^n) &\geq \|p^n\|_{\log} \log_3 p^n - 3 \log_3 p^n \geq \\ &\geq (3 + \epsilon) n \log_3 p - 3n \log_3 p = \\ &= \epsilon n \log_3 p. \end{aligned}$$

□

Definition

Integer complexity in basis $\{1, +, \cdot, -\}$

Integer complexity (in basis $\{1, +, \cdot, -\}$) of a positive integer n, denoted by $\|n\|_-$, is the least amount of 1's in an arithmetic expression for n consisting of 1's, $+$, $\cdot$, $-$ and brackets.
The corresponding logarithmic complexity is denoted by $\|n\|_{-\log}$.

Having computed $\|n\|_-$ for n up to $2 \cdot 10^{11}$ we present our observations.

Experimental Results in Basis $\{1, +, \cdot, -\}$ [1]

Smallest number with $\|n\|_- < \|n\|$:

$$\|23\|_- = 10; \|23\| = 11;$$

$$23 = 2^3 \cdot 3 - 1 = 2^2 \cdot 5 + 2.$$

There are numbers for which subtraction of 6 is necessary:

$$\|n\|_- = 75; n = 55\,659\,409\,816 = (2^4 \cdot 3^3 - 1)(3^{17} - 1) - 2 \cdot 3;$$

$$\|n\|_- = 77; n = 111\,534\,056\,696 = (2^5 \cdot 3^4 - 1)(3^{16} + 1) - 2 \cdot 3;$$

$$\|n\|_- = 78; n = 167\,494\,790\,108 = (2^4 \cdot 3^4 + 1)(3^{17} - 1) - 2 \cdot 3.$$

Experimental Results in Basis $\{1, +, \cdot, -\}$ [2]

"Worst" numbers

Let

- $e(n)$ denote $\min\{k \mid \|k\| = n\}$ and
- $e_-(n)$ denote $\min\{k \mid \|k\|_- = n\}$.

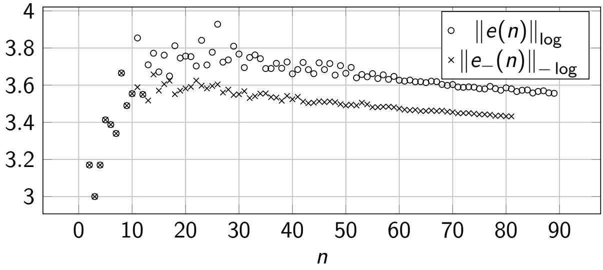

Figure : Logarithmic complexities of the numbers $e(n)$ and $e_-(n)$

Upper Bound

Theorem (Čerņenoks et al.)

$$\|n\|_{-\log} \leq 3.679 + \frac{5.890}{\log_3 n}$$

Sketch of proof.

- $n = 6k$; write n as $3 \cdot 2 \cdot k$;
- $n = 6k + 1$; write n as $3 \cdot 2 \cdot k + 1$;
- $n = 6k + 2$; write n as $2 \cdot (3 \cdot k + 1)$;
- $n = 6k + 3$; write n as $3 \cdot (2 \cdot k + 1)$;
- $n = 6k + 4$; write n as $2 \cdot (3 \cdot (k + 1) - 1)$;
- $n = 6k + 5$; write n as $3 \cdot 2 \cdot (k + 1) - 1$;

Upper Bound

Theorem (Čerņenoks et al.)

$$\|n\|_{-\log} \leq 3.679 + \frac{5.890}{\log_3 n}$$

Sketch of proof.

- Apply the rules iteratively
- Each iteration uses at most 6 ones
- Each iteration reduces the problem from n to some $k \leq \frac{n}{6} + \frac{1}{3}$
- After m applications we arrive at a number

$$k < \frac{n}{6^m} + \frac{2}{5}$$

□

Our Results

Digit sum problem

Hypothesis $\|2^n\| = 2n$ implies a **linear** lower bound on the sum of digits:

$$S_3(2^n) > 0.107n.$$

Upper bound in base $\{1, +, \cdot, -\}$

$$\limsup_{n\to\infty} \|n\|_{-\log} \leq 3.679$$

Open Problems

H1

Is $\|2^n\| = 2n$?

Spectrum of $\|n\|_{\log}$

1. Is $\limsup_{n\to\infty} \|n\|_{\log} = 3$? (H2)
2. Can we at least show

$$\limsup_{n\to\infty} \|n\|_{\log} < 3\log_2 3?$$

Digit sum

Can we improve the sum of digits bound

$$S_p(q^n) \geq \frac{\log n}{\log\log n + C_{p,q}} - 1?$$

Questions?