MERKLE DAMGARD TRANSFORMATION CONCEPTS

Overview:

The Merkle–Damgård transform allows us to convert a fixed-length hash functions to handle inputs of arbitrary length.
This approach for domain extension of hash functions has been used frequently in practice.

Code Construction:

CONSTRUCTION 6.3

Let (Gen, h) be a compression function for inputs of length $n + n' \geq 2n$ with output length n. Fix $\ell \leq n'$ and $IV \in \{0,1\}^n$. Construct hash function (Gen, H) as follows:

- Gen: remains unchanged.
- H: on input a key s and a string $x \in \{0,1\}^*$ of length $L < 2^\ell$, do:
  1. Append a 1 to x, followed by enough zeros so that the length of the resulting string is ℓ less than a multiple of n'. Then append L, encoded as an ℓ-bit string. Parse the resulting string as the sequence of n'-bit blocks $x_1, \ldots, x_B$.
  2. Set $z_0 := IV$.
  3. For $i = 1, \ldots, B$, compute $z_i := h^s(z_{i-1} \| x_i)$.
  4. Output z_B.

Proof:

THEOREM 6.4 *If* (Gen, h) *is collision resistant, then so is* (Gen, H)*.*

PROOF We show that for any s, a collision in H^s yields a collision in h^s. Let x and x' be two different strings of length L and L', respectively, such that $H^s(x) = H^s(x')$. Let $x_1, \ldots, x_B$ be the B blocks of the padded x, and let $x'_1, \ldots, x'_{B'}$ be the B' blocks of the padded x'. Let $z_0, z_1, \ldots, z_B$ (resp., $z'_0, z'_1, \ldots, z'_{B'}$) be the intermediate results during computation of $H^s(x)$ (resp., $H^s(x')$). There are two cases to consider:

Case 1: $L \neq L'$. In this case, the last step of the computation of $H^s(x)$ is $z_B := h^s(z_{B-1} \| x_B)$, and the last step of the computation of $H^s(x')$ is $z'_{B'} := h^s(z'_{B'-1} \| x'_{B'})$. Since $H^s(x) = H^s(x')$ we have $h^s(z_{B-1} \| x_B) = h^s(z'_{B'-1} \| x'_{B'})$. However, $L \neq L'$ and so $x_B \neq x'_{B'}$. (Recall that the last ℓ bits of x_B encode L, and the last ℓ bits of $x'_{B'}$ encode L'.) Thus, $z_{B-1} \| x_B$ and $z'_{B'-1} \| x'_{B'}$ are a collision with respect to h^s.

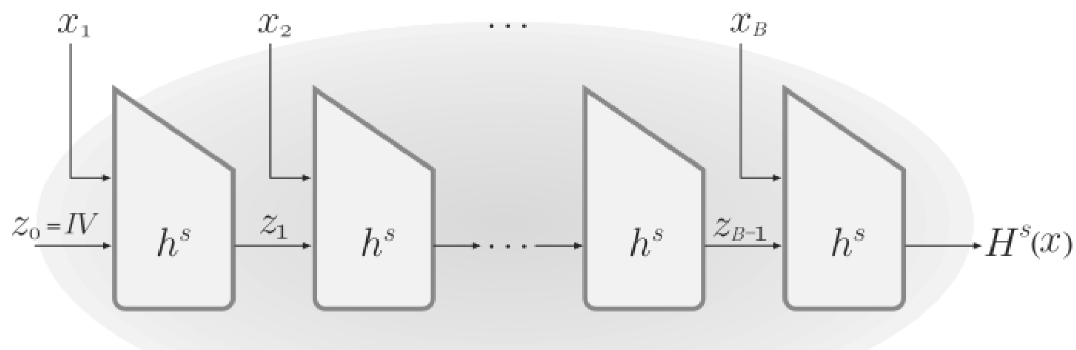

FIGURE 6.1: The Merkle–Damgård transform.

Case 2: $L = L'$. This means that $B = B'$. Let $I_i \stackrel{\text{def}}{=} z_{i-1} \| x_i$ denote the ith input to h^s during computation of $H^s(x)$, and define $I_{B+1} \stackrel{\text{def}}{=} z_B$. Define $I'_1, \ldots, I'_{B+1}$ analogously with respect to x'. Let N be the largest index for which $I_N \neq I'_N$. Since $|x| = |x'|$ but $x \neq x'$, there is an i with $x_i \neq x'_i$ and so such an N certainly exists. Because

$$I_{B+1} = z_B = H^s(x) = H^s(x') = z'_B = I'_{B+1},$$

we have $N \leq B$. By maximality of N, we have $I_{N+1} = I'_{N+1}$ and in particular $z_N = z'_N$. But this means that I_N, I'_N collide under h^s.