

22C:44 Homework 5 Solution

The probabilities in the problem were changed via an announcement in class and the problem with the new probabilities is as follows.

- 2 Suppose that the universe contains the million numbers 1 through 10^6 . Further suppose that the probability distribution P telling us the likelihood of an element in the universe being chosen into S is as follows:

$$P(k) = \begin{cases} 2/(3 \times 10^4) & 1 \leq k \leq 10^4 \\ 1/(54 \times 10^4) & 10^4 < k \leq 10^5 \\ 1/(54 \times 10^5) & 10^5 < k \leq 10^6 \end{cases}$$

Given this probability distribution, prove or disprove the following claim: $h(k) = k \bmod 1000$ is a simple uniform hashing function.

Let $U_i = \{k \in U \mid h(k) = i\}$ for $i = 0, 1, \dots, 999$. Each element in U_i has the form $1000q + i$ for some non-negative integer q . It is easy to see that there are 10 elements in U_i in the range $[1, 10^4]$, 90 in the range $[10^4, 10^5]$ and 900 in the range $[10^5, 10^6]$. This implies that

$$\begin{aligned} \sum_{k \in U_i} P(k) &= 10 \cdot \frac{2}{3 \times 10^4} + 90 \cdot \frac{1}{54 \times 10^4} + 900 \cdot \frac{1}{54 \times 10^5} \\ &= \frac{1}{1000} \left[\frac{2}{3} + \frac{1}{6} + \frac{1}{6} \right] \\ &= \frac{1}{1000}. \end{aligned}$$

This shows that the given hash function is indeed a simple uniform hash function.
