

let S be the set of the 10^N , N -digit numbers assuming $N \geq 2$. Consider subsets of S , A_k , where the elements of A_k are all the elements of S having at least k occurrences of the substring “42”, where $k = 1, 2, \dots, \lfloor \frac{N}{2} \rfloor$. Note that since $A_{\lfloor \frac{N}{2} \rfloor} \subset \dots \subset A_3 \subset A_2 \subset A_1$, the cardinality of the union of all subsets A_k is equal to the cardinality $|A_1|$, which is the quantity we want. By inclusion-exclusion the cardinality of the union of all such subsets is

$$\left| \bigcup_{k=1}^{\lfloor \frac{N}{2} \rfloor} A_k \right| = \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} |A_k| - \sum_{1 \leq k < j \leq \lfloor \frac{N}{2} \rfloor} |A_k \cap A_j| + \dots + (-1)^{\lfloor \frac{N}{2} \rfloor - 1} |A_1 \cap A_1 \cap \dots \cap A_{\lfloor \frac{N}{2} \rfloor}| \quad (1)$$

The first summation term on the right hand side of the above equation is a sum over k of cardinalities $|A_k|$,

$$\sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} |A_k| = (N-1)10^{N-2} \quad (2)$$

One can see this by thinking about the number of positions at which a single “42” can sit. There are clearly $N-1$ such positions. The remaining $N-2$ digits can take any value, hence the 10^{N-2} term. The sum, Eq. (2), is not equal to $|A_1|$ because it includes overcounting terms due to subset intersections. Each $A_k \cap A_j$ is the set of elements having two or more occurrences. Adding up all the possible combinations of two or more occurrences gives

$$\sum_{1 \leq k < j \leq \lfloor \frac{N}{2} \rfloor} |A_k \cap A_j| = 10^{N-4} \binom{N-2}{2}$$

Treating all terms in Eq. (1) in similar fashion we get

$$\left| \bigcup_{k=1}^{\lfloor \frac{N}{2} \rfloor} A_k \right| = \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} (-1)^{k+1} 10^{N-2k} \binom{N-2k+k}{k} \quad (3)$$

For $N = 6$ this gives 49401 occurrences