3b1b puzzle

lewisxy

May 2022

1 Problem Statement

Find the number of subsets of $\{1, ..., 2000\}$, the sum of whose elements is divisible by 5.

2 Approaching the problem

The total number of subsets of set $\{1, ..., k\}$ is 2^k. 2^{2000} is quite a large number. 2000 seems quite arbitrary, so we can replace it with k. If we can solve the problem for small k (hopefully not too difficult), we might be able to use a similar method to solve the real problem ($k = 2000$).

3 Initial Attempt

Let U_k be the set of all subset of $\{1, ..., k\}$ (for natural number k). Let $A_k \subseteq U_k$ be the subsets sum of whose elements is divisible by 5. Let $B_k = U_k \setminus A_k$, which is the rest of subsets (the sum of whose elements is not divisible by 5). The goal is to find $|A_k|$.

Since $k = 1, 2, ...$, in the spirit of induction, if we can find the relationship between $|A_k|$ and $|A_{k-1}|$ as well as $|A_0|$, then the problem is solved. Let's try that.

We know $|U_k| = 2|U_{k-1}|$ for any k since for any subset $S \in U_{k-1}$, $S \in U_k$ and $S \cup \{k\} \in U_k$. Similarly, suppose we know $|A_{k-1}|$ for some k, then if k is divisible by 5, we know that $|A_k| \geq 2|A_{k-1}|$ since for any $S \in A_{k-1}$, $S \in A_k$ and $S \cup \{k\} \in A_k$. For those of B_{k-1}, adding a number (k) divisible by 5 will not make the sum divisible by 5. So $|A_k| = 2|A_{k-1}|$.

If k is not divisible by 5, it's a bit more complex. Since for all $S \in A_{k-1}$, $S \cup \{k\} \notin A_k$, and there are elements of $T \in B_{k-1}$ that $T \cup \{k\} \in A_k$. So $|A_k| = |A_{k-1}| + X$. In which X is the number of such $T \in B_{k-1}$.

There is no straightforward way to compute X using the current formulation, so we need to be a bit systematic.

4 Improved Formulation

In our initial formulation, we partition the entire space U_k into A_k and B_k. The elements in A_k has sum divisible by 5, The elements in B_k has sum that is not divisible by 5. Observe that there are only 5 outcomes if we take the remainder of any integer divided by 5. That is 0, 1, 2, 3, 4.

Therefore, we can partition U_k into $P_0(k), P_1(k), P_2(k), P_3(k), P_4(k)$, in which $P_i(k) \subseteq U_k$ is the subset the sum of whose elements is i mod 5. Observe that $A_k = P_0(k)$ and $B_k = P_1(k) \cup P_2(k) \cup P_3(k) \cup P_4(k)$.

With this formulation, we can derive the relationship between $|P_i(k-1)|$ and $|P_i(k)|$.

Using the result of initial attempt. If $k \equiv 0 \bmod 5$, then $|P_0(k)| = |P_0(k-1)| + |P_0(k-1)| = 2|P_0(k-1)|$. Similarly, $|P_1(k)| = |P_1(k-1)| + |P_1(k-1)| = 2|P_1(k-1)|$, and so on. This is because for all $S \in P_i(k-1)$, $S \in P_i(k-1)$ and $S \cup \{k\} \in P_i(k-1)$.

If $k \equiv 1 \bmod 5$, then $|P_0(k)| = |P_0(k-1)| + |P_4(k-1)|$. Observe that for all $S \in P_4(k-1)$, $S \cup \{k\} \in P_0(k-1)$ because $1(\mod 5) + 4(\mod 5) = 0(\mod)5$. Similarly, $|P_1(k)| = |P_1(k-1)| + |P_0(k-1)|$ and so on.

If we apply this approach for all 5 cases of k for all 5 cases of i. We can use the following table to describe the relationship between $|P_i(k-1)|$ and $|P_i(k)|$.

	$\lvert P_0\rvert$	$\lvert P_1\rvert$	$\lvert P_2\rvert$	$\lvert P_3\rvert$	$\lvert P_4\rvert$
0	$\lvert P_0\rvert + \lvert P_0\rvert$	$\lvert P_1\rvert + \lvert P_1\rvert$	$\lvert P_2\rvert + \lvert P_2\rvert$	$\lvert P_3\rvert + \lvert P_3\rvert$	$\lvert P_4\rvert + \lvert P_4\rvert$
1	$\lvert P_0\rvert + \lvert P_4\rvert$	$\lvert P_1\rvert + \lvert P_0\rvert$	$\lvert P_2\rvert + \lvert P_1\rvert$	$\lvert P_3\rvert + \lvert P_2\rvert$	$\lvert P_4\rvert + \lvert P_3\rvert$
2	$\lvert P_0\rvert + \lvert P_3\rvert$	$\lvert P_1\rvert + \lvert P_4\rvert$	$\lvert P_2\rvert + \lvert P_0\rvert$	$\lvert P_3\rvert + \lvert P_1\rvert$	$\lvert P_4\rvert + \lvert P_2\rvert$
3	$\lvert P_0\rvert + \lvert P_2\rvert$	$\lvert P_1\rvert + \lvert P_3\rvert$	$\lvert P_2\rvert + \lvert P_4\rvert$	$\lvert P_3\rvert + \lvert P_0\rvert$	$\lvert P_4\rvert + \lvert P_1\rvert$
4	$\lvert P_0\rvert + \lvert P_1\rvert$	$\lvert P_1\rvert + \lvert P_2\rvert$	$\lvert P_2\rvert + \lvert P_3\rvert$	$\lvert P_3\rvert + \lvert P_4\rvert$	$\lvert P_4\rvert + \lvert P_0\rvert$

Table 1: The row label represents the k (from 0 mod 5 to 4 mod 5), the column label represents the $|P_i(k)|$, the element of the table shows how to compute it, the $(k-1)$ part is omitted to save space

Next, we find the base case with $k = 0$. $U_0 = \{\emptyset\}$, and $|P_0(0)| = 1$ and $|P_1(0)| = |P_2(0)| = |P_3(0)| = |P_4(0)| = 0$.

5 The Solution

With the base case and table to compute the case for k from $k-1$. We can compute $|P_0(2000)|$ easily with Python, thanks to its built-in big integer support.

The answer is: `2296261390548509048...74548427800576`, which is exactly $\frac{1}{5}(2^{2000} + 4 \cdot 2^{200})$. You can verify this using Wolfarm Alpha.