LARGEST SMITH NUMBER
(PAPER IN PROGRESS)

MARLON TRIFUNOVIC

Abstract. We find large Smith numbers by explicitly calculating digit sums through several methods relying on computer programs. This paper explicitly constructs a Smith number with 1 094 654 215 464 digits, exceeding previous record of 32 066 910 digits.

Contents

1. Introduction 1
2. Notation and Basic Facts 1
3. An update on Costello 2002 1
4. Best by adding coefficients 3
5. Implementation & Afterword 4
References 4

1. Introduction

A Smith number is defined by A. Wilanksy as "a composite number the sum of whose digits is the sum of all digits of all its prime factors"[5].

2. Notation and Basic Facts

The following notation and basic facts are taken from Patrick Costello [3]. For any positive integer n, let $S(n)$ denote the sum of the digits of n. For any positive integer n, let $S_p(n)$ denote the sum of digits of the prime factorization of n. For example, $S(12) = 1 + 2 = 3$ and $S_p(12) = S_p(2 \cdot 2 \cdot 3) = 2 + 2 + 3 = 7$.

3. An update on Costello 2002

Patrick Costello was able to construct a $32,066,910$ digit Smith number by using the known prime repunit R_{1031} and Chris Caldwell's large palindromic prime $M = 10^{28572} + 8 \cdot 10^{14286} + 1$. I will briefly go through the similar steps as Costello with more recently verified primes to construct a new largest Smith number.

Firstly, two facts are necessary

Fact 1 (Lewis [4])**.** If you multiply $9R_n$ by any natural number less than $9R_n$, then the digit sum is $9n$, i.e., $S(M \cdot 9R_n) = 9M = S(9R_n)$ when $M < 9R_n$.

Fact 2 (Wayland, Oltikar [6])**.** If $S(u) > S_p(u)$ and $S(u) = S_p(u) \pmod 7$, then $10^k \cdot u$ is a Smith number, where $k = \frac{S(u)-S_p(u)}{7}$.

Date: July 4, 2022.

Chris Caldwell's list of large proven primes[7] lists the prime $M = 3\cdot 10^{665829}+1$. It additionally lists the large prime repunit R_{49081} discovered and proven by Paul Underwood[8]. Notice that M is not palindromic and requires different method for bounding coefficients than Costello's 2002 paper.

For a power t, the term M^t can be represented as a sum of coefficients multiplied by powers of 10^{665829}.

$$M^t = \sum_{k=0}^{t} c_k 10^{665829k}, \quad c_{t,k} = \binom{t}{k} 3^k$$

When t is fixed, I will drop the t for ease of notation: $c_k = c_{t,k}$.

Theorem 3.1. *Fix $t > 0$ and let $k(t) = \lceil \frac{3t-1}{4} \rceil$, then $c_k \leq c_{k(t)}$ for $0 \leq k \leq t$.*

Proof. Define the ratio of coefficients $b_k = \frac{c_{k+1}}{c_k}$, then

$$\begin{aligned} b_k &= \frac{c_{k+1}}{c_k} \\ &= \frac{\binom{t}{k+1}3^{k+1}}{\binom{t}{k}3^k} \\ &= \frac{3\frac{1}{(k+1)!(t-(k+1))!}}{\frac{1}{k!(t-k)!}} \\ b_k &= \frac{3(t-k)}{k+1} \end{aligned}$$

Notice that the b_k ratios are decreasing with $b_{t-1} < 1 < b_0$. Then $k(t) = \lceil \frac{3t-1}{4} \rceil$ is the minimal integer so that $b_{k(t)} \leq 1$. Since these b_k are ratios of c_k, $c_{k(t)}$ is the largest coefficient. □

Theorem 3.2. *Fix $t \leq 81525$ and suppose $N = 9R_{49081}M^t$, then for all $0 \leq k \leq t$, $c_{t,k} < 9R_{49081}$ and $9R_{49081}c_{t,k} < 10^{665829}$.*

Proof. Fix t, k as in the theorem, then by Theorem 3.1,

$$\begin{aligned} c_{t,k} &\leq c_{t,k(t)} \\ &= \binom{t}{k(t)} 3^{k(t)} \end{aligned}$$

It's clear that the function $t \mapsto \binom{t}{k(t)}3^{k(t)}$ is an increasing function. By explicit calculation,

$$c_{t,k(t)} \leq c_{81525,k(81525)} < 9R_{49081} < c_{81526,k(81526)}$$

Since $(9R_{49081})^2 < 10^{665829}$, then $9R_{49081}c_{t,k} < 10^{665829}$ as well.

□

Now suppose $N = 9R_{49081}M^t$ for a power $t \leq 81525$. We know each coeffient $c_k < 9R_{49081}$ and $9R_{49081}c_k < 10^{665829}$. The latter constraint means the digit sum of N is the sum of the digit sums of $9R_{49081}c_k$. The first constraint allows us to apply fact 1 to prove the digit sum of $9R_{49081}c_k$ is $9 \cdot 49081$. Since k varies from $0 \leq k \leq t$, then

$$S(N) = (t+1) \cdot 9 \cdot 49081$$

The prime factorization of N is simply $3 \cdot 3 \cdot R_{49081} \cdot M^t$. Keeping in mind $S(M) = 4$, then

$$S_p(N) = 3 + 3 + 49081 + 4t$$

Note that $S(N) > S_p(N)$ and

$$\begin{aligned} S(N) - S_p(N) &= (t+1) \cdot 9 \cdot 49081 - (3 + 3 + 49081 + 4t) \\ &= 441725t + 392642 \\ &= 4t + 5 \pmod 7 \\ &= 4(t+3) \pmod 7 \end{aligned}$$

Fix $t = 81519$, then $t \equiv 4 \pmod 7$ and $t \leq 81525$. By above, $S(N) - S_p(N) \equiv 0 \pmod 7$. Calculate k

$$k = \frac{S(N) - S_p(N)}{7} = \frac{441725t + 392642}{7} = 5144196131$$

then Fact 2 to proves $10^k N$ is a Smith number. The explicit description for this Smith number is

$$\begin{aligned} 10^k N &= (3 \cdot 10^{665829} + 1)^t \cdot 9R_{49081} \cdot 10^k \\ &= (3 \cdot 10^{665829} + 1)^{81519} \cdot (10^{49081} - 1) \cdot 10^{5144196131} \end{aligned}$$

This Smith number has 59 421 998 358 digits.

4. Best by adding coefficients

The previous section relied on **Fact 1** to greatly simplify calculation of the coefficient digit sums. This caused the coefficient $t = 81519$ to be limited by the size of a provably-prime repunit R_{49081}. With modern processors, it is possible to explicitly calculate each coefficient and the resulting digit sum, allowing us to discard the R_{49081} from the product and achieve far larger t powers.

For selecting a power t of M^t while planning to explicitly calculate each $S(c_{t,k})$, we only need to worry about two constraints. Firstly, it is necessary that $S(M^t) = S_p(M^t) \pmod 7$ to apply **Fact 2**. Secondly, t will need to be small enough that each coefficient $c_{t,k} < 10^{665829}$ so that the coefficients can be separated in the digit representation. By Theorem 3.1, it's sufficient to show that only $c_{t,k(t)} < 10^{665829}$.

After a somewhat quick search for a bound on the second constraint, I found that fixing $t_0 = 1105923$ yields

$$c_{t_0,k(t_0)} = 8.5967... \cdot 10^{665828} < 10^{665829}$$

This is maximal as

$$10^{665829} < c_{t_0+1,k(t_0+1)} = 3.4387... \cdot 10^{665829}$$

As long as $t \leq t_0$, then the second constraint is satisfied. Since we are not relying on **Fact1**, the digit sum of M^t is much less predictable. Therefor, finding a $t \leq t_0$ so that $S(M^t) = S_p(M^t) \pmod 7$ simply involves checking $t_0, t_0 - 1, t_0 - 2, ...$ until one of these values happen to satisfy the congruence.

By a lucky $\frac{1}{7}$ chance, it turns out $S(M^{t_0}) = S_p(M^{t_0}) \pmod 7$. This took 4-ish hours to calculate using all cores on an AMD Ryzen 7 5800X 8-Core Processor.

$$\begin{aligned} S(M^{t_0}) &= 2508098743612 \\ S_p(M^{t_0}) &= (3+1) \cdot 1105923 = 4423692 \\ S(M^{t_0}) - S_p(M^{t_0}) &= 2508094319920 \end{aligned}$$

Let $k = \frac{S(M^{t_0}) - S_p(M^{t_0})}{7} = 358299188560$, then finally let $N = M^{t_0} 10^k$. N is a Smith number with explicit form

$$N = (3 \cdot 10^{665829} + 1)^{1105923} \cdot 10^{358299188560}$$

This Smith number has 1 094 654 215 464 digits.

5. Implementation & Afterword

References

[1] Amin Witno, *Another Simple Construction of Smith Numbers*, Missouri Journal of Mathematical Sciences **22** (2010), no. 2, 97 – 101, DOI 10.35834/mjms/1312233139.

[2] Patrick Costello, *Smith Numbers From Primes with Small Digits*, Missouri Journal of Mathematical Sciences **27** (2015), no. 1, 10 – 15, DOI 10.35834/mjms/1449161363.

[3] Patrick Costello, *A new largest Smith number*, The Fibonacci Quarterly **40** (2002).

[4] Kathy Sharon Lewis, *Smith numbers: an infinite subset of N*, Eastern Kentucky University, 1994.

[5] A. Wilansky, *Smith Numbers*, The Two-Year College Mathematics Journal **13** (1982), no. 1, 21-21, DOI 10.1080/00494925.1982.11972573.

[6] Sham Oltikar and Keith Wayland, *Construction of Smith Numbers*, Mathematics Magazine **56** (1983), no. 1, 36-37, DOI 10.1080/0025570X.1983.11977013.

[7] Chris K Caldwell, *PrimePage Primes: 3 · 10^665829 + 1* (2012), `https://primes.utm.edu/primes/page.php?id=109247`.

[8] Chris K Caldwell and Paul Underwood, *PrimePage Primes: R(49081)* (2021), `https://primes.utm.edu/primes/page.php?id=133761`.