6.857 Homework Changping Chen Peinan Chenn Skanda Koppula	**Problem Set 2**	# 2-2 Hash Functions March 8, 2015

(a) $E[collisions] = \sum P(h(x_i) = P(x_j)) = \binom{n}{2} \cdot 2^{-d}$. When n, the number of hashes we compute, is $c \cdot 2^{d/2}$:

$$E[collisions] = \binom{c \cdot 2^{d/2}}{2} \cdot 2^{-d}$$

This approximates to $\frac{(c \cdot 2^{d/2})^2}{2} \cdot 2^{-d} = \frac{c^2}{2}$.

(b) XOR'ing two inputs does not change one-way of the function. $x \oplus y$ distributes evenly across the input space $\{0,1\}^n$ of $h(x)$; if finding n given $h(n)$ is is worst case $O(2^d)$ operations, finding $x \oplus y$ (and thus x, y) given $h'(x, y) = h(x \oplus y)$, is still $O(2^d)$.

However, in the case of AND'ing the two inputs, $x \wedge y$ distributes unevenly across $\{0,1\}^n$; for example, given a random x, y, our input to $h(n)$ is much more likely to have be entirely 0's than entirely 1's. Because our input space collapsed in a certain direction, iterating through x, y to find our $h'(x, y)$ no longer is expected $\Theta(2^d)$.

(c) Not collision resistant. For any x_1, x_2, pick y_1, y_2 such that $y_1 = h(x_2)$ and $y_2 = h(x_1)$.

$$h'(x_1, y_1) = h(x_1) \oplus y_1 = h(x_1) \oplus h(x_2) = h(x_2) \oplus h(x_1) = h(x_2) + y_2 = h'(x_2, y_2)$$

We have a collision.

(d) Not weak collision resistant. Because $h(x)$ is only TCR, we can assume that finding x_1, x_2 s.t. $h(x_1) = h(x_2)$ is easy. That means for target $h'(x, y) = 0$, finding colliding x, y is easy with the aforementioned pairs because $h(x_1) \oplus h(x_2) = h(x_1) \oplus h(x_1) = 0$. More generally, for any target $h'(x, y)$ and input (x, y), the input (y, x) always collides with the target, because of the commutivity of XOR.