Hands On 1: Universal Hash Families

Algorithm Design

Ferraro Domenico
559813

1 Problem

Prove that the family H of functions is **universal** for given $m > 1$ and $p \in [m+1, 2m]$ prime:

$$H = \{\, h_{ab}(x) = (ax + b)\%\, p\, \%\, m, where\ a \in [1, p-1]\ and\ b \in [0, p-1] \,\}$$

That is, for any $k1 \neq k2$, it holds that $|\,\{h \in H : h(k_1) = h(k_2)\}\,| = \frac{|H|}{m}$

Hint: consider first

- $r = (a\, k_1 + b)\, \%\, p$
- $s = (a\, k_2 + b)\, \%\, p$

where $k_1, k_2 \in [0,\ p-1]$

2 Solution

$\mathcal{H}$ is a universal hash family if $\forall k_1, k_2 \in U, k_1 \neq k_2$: $\Pr_{h \in H}[h(k_1) = h(k_2)] \leq \frac{1}{m}$. In other words, any two different keys collide with probability $\frac{1}{m}$ when the hash function is randomly chosen from $\mathcal{H}$.

Consider first that, given $k_1 \neq k_2$, if we have collision, we have that $h(k_1) = h(k_2)$ which can be written as

$$ak_1 + b \equiv\ ak_2 + b + i \cdot m \pmod{p}$$

$$a(k_1 - k_2) \equiv\ i \cdot m \pmod{p}$$

For some integer i between 0 and $p - 1/m$. Because $k_1 \neq k_2$ then their difference is nonzero, so it has an inverse modulo p. If we solve for a we have

$$a \equiv\ i \cdot m(k_1 - k_2)^{-1} \pmod{p}$$

a is nonzero, so there are $p - 1$ possible choices for a. Varying i in its range, $i \cdot m(k_1 - k_2)^{-1}$ has $p - 1/m$ nonzero values. Finally, the collision probability is $(p - 1/m)/(p - 1) = 1/m$.