University of Wisconsin Platteville
Department of Mathematics

Problem of the Week #7

(Spring 2022)

A positive integer n is a *semi*-1 *number* if exactly half of the integers from 1 through n contain the digit 1. For example, 16 is semi-1, because exactly 8 of the integers between 1 and 16 contain the digit 1:

$$\{1, 10, 11, 12, 13, 14, 15, 16\}.$$

Your challenge: submit the largest semi-1 number you can find! How do you know that your number is semi-1?

Bonus: Are there infinitely many semi-1 numbers, or only finitely many?

Solution:
A bit of computer programming suggests that there are exactly 16 semi-1 numbers:

$$\{2, 16, 24, 160, 270, 272, 1456, 3398, 3418, 3420, 3422, 13120, 44686, 118096, 674934, 1062880\}.$$

For the bonus, let's prove that there are only finitely many semi-1 numbers.

Proof. For any positive integer n, let $c(n)$ be the number of integers from 1 through n that contain the digit 1. Then $d(n)$ integers from 1 through n do not contain the digit 1, where $c(n) + d(n) = n$. Let $f(n) = c(n) - d(n)$; thus between 1 and n, there are $f(n)$ more integers that do contain the digit 1 than don't. By definition, n is semi-1 if and only if $f(n) = 0$.
For any t we have $f(t + 1) = f(t) \pm 1$, and by induction, $f(t) - m \le f(t + m) \le f(t) + m$ for any t and m.
Also, for any positive integer k, we have $d(10^k) = 9^k - 1$, so $c(10^k) = 10^k - 9^k + 1$ and $f(10^k) = 10^k - 2 \cdot 9^k + 2$. Specifically, $f(10^7) = 434{,}064 > 0$.
[Note in passing that $f(10^6) = 10^6 - 2 \cdot 9^6 + 2 = -62880$, which is why $f(1062880) = 0$.]
Next, note that from $10^k + 1$ through $2 \cdot 10^k$ there are $10^k - 1$ integers that contain the digit 1, and one integer that does not, so $f(2 \cdot 10^k) = f(10^k) + 10^k - 2$. So if $k > 1$, then $f(2 \cdot 10^k) > f(10^k)$.
On the other hand, suppose that a is an integer and $2 \le a \le 9$. If $a \cdot 10^k < x \le (a + 1) \cdot 10^k$, then $0 < x - a \cdot 10^k \le 10^k$, and x and $x - a \cdot 10^k$ contain the digit 1 the same number of times. Therefore, $f((a + 1)10^k) = f(a \cdot 10^k) + f(10^k)$. By induction, this shows that if $k \ge 7$, then

$$0 < f(10^k) < f(2 \cdot 10^k) < f(3 \cdot 10^k) < \cdots < f(9 \cdot 10^k) < f(10^{k+1}).$$

Specifically, if $k > 7$, then $f(10^k) > f(10^7)$.
Finally, let $x \ge 10^7$. Let a be the most significant digit of x, so that $x = a \cdot 10^k + m$ for some integers k and m with $k \ge 7$ and $0 \le m < 10^k$. If $a = 1$, then $f(x) = f(10^k + m) = f(10^k) + m > m \ge 0$. Otherwise,

$$f(x) = f(a{\cdot}10^k{+}m) \ge f(a{\cdot}10^k){-}m \ge f(2{\cdot}10^k){-}m = f(10^k){+}10^k{-}2{-}m > f(10^k){-}2 \ge f(10^7){-}2 > 0.$$

In short, if $x \ge 10^7$, then $f(x) > 0$, so x is not semi-1. Therefore, there are only finitely many semi-1 numbers. [For the programming mentioned in the first line of the solution, it's a big help to know that you can stop looking at $n = 10^7$.]

□

Source: Adam Atkinson, University of Pisa.