# 1

# Eight Hateful Sequences

**N. J. A. Sloane**

> Dear Martin Gardner:
>
> In your July 1974 *Scientific American* column you mentioned the *Handbook of Integer Sequences*, which then contained 2372 sequences. Today the *On-Line Encyclopedia of Integer Sequences* (the *OEIS*) (Sloane, 2010) contains 140000[1] sequences. Here are eight of them, suggested by the theme of the Eighth Gathering For Gardner: they are all infinite, and all 'ateful in one way or another. I hope you like 'em! Each one is connected with an interesting unsolved problem.

Since this is a 15-minute talk, I can't give many details—see the entries in the OEIS for more information, and for links to related sequences.

## 1.1 Hateful or Beastly Numbers

The most hateful sequence of all! These are the numbers that contain the string 666 in their decimal expansion:

$$666, 1666, 2666, 3666, 4666, 5666, 6660, 6661,$$
$$6662, 6663, 6664, 6665, 6666, 6667, \ldots$$

(A051003). This sequence is doubly hateful because it is based on superstition and because it depends on the fact that we write numbers in base 10.

[1] September 10, 2013: there are now 228560 sequences.

It has been said that if the number of the Beast is 666, its fax number must be 667; fortunately this has not yet been made the basis for any sequence. (Added later: unfortunately this is no longer true—see A138563!)

Base-dependent sequences are not encouraged in the OEIS, but nevertheless many are included because someone has found them interesting or they have appeared on a quiz or a website, etc. One has to admit that some of these sequences are very appealing.

For example: start with $n$; if it is a palindrome, stop; otherwise add $n$ to itself with the digits reversed; repeat until you reach a palindrome; or set the value to $-1$ if you never reach a palindrome. Starting at 19, we have

$$19 \to 19 + 91 = 110 \to 110 + 011 = 121,$$

which is a palindrome, so we stop, and the 19th term of the sequence is 121.

The sequence (A033865) which gives the first palindrome that is reached, or $-1$ if no palindrome is ever reached, begins:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 11, 33, 44, 55,$$
$$66, 77, 88, 99, 121, 22, 33, 22, 55, 66, 77, \ldots .$$

The first unsolved case occurs when we start with 196. Sequence A006960 gives its trajectory, which begins

$$196, 887, 1675, 7436, 13783, 52514, 94039,$$
$$187088, 1067869, 10755470, 18211171, \ldots .$$

It *appears* that it never reaches a palindrome, but it would be nice to have a proof! This is hateful because it is one of those problems that seem too difficult for twenty-first century mathematics to solve.

## 1.2 Éric Angelini's "1995" puzzle

The following puzzle was invented by Éric Angelini in September 2007: Find the rule that generates the sequence

o n e   n i n e   n i n e   f i v e   f i v e   n i n e   n i n e   f i v e . . . .

The answer is that if each letter is replaced by its rank in the English alphabet, then the absolute values of the differences between successive numbers produce the same sequence:

$$1, 9, 9, 5, 5, 9, 9, 5, 5, 9, 1, 3, 13, 17, 1, 3,$$
$$13, 17, 9, 5, 5, 9, 9, 5, 5, 9, 1, 3, 13, 17, \ldots$$

(A131744). In (Applegate and Sloane, n.d.), David Applegate and I analyzed this sequence, and showed among other things that only 19 numbers occur (16, 19, 20, and 22–26 never appear). We determined the relative frequencies of these 19 numbers: 9 occurs the most often, with density $0.173\cdots$.

In English this sequence is unique: “one” is the only number with the property that it is equal to the absolute value of the difference in rank between the first two letters of its name (“o” is the fifteenth letter of the alphabet, “n” is the fourteenth, and $1 = |14 - 15|$). We also discuss versions in other languages. In French one can begin with either 4 (see A131745) or 9 (A131746), in German with either 9 (A133816) or 15 (A133817), in Italian with 4 (A130316), in Russian with either 1 (A131286) or 2 (A131287), and so on.

## 1.3 Powertrains

What is the next term in the following sequence?

$$679 \to 378 \to 168 \to 48 \to 32 \to ?$$

Answer: 6. The reason is that each term is the product of the digits of the previous term. Eventually every number $n$ reaches a single-digit number (these are the only fixed points), and the number of steps for this to happen is called the *persistence* of $n$. 679 has persistence 5, and it is in fact the smallest number with persistence 5. The smallest numbers with persistence $n = 1, 2, \ldots, 11$ are given in sequence A003001:

$$10, 25, 39, 77, 679, 6788, 68889, 2677889,$$
$$26888999, 3778888999, 277777788888899.$$

This sequence was the subject of an article I wrote in 1973 (Sloane, 1973), which Martin Gardner may remember! I conjectured that the sequence is finite, and even today no number of persistence greater than 11 has been found.

In December 2007, John Conway proposed some variations of “persistence”, one of which I will discuss here (see (Conway and Sloane, n.d.) for further information). If $n$ has decimal expansion $abcd\cdots$, the *powertrain* of $n$ is the number $a^b c^d \cdots$, which ends in an exponent or a base according as the number of digits in $n$ is even or odd. We take $0^0 = 1$ and define the powertrain of 0 to be 0.

The OEIS now contains numerous sequences related to powertrains; see A133500 and the sequences cross-referenced there. For example, the following numbers are fixed under the powertrain map:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 2592, 24547284284866560000000000$$

(A135385), and we conjecture that there are no others. Certainly, there are no other fixed points below $10^{100}$.

## 1.4 Alekseyev's "123" sequence

This is a question about strings, proposed by Max Alekseyev when he was a graduate student in the Computer Science Department of the University of California at San Diego (personal communication). If we start with the string 12, and repeatedly duplicate any substring in place, the strings we obtain are

$$\begin{array}{c} 12 \\ 112 \quad 122 \\ 1112 \quad 1122 \quad 1212 \quad 1222 \\ 11112 \quad 11122 \quad 11212 \quad 11222 \quad 12112 \quad 12122 \quad 12212 \quad 12222 \\ \cdots \end{array}$$

(A130838). These strings must start with 1 and end with 2, but are otherwise arbitrary. So the number of such strings of length $n$ is $2^{n-2}$ for $n \geq 2$.

But what if we start with the string 123? Now the strings we obtain are

$$\begin{array}{c} 123 \\ 1123 \quad 1223 \quad 1233 \\ 11123 \quad 11223 \quad 11233 \quad 12223 \quad 12233 \quad 12333 \quad 12123 \quad 12323 \\ \cdots \end{array}$$

(A135475), and the number of such strings of length $n \geq 3$ is

$$1, 3, 8, 21, 54, 138, 355, 924, 2432, 6461, 17301, 46657, 126656, 345972, \ldots$$

(A135473). The question is, what is the $n$th term in the latter sequence? This is hateful because one feels that if this sequence was only looked at in the right way, there would be a simple recurrence or generating function.[2]

## 1.5 The Curling Number Conjecture

Let $S = S_1 S_2 S_3 \cdots S_n$ be a finite string (the symbols can be anything you like). Write $S$ in the form $XYY \cdots Y = XY^k$, consisting of a prefix $X$ (which may be empty), followed by $k$ (say) copies of a nonempty string $Y$. In general there will be several ways to do this; pick one with the greatest value of $k$. This $k$ is called the *curling number* of $S$.

A few years ago Dion Gijswijt proposed the sequence that is obtained by starting with the string 1, and extending it by continually appending the curling number of the current string. The resulting sequence (A090822)

$$1, 1, 2, 1, 1, 2, 2, 2, 3, 1, 1, 2, 1, 1, 2, 2, 2,$$
$$3, 2, 1, 1, 2, 1, 1, 2, 2, 2, 3, 1, 1, 2, 1, 1, \ldots$$

[2] It is known from the work of Ming-Wei Wang (Wang, 2000) that the set of strings produced does not form a regular language, suggesting that this question may not have a simple answer.

was analyzed in van der Bult et al. (2007), and I talked about it at the Seventh Gathering 4 Gardner (Sloane, 2009). It is remarkable because, although it is unbounded, it grows *very* slowly. For instance, the first 5 only appears after about $10^{10^{23}}$ terms.

Some of the proofs in that paper could have been shortened if we had been able to prove a certain conjecture, which remains open to this day. This is the *Curling Number Conjecture*, which states that if one starts with any finite string, over any alphabet, and repeatedly extends it by appending the curling number of the current string, then eventually one must reach a 1.

One way to attack this problem is to start with a string that only contains 2's and 3's, and see how far we can get before a 1 appears. For an initial string of $n$ 2's and 3's, for $n = 1, 2, \ldots, 30$, respectively, the longest string that can be obtained before a 1 appears is

$$1, 4, 5, 8, 9, 14, 15, 66, 68, 70, 123, 124, 125, 132, 133, 134, 135, 136,$$
$$138, 139, 140, 142, 143, 144, 145, 146, 147, 148, 149, 150, \ldots$$

(A094004). For example, the best initial string of length 6 is $2, 2, 2, 3, 2, 2$. This produces the string

$$2, 2, 2, 3, 2, 2, 2, 3, 2, 2, 2, 3, 3, 2, 1, \ldots,$$

which extends for 14 terms before a 1 appears. It's hateful because it seems hard to predict the asymptotic behavior. Does it continue in a roughly linear manner for ever, or are there bigger and bigger jumps? If the Curling Number Conjecture is false, the terms could be a "lazy eight" ($\infty$) from some point on![3]

## 1.6 Leroy Quet's prime-generating recurrence

Leroy Quet has contributed many original sequences to the OEIS. Here is one which so far has resisted attempts to analyze it. Let $a(0) = 2$, and for $n \geq 1$, define $a(n)$ to be the smallest prime $p$, not already in the sequence, such that $n$ divides $a(n-1) + p$. The sequence begins

$$2, 3, 5, 7, 13, 17, 19, 23, 41, 31, 29, 37, 11, 67,$$
$$59, 61, 83, 53, 73, 79, 101, 109, 89, 233, \ldots$$

(A134204). Is it infinite? To turn the question around, does $a(n-1)$ ever divide $n$? If it *is* infinite, is it a permutation of the primes? (This is somewhat reminiscent of the EKG sequence A064413 (Lagarias et al., 2002).) This is hateful because we do not know the answers!

[3] For recent work on this problem, see (Chaffin et al., 2013).

David Applegate has checked that the sequence exists for at least $450 \cdot 10^6$ terms. Note that $a(n)$ can be less than $n$. This happens for the following values of $n$ (A133242):

$$12, 201, 379, 474, 588, 868, 932, 1604, 1942, 2006, 3084, 4800, 7800, \ldots.$$

## 1.7 The $n$-point traveling salesman problem

My colleagues David Johnson and David Applegate have been re-examining the old question of the expected length of a traveling salesman tour through $n$ random points in the unit square (cf. Applegate et al. (2007); Beardwood et al. (1959)). To reduce the effects of the boundary, they identify the edges of the square so as to obtain a flat torus and ask for the expected length, $L(n)$, of the optimal tour through $n$ random points (Johnson, 2008). It is convenient to express $L(n)$ in units of "eels" (the *expected Euclidean length* of the line joining a random point in the unit square to the center), a quantity which is well-known (Finch, 2003, p. 479) to be

$$\frac{\sqrt{2} + \log(1 + \sqrt{2})}{6} = 0.382597858232 \cdots$$

(although our name for it is new!). Then $L(1) = 0$, $L(2) = 2$ eels, and it is not difficult to show that $L(3) = 3$ eels. David Applegate has made Monte Carlo estimates of $L(n)$ for $n \leq 50$ based on finding optimal tours through a million sets of random points. His estimates for $L(2)$ through $L(10)$, expressed in eels, are

$$1.99983, 2.99929, 3.60972, 4.08928, 4.5075,$$

$$4.88863, 5.24065, 5.5712, 5.8825, 6.17719.$$

Observe that the first two terms are a close match for the true values. It would be nice to know the exact value of $L(4)$. The sequence $\{L(n)\}$ is hateful because for $n \geq 4$ it may be irrational, even when expressed in terms of eels, and so will be difficult to include in the OEIS.

## 1.8 Lagarias's Riemann hypothesis sequence

Every inequality in number theory is potentially the source of a number sequence (if the inequality states that $f(n) \geq g(n)$, the corresponding sequence is $a(n) = \lfloor f(n) - g(n) \rfloor$). Here is one of the most remarkable examples. Consider the sequence defined by

$$a(n) = \lfloor H(n) + \exp(H(n)) \log(H(n)) \rfloor - \sigma(n),$$

where $H(n) = \sum_{k=1}^{n} 1/k$ is the $n$th harmonic number (see A001008 and A002805) and $\sigma(n)$ is the sum of the divisors of $n$ (A000203). This begins

$$0, 0, 1, 0, 4, 0, 7, 2, 7, 5, 13, 0, 17, 9, 12, 8, 23, 5, 27, 8, 21, 20, 34, 1, 33, 25, \ldots$$

(A057641). Jeff Lagarias (Lagarias, 2002), extending earlier work of G. Robin (Robin, 1984), has shown that proving that $a(n) \geq 0$ for all $n$ is equivalent to proving the Riemann hypothesis! Hateful because it's *hard*.

## 1.9 Acknowledgment

I would like to thank my colleague David Applegate for being always ready to help analyze a new sequence, as well as for stepping in on several occasions to keep the OEIS website up and running when it was in difficulties.

## Further Reading

D. L. Applegate and N. J. A. Sloane. *Éric Angelini's "1995" puzzle sequence* (in preparation).

D. L. Applegate, R. E. Bixby, V. Chvtal, and W. J. Cook (2007). *The Traveling Salesman Problem: A Computational Study*. Princeton University Press.

J. Beardwood, J. H. Halton, and J. M. Hammersley (1959). The shortest path through many points. *Proceedings of the Cambridge Philosophical Society*, 55: 299–327.

B. Chaffin, J. P. Linderman, N. J. A. Sloane, and A. R. Wilks (2013). On curling numbers of integer sequences. *Journal of Integer Sequences*, 16(13.4.3).

J. H. Conway and N. J. A. Sloane. *Powertrains and other sequences* (in preparation).

S. R. Finch (2003). *Mathematical constants*. Cambridge University Press, New York. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=589155.

D. S. Johnson (2008). *Comparability*. A talk given at Workshop on Experimental Analysis of Algorithms: Interfaces between the Statistical and Computational Sciences, Research Triangle Park, NC.

J. C. Lagarias (2002). An elementary problem equivalent to the Riemann hypothesis. *The American Mathematical Monthly*, 109(6): 534–543. http://www.jstor.org/stable/2695443.

J. C. Lagarias, E. M. Rains, and N. J. Sloane (2002). The EKG sequence. *Experimental Mathematics*, 11(3): 437–446. DOI 10.1080/10586458.2002.10504486.

G. Robin (1984). Grandes valeurs de la fonction somme des diviseurs et hypothèse de Riemann, *Journal de Mathèmatiques Pures et Appliquées*, 63: 187–213.

N. J. A. Sloane (1973). The persistence of a number. *Journal of Recreational Mathematics*, 6(2): 97–98.

N. J. A. Sloane (2009). *Homage to a pied puzzler*; see chapter seven, Staggering sequences, pp. 93–110. A. K. Peters, Wellesley, Mass.

N. J. A. Sloane (2010). *The on-line encyclopedia of integer sequences*. `http://oeis.org`.

F. J. van der Bult, D. C. Gijswijt, J. P. Linderman, N. J. A. Sloane, and A. R. Wilks (2007). A slow-growing sequence defined by an unusual recurrence. *Journal of Integer Sequences*, 10(#07.1.2).

M. W. Wang (2000). On the irregularity of the duplication closure. *Bulletin of the European Association for Theoretical Computer Science*, 70: 162–163.