PROBLEM 247. GOLOMB'S SELF-DESCRIBING SEQUENCE.

The *Golomb's self-describing sequence* $\{G(n)\}$ is the only nondecreasing sequence of natural numbers such that n appears exactly $G(n)$ times in the sequence. The values of $G(n)$ for the first few n are

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$G(n)$	1	2	2	3	3	4	4	4	5	5	5	6	6	6	6

You are given that $G(10^3) = 86$, $G(10^6) = 6137$. You are also given that $\sum_{n=1}^{999} G(n^3) = 153506976$.

Find $\sum_{n=1}^{999999} G(n^3)$.

SOLUTION

Since there is no explicit formula for $G(n)$ and apparently no special property for $G(n^3)$, this problem is about storing and retrieving the terms in $\{G(n)\}$ efficiently.

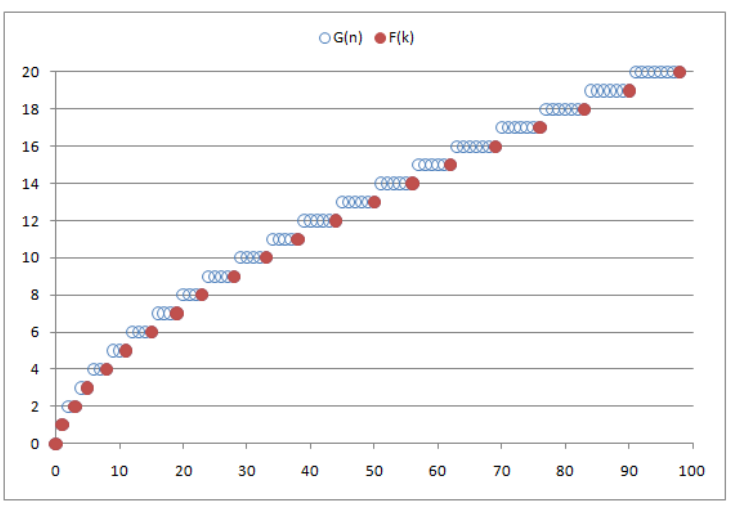
First, note that $\{G(n)\}$ can be stored in compact form by omitting duplicate terms and keeping only the largest n for a given value. Let F denote the generalized inverse function of G: $F(k)$ is the location of the last appearance of k in $\{G(n)\}$. Formally,

$$F(k) = \max\{n \,|\, G(n) = k\}.$$

The first 10 values of $F(k)$ are

k	1	2	3	4	5	6	7	8	9	10
$F(k)$	1	3	5	8	11	15	19	23	28	33

Below is a plot of the first 98 values of $G(n)$ and first 20 values of $F(k)$:



It is easy to find the recurrence relation

$$F(k) = F(k-1) + G(k), \tag{1}$$

with the starting value $F(0) = 0$. To understand this, note that $F(k-1)$ is the location of the last appearance of $(k-1)$ in $\{G(n)\}$, and k appears after $(k-1)$ for $G(k)$ times. Therefore the location of the last appearance of k is $F(k-1) + G(k)$.

Once we have obtained the values of $\{F(k)\}$, we can back out $G(n)$ as follows. Find the unique k such that $F(k-1) < n \leq F(k)$, then $G(n) = k$.

Now we move one step further, and claim that $F(k)$ is also the largest value that appears k times in $\{G(n)\}$. This is because all $\{G(n)\}$ where $F(k-1) < n \leq F(k)$ are equal to k, so every n where $F(k-1) < n \leq F(k)$ appears k times in the sequence. But $F(k)$ is, by definition, the largest such n. So it is the largest value that appears k times.

Next, let $F(F(k))$ be the location of the last appearance of $F(k)$ in $\{G(n)\}$. The first few values of $F(F(k))$ are

k	1	2	3	4	5	6	7	8	9	10
$F(k)$	1	3	5	8	11	15	19	23	28	33
$F(F(k))$	1	5	11	23	38	62	90	122	167	217

We find the following recurrence relation for $F(F(k))$:

$$F(F(k)) = F(F(k-1)) + kG(k), \tag{2}$$

with the starting value $F(F(0)) = 0$. To understand this, note that each n where $F(k-1) < n \leq F(k)$ appears k times in the sequence, and there are $G(k)$ values that appear k times.

Once we have obtained the sequence $\{F(k), F(F(k))\}$, we can back out $G(n)$ as follows. Given n, find the unique k such that $F(F(k-1)) < n \leq F(F(k))$. By construction, every n where $F(k-1) < n \leq F(k)$ appears k times, and $G(F(F(k))) = F(k)$. It follows that

$$G(n) = F(k) - \left\lfloor \frac{F(F(k)) - n}{k} \right\rfloor.$$

The algorithm is outlined as follows. We use equation (1) to build and store the sequence $\{F(k)\}$, which is then used to find out $G(n)$ for small n. Then we use equation (2) to compute the sequence $\{F(F(k))\}$. As we compute, we back out the values of $G(n^3)$. So there is no need to store $\{F(F(k))\}$.

To estimate the complexity of the algorithm, we run a regression on the first one million values of $F(k)$ (with 1000 samples at $k = 1000, 2000, \ldots$) and get the following empirical relation:

$$F(k) \approx \varphi^{1-\varphi} k^{\varphi},$$

where $\varphi = (1+\sqrt{5})/2 \approx 1.618$ is the golden ratio. A few useful results follow:

$$F^{-1}(n) \approx \varphi^{2-\varphi} n^{\varphi-1}.$$
$$F^{-1}(F^{-1}(n)) \approx \varphi^{\varphi-1} n^{2-\varphi}.$$
$$F^{-1}(F^{-1}(F^{-1}(n))) \approx \varphi^{4-2\varphi} n^{2\varphi-3}.$$

In this problem, we are asked to find $F(F(k)) \geq n$ for $n = 10^{18}$. This corresponds to $k \approx F^{-1}(F^{-1}(n)) = 10105311$. To use equation (2) to compute $F(F(k))$, we need to know $G(k)$, which can in turn be backed out using equation (1). This requires a storage of $F^{-1}(k) \approx 25638$.

Complexity

Time complexity: $\mathcal{O}(1.35 \times n^{1.15})$

Space complexity: $\mathcal{O}(1.44 \times n^{0.71})$

Answer

56098610614277014