Tutorial 7: CS21003 Algorithms I

Prof. Partha Pratim Chakrabarti and Palash Dey
Indian Institute of Technology, Kharagpur

March 18, 2021

For any hash function, there exists a sequence of keys which makes the hash function perform badly — there can be too many collisions. To tackle this problem, we can pick our hash function randomly from a set of hash functions at runtime. This approach does not suffer from the problem discussed above — hopefully there no sequence which is bad for all hash functions in the set. Clearly, such a set of hash functions need to satisfy certain properties; for example a singleton set is not appropriate. In this tutorial, we formalize this notion which is called *universal hash family*.

Let $\mathcal{U}$ be the universe of keys. Hash functions map from $\mathcal{U}$ to $\{0, \ldots, m-1\}$. Let $\mathcal{H}$ be a set of hash functions. We call $\mathcal{H}$*universal* if for every two distinct keys $x, y \in \mathcal{U}, x \neq y$, we have the following.

$$\Pr_{h \text{ picked u.a.r from } \mathcal{H}}[h(x) = h(y)] \leqslant \frac{1}{m}$$

1. Let $\mathcal{H}$ be the set of all functions from $\mathcal{U}$ to $\{0, \ldots, m-1\}$. Prove that $\mathcal{H}$ is universal.

2. Let p be a prime number such that $\mathcal{U} \subseteq \{0, 1, \ldots, p-1\}$. We define

$$\mathcal{H} = \{h_{ab}(k) = ((ak + b) \mod p) \mod m : a \in \{1, \ldots, p-1\}, b \in \{0, \ldots, p-1\}\}$$

   Prove that $\mathcal{H}$ is universal.