1 2-Universal Hashing

Let $\mathcal{H}$ be a class of hash functions in which each $h \in \mathcal{H}$ maps the universe $\mathcal{U}$ of keys to $\{0, 1, \ldots, m-1\}$. Recall that $\mathcal{H}$ is *universal* if for any $x \neq y \in \mathcal{U}$, $\Pr_{h \in \mathcal{H}}[h(x) = h(y)] \leq 1/m$.

We say that $\mathcal{H}$ is 2-universal if, for every fixed pair (x, y) of keys where $x \neq y$, and for any h chosen uniformly at random from $\mathcal{H}$, the pair $(h(x), h(y))$ is equally likely to be any of the m^2 pairs of elements from $\{0, 1, \ldots, m-1\}$. (The probability is taken only over the random choice of the hash function.)

(a) Show that, if $\mathcal{H}$ is 2-universal, then it is universal.

(b) Suppose that you choose a hash function $h \in \mathcal{H}$ uniformly at random. Your friend, who does not know which hash function you picked, tells you a key x, and you tell her $h(x)$. Can your friend tell you $y \neq x$ such that $h(x) = h(y)$ with probability greater than $1/m$ (over your choice of h) if:

(i) $\mathcal{H}$ is universal?

(ii) $\mathcal{H}$ is 2-universal?

In each case, either give a choice of $\mathcal{H}$ which allows your friend to find a collision, or prove that they cannot for any choice of $\mathcal{H}$.

Solution:

(a) If $\mathcal{H}$ is 2-universal, then for every pair of distinct keys x and y, and for every $i \in \{0, 1, \ldots, m-1\}$,

$$\Pr_{h \in \mathcal{H}}[\langle h(x), h(y) \rangle = \langle i, i \rangle] = \frac{1}{m^2}$$

There are exactly m possible ways for us to have x and y collide, i.e., $h(x) = h(y) = i$ for $i \in \{0, 1, \ldots, m-1\}$. Thus,

$$\Pr_{h \in \mathcal{H}}[h(x) = h(y)] = \sum_{i=0}^{m-1} \left(\Pr_{h \in \mathcal{H}}[\langle h(x), h(y) \rangle = \langle i, i \rangle] \right) = \frac{m}{m^2} = \frac{1}{m}$$

Therefore, by definition, $\mathcal{H}$ is universal.

(b) (i) We can construct a scenario where the adversary can force a collision. On a universe $\mathcal{U} = \{x, y, z\}$, consider the following family $\mathcal{H}$:

	x	y	z
h_1	0	0	1
h_2	1	0	1

$\mathcal{H}$ is a universal hash family: x and y collide with probability $1/2$, x and z collide with probability $1/2$, and y and z collide with probability $0 < 1/2$.

The adversary can determine whether we have selected h_1 or h_2 by giving us x to hash. If $h(x) = 0$, then we have chosen h_1, and the adversary then gives us y. Otherwise, if $h(x) = 1$, we have chosen h_2 and the adversary gives us z.

(ii) Suppose that your friend uses the function $f\colon \mathcal{U} \times \{0, \ldots m-1\} \to \mathcal{U}$ to find a collision. We can assume that $f(x,i) \neq x$ for all x, i. The probability that your friend wins is then

$$\Pr_{h \in \mathcal{H}}[h(x) = h(f(x, h(x)))] = \sum_{i=0}^{m-1} \Pr_{h \in \mathcal{H}}[(h(x), h(f(x,i)) = (i,i)] = \frac{1}{m} \ .$$