Merkle-Damgård, Length Extension

Application of hash functions:

- MAC is a MAC scheme for $\{0,1\}^n$
- Define new MAC scheme for $\{0,1\}^*$ Hash-then-MAC

$$MAC^*(k, m \in \{0,1\}^*) = MAC(k, H(m))$$

Claim: If underlying MAC is secure, and H is collision-resistant, then MAC^* is secure

Idea: Adv (attacking MAC^*) sees

$$t_1 = MAC^*(k, m_1) = MAC(k, H(m_1))$$
$$\vdots$$
$$t_q = MAC^*(k, m_q) = MAC(k, H(m_q))$$

Suppose Adv produces forgery: (m^*, t^*)

$$t^* = MAC^*(k, m^*) = MAC(k, H(m^*))$$
where $m^* \notin \{m_1, \ldots, m_q\}$

Case 1: $H(m^*) = H(m_i)$ for some m_i
$\Rightarrow$ Adv found collision in H

Case 2: $H(m^*) \notin \{H(m_1), \ldots, H(m_q)\}$

$\Rightarrow$ Adv knows $MAC(k, H(m^*))$
after seeing $MAC(k, x)$
for many $x \neq H(m^*)$

} forgery of underlying MAC

Common Pitfall:

Construct $MAC(k, m) = H(k \| m)$

(secure if $H =$ SHA-3, insecure if $H =$ MD5, SHA-1)

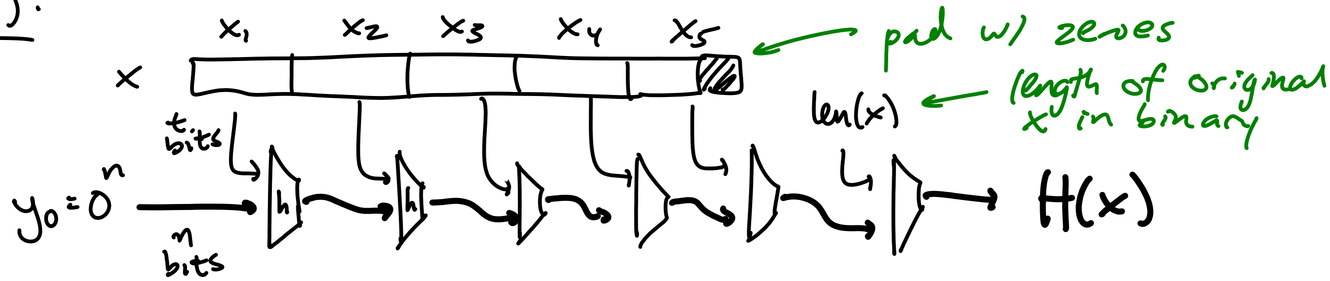
Merkle-Damgård Construction:

- Design a "compression function" $h : \{0,1\}^{n+t} \to \{0,1\}^n$

  (not "compression" like ZIP, gzip, etc)
- Extend h to a full-fledged Hash function:

each t bits

$H(x)$:



Security: If h is collision-resistant, then H is too

Idea: If $H(x) = H(x')$ then within these 2 computations, a collision under h is guaranteed to exist

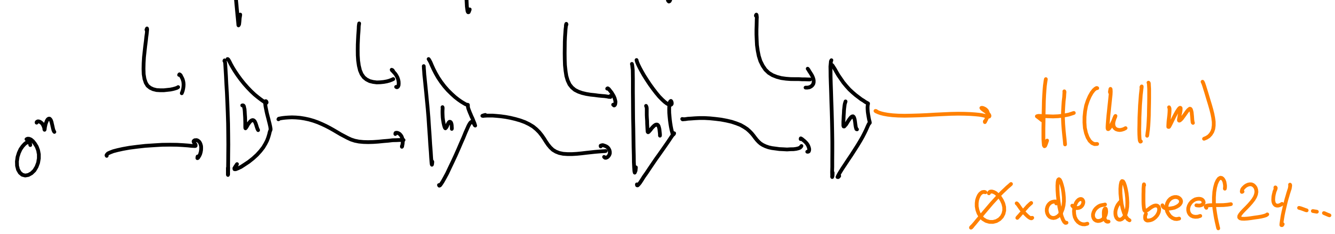
Ex: Suppose $MAC(k, m) = H(k \| m)$, where H is Merkle-Damgård

$n = 16$ bytes
$t = 2$ bytes

key k = e93a3527 (4 bytes)
m = 7d6 (12 bits)

$H(k \| m)$: $\quad$ len$(k \| m)$ = 44 bits = "2c" in hex

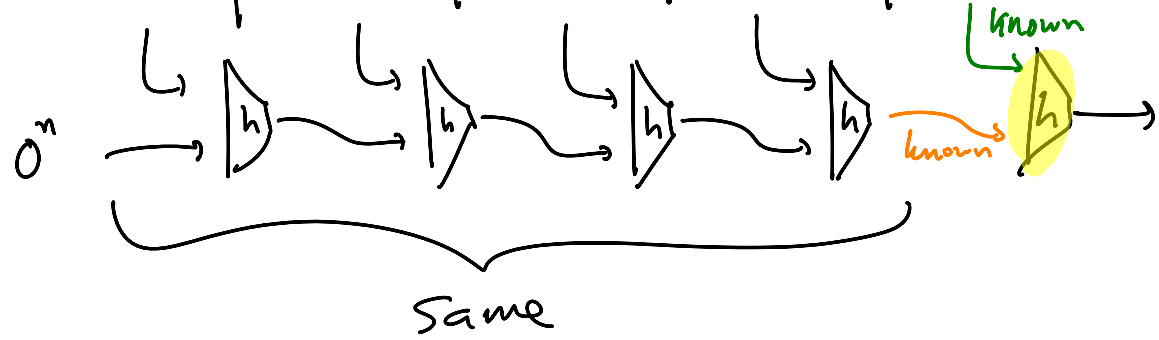
e93a | 3527 | 7d6∅ | ∅∅2c



Q: How is $H(k \| m)$ related to $H(k \| m \| 002c)$?

$H(k \| m \| 002c)$: $\quad$ len(...) = 64 bits = "40" hex

e93a | 3527 | 7d60 | 002c | 0040



A:

Q: Do you need to know k to predict $H(k \| m \| 002c)$ given $H(k \| m)$?

A: No, $\quad H(k \| m \| 002c) = h(0040 \| H(k \| m))$

Length extension attack!

Ask for MAC of m

predict MAC of $m \| 002c \| \text{anything}$