Theorem \mathscr{H}_{pm} is universal.

Proof. Suppose p and m as specified earlier. Suppose $k, \ell \in Keys$, and $h_{ab} \in \mathscr{H}_{pm}$ (which implies supposing that $a \in [1, p)$ and $b \in [0, p)$). Let $r = (a \cdot k + b) \mod p$ and $s = (a \cdot \ell + b) \mod p$ Subtracting gives us

$$r-s \equiv (a \cdot k + b) - (a \cdot \ell + b) \mod p$$

 $\equiv a \cdot (k - \ell) \mod p$

Now a cannot be 0 because $a \in [1, p)$. Similarly $k - \ell$ cannot be 0, since $k \neq \ell$. Hence $a \cdot (k - \ell) \neq 0$. Since p is prime and greater than a, k, and ℓ , it cannot be a factor of $a \cdot (k - \ell)$. In other words, $a \cdot (k - \ell) \mod p \neq 0$. By substitution, $r - s \neq 0$, and so $r \neq s$. By another substitution, $(a \cdot k + b) \mod p \neq (a \cdot \ell + b) \mod p$.

▲ロト ▲圖ト ▲画ト ▲画ト 三回 - のへで

Define the following function, given k and ℓ , which maps from (a, b) pairs to (r, s) pairs (formally, $[1, p) \times [0, p) \rightarrow [1, p) \times [0, p)$):

$$\phi_{k\ell}(a,b) = ((a \cdot k + b) \mod p, (a \cdot \ell + b) \mod p)$$

Now consider the inverse of that function.

$$\phi_{k\ell}^{-1}(r,s) = (((r-s) \cdot (k-\ell)^{-1}) \mod p), (r-ak) \mod p)$$

= (a,b)

The existence of ϕ^{-1} implies that ϕ is a one-to-one correspondence. Hence for each (a, b) pair, there is a unique (r, s) pair. Since the pair (a, b) specifies a hash function, that means that for each hash function in the family \mathscr{H}_{pm} , there is a unique (r, s) pair.

▲□▶ ▲圖▶ ▲필▶ ▲필▶ 三里

There are p-1 possible choices for a and p choices for b, so there are $p \cdot (p-1)$ hash functions in family \mathscr{H}_{pm} . Likewise there are p choices for r, and for each r there are p-1 choices for s (since $s \neq r$). Thus we can partition the set \mathscr{H}_{pm} into p subsets by r value, each subset having p-1 hash functions.

For a given r, at most one out of every m can have an s that is equivalent to r mod m, in other words, at most $\frac{p-1}{m}$ hash functions. Now sum that for all p of the subsets of \mathscr{H}_{pm} , and we find that the number of hash functions for which k and ℓ collide are

$$p \cdot \frac{p-1}{m} = \frac{p \cdot (p-1)}{m} = \frac{|\mathscr{H}_{pm}|}{m}$$

▲ロト ▲圖ト ▲画ト ▲画ト 三回 - のへで

Therefore \mathscr{H}_{pm} is universal by definition. \Box

Theorem [Probability of any collisions.] If Keys is a set of keys, $m = |Keys|^2$, p is a prime greater than all keys, and $h \in \mathscr{H}_{pm}$, then the probability that any two distinct keys collide in h is less than $\frac{1}{2}$.

Proof. Suppose we have a set Keys, $m = |Keys|^2$, p is a prime greater than all keys, and $h \in \mathscr{H}_{pm}$. Consider the number of pairs of unique keys. The number of pairs of keys is

$$\binom{n}{2} = \frac{n!}{2! \cdot (n-2)!} = \frac{n!}{2 \cdot (n-2)!} = \frac{n \cdot (n-1) \cdot (n-2)!}{2 \cdot (n-2)!} = \frac{n \cdot (n-1)}{2}$$

▲□▶ ▲圖▶ ▲≣▶ ▲≣▶ = 三 ∽੧<⊙

Since \mathscr{H}_{pm} is universal, each pair collides with probability $\frac{1}{m}$. Multiply that by the number of pairs, and the expected number of collisions is

$$\frac{n \cdot (n-1)}{2} \cdot \frac{1}{m} < \frac{n^2}{2} \cdot \frac{1}{m} \quad \text{since } n \cdot (n-1) < n^2$$
$$= \frac{n^2}{2} \cdot \frac{1}{n^2} \quad \text{since } m = n^2$$
$$= \frac{1}{2} \qquad \text{by cancelling } n^2$$

With the expected number of collisions less than one half, the probability there are any collisions is also less than $\frac{1}{2}$. \Box

(日) (문) (문) (문) (문)

