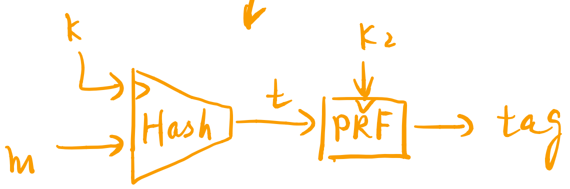
Motivation: (previously: PRF to construct MAC), now a general "hash-then-PRF" paradigm using "universal hash function" (UHF)

satisfy "weak collision resistant"

adversary knows nothing about the key.



UHF Attack Game:

chal   adv

$k \xleftarrow{R} \mathcal{K}$

$\xleftarrow{m_0, m_1 \in \mathcal{M}}$

$$\text{UHFadv}[\mathcal{A}, H] = \Pr[H(k, m_0) = H(k, m_1)]$$

$\text{UHFadv}[\mathcal{A}, H] \le \epsilon$

ϵ-UHF } unbounded

statistical-UHF

ϵ is negl.

computational UHF

adv is efficient & ϵ is negl.

Multi-query UHF (MUHF):

chal   adv

$k \xleftarrow{R} \mathcal{K}$

$\xleftarrow[s \le Q]{m_0, \cdots m_s \in \mathcal{M}}$

$$\text{MUHFadv}[\mathcal{A}, H] \le \frac{Q^2}{2} \cdot \text{UHFadv}[\mathcal{B}, H]$$

Constructing statistical UHF using polynomials:

$$H_{poly}(k, (a_1, \cdots, a_v)) := k^v + a_1 \cdot k^{v-1} + a_2 k^{v-2} + \cdots + a_{v-1} k + a_v \in \mathbb{Z}_p$$

NOTE: 1. evaluation w/o knowledge of len(m) ahead of time

Horner's method:

Input: $m = (a_1, \cdots, a_v)$, $k \in \mathbb{Z}_p$

Output: $t := H_{poly}(k, m)$

Set $t \leftarrow 1$

For $i \leftarrow 1$ to v:

$\quad t \leftarrow t \cdot k + a_i \in \mathbb{Z}_p$

Output t

4-way parallel:

$\Longrightarrow$ For $i = 1$ to v, increment i by 4

$t \leftarrow t \cdot k^4 + a_i \cdot k^3 + a_{i+1} \cdot k^2 + \cdots + a_{i+3} \in \mathbb{Z}_p$

2. H_{poly} over $(\mathbb{Z}_p, \mathbb{Z}_p^{\le l}, \mathbb{Z}_p)$ is a (l/p)-UHF.

3. the leading term k^v is necessary to be UHF.
   (counter-example: $m_0 = (a_1, a_2)$, $m_2 = (0, a_1, a_2)$ are collisions)
   if we restrict message to fixed length, then H_{fpoly} is $(\frac{l-1}{p})$-UHF

4. just by adapting H_{poly} from $\mathbb{Z}_p$ to $GF(2^n)$ would result in an insecure UHF
   ↳ see Ex7.1. But why? What fundamentally about $GF(2^n)$ makes it unsuitable?

5. revealing few points about the function could recover the key.

Constructing Computational UHF using CBC & Cascade:

\# 1st theorem: prefix-free, extendable secure PRF is a computational UHF,

and

$$UHF adv[\mathcal{A}, PF] \le PRF^{pf} adv[\mathcal{B}, PF] + \frac{1}{|\mathcal{Y}|}$$

PF: over $(\mathcal{K}, \mathcal{X}^{\le l+1}, \mathcal{Y})$, UHF: over $(\mathcal{K}, \mathcal{X}^{\le l}, \mathcal{Y})$

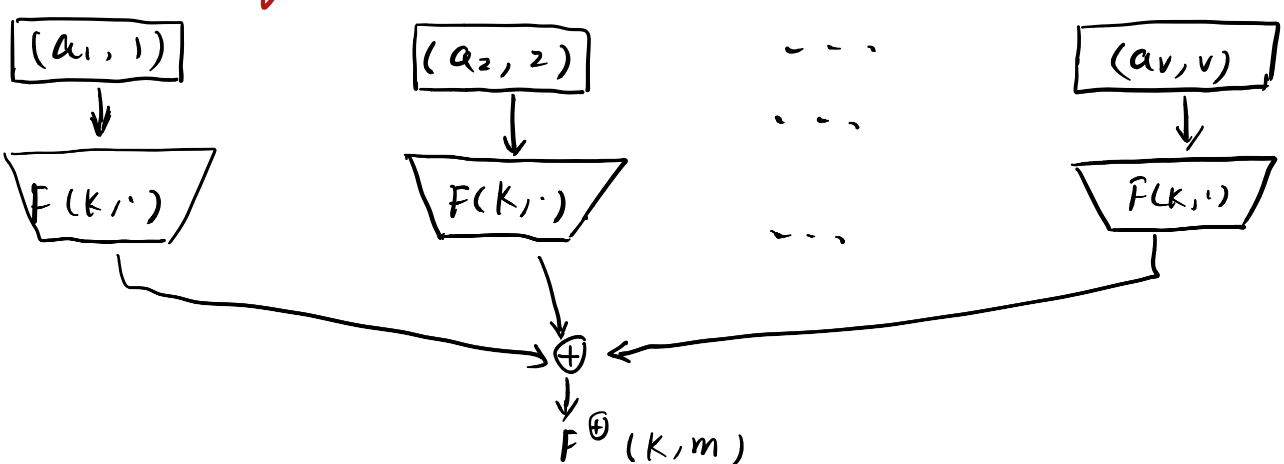
NOTE: require extra 1 block to build adv. $\mathcal{B}$ since its queries to PRF chal. have to be "prefix-free", whereas $\mathcal{A}$ to $\mathcal{B}$ (or $\mathcal{A}$'s chal.) doesn't have to. Thus by padding 1 block to achieve prefix-free.

\# 2nd theorem: PF is also a multi-query UHF,

$$MUHF adv[\mathcal{A}, PF] \le PRF^{pf} adv[\mathcal{B}, PF] + \frac{Q^2}{2|\mathcal{Y}|}$$

unique number of pairs/possibilities: $\frac{Q(Q-1)}{2}$

Constructing parallel UHF:



$$\mathrm{UHFadv}[\mathcal{A}, F^{\oplus}] \leq \mathrm{PRFadv}[\mathcal{B}, F] + \frac{1}{|\mathcal{Y}|}$$