Hands On 1: Universal Hash Families

Algorithm Design

Ferraro Domenico
559813

1 Problem

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

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

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

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

First consider 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

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

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, varying i in its range, $i \cdot m(k_1 - k_2)^{-1}$ has $\frac{p-1}{m}$ nonzero values. Finally, the collision probability is

$$\frac{\text{\# a and b that give collision}}{\text{\# all the possible a and b}} = \frac{p\left(\frac{p-1}{m}\right)}{p(p-1)} = \frac{1}{m}$$

2.1 A different approach

Consider first that, given $k_1 \neq k_2$, we can define

$$\mathrm{r} = \mathrm{a}k_1 + \mathrm{b} \pmod{\mathrm{p}}$$
$$\mathrm{s} = \mathrm{a}k_2 + \mathrm{b} \pmod{\mathrm{p}}$$

We can demonstrate that the modulo p operation doesn't produce any collisions. We have collision with modulo p if, for $k_1 \neq k_2$, r and s are equal

$$r = s \Leftrightarrow r \equiv \mathrm{s} \pmod{\mathrm{p}} \Leftrightarrow \mathrm{a}k_1 + \mathrm{b} \equiv \mathrm{a}k_2 + \mathrm{b} \pmod{\mathrm{p}} \Leftrightarrow \mathrm{a}(\mathrm{k}_1 - k_2) \equiv 0 \pmod{\mathrm{p}}$$

a is nonzero as well as $\mathrm{k}_1 - k_2$ because $k_1 \neq k_2$, so the above never happen for any a and b. We can also conclude that $\Pr_{h \in \mathcal{H}}[\mathrm{a}k_1 + \mathrm{b} = r \wedge \mathrm{a}k_2 + \mathrm{b} = \mathrm{s}] = \frac{1}{p(p-1)}$.

Finally, we have collision when, for $r \neq s$, we have that $r \equiv \mathrm{s} \pmod{\mathrm{m}}$.

$$\begin{aligned} r \equiv \mathrm{s} \pmod{\mathrm{m}} &\Leftrightarrow \mathrm{r} = \mathrm{s} + \mathrm{i} \cdot \mathrm{m} \pmod{\mathrm{p}} \\ &\Leftrightarrow \mathrm{a}k_1 + \mathrm{b} = \mathrm{a}k_2 + \mathrm{b} + \mathrm{i} \cdot \mathrm{m} \pmod{\mathrm{p}} \\ &\Leftrightarrow \mathrm{a}k_1 = \mathrm{a}k_2 + \mathrm{i} \cdot \mathrm{m} \pmod{\mathrm{p}} \\ &\Leftrightarrow \mathrm{a}(k_1 - k_2) = \mathrm{i} \cdot \mathrm{m} \pmod{\mathrm{p}} \\ &\Leftrightarrow \mathrm{a} = \mathrm{i} \cdot \mathrm{m}(k_1 - k_2)^{-1} \pmod{\mathrm{p}} \end{aligned}$$

For some integer i between 0 and $\frac{p-1}{m}$. Finally, the collision probability is

$$\begin{aligned} \Pr_{h \in \mathcal{H}}[h(k_1) = h(k_2)] &= \Pr_{h \in \mathcal{H}}[\mathrm{a}k_1 + \mathrm{b} = r \wedge \mathrm{a}k_2 + \mathrm{b} = \mathrm{s} \wedge r \neq s \wedge r \equiv \mathrm{s} \pmod{\mathrm{m}}] \\ &= \frac{1}{p(p-1)} * p\left(\frac{p-1}{m}\right) = \frac{1}{m} \end{aligned}$$