We'll go with secret-key approach (as this is what's used in practice).

DEF (ALMOST UNIVERSAL) Family $\mathcal{H}$ is ε-AU if: $\forall\ x, x' \in \{0,1\}^{N}$ s.t. $x \neq x'$:

$$\Pr_{S}\left[h_S(x) = h_S(x') \right] \leq \varepsilon.$$

(This is strong: it implies no A can find a collision!).

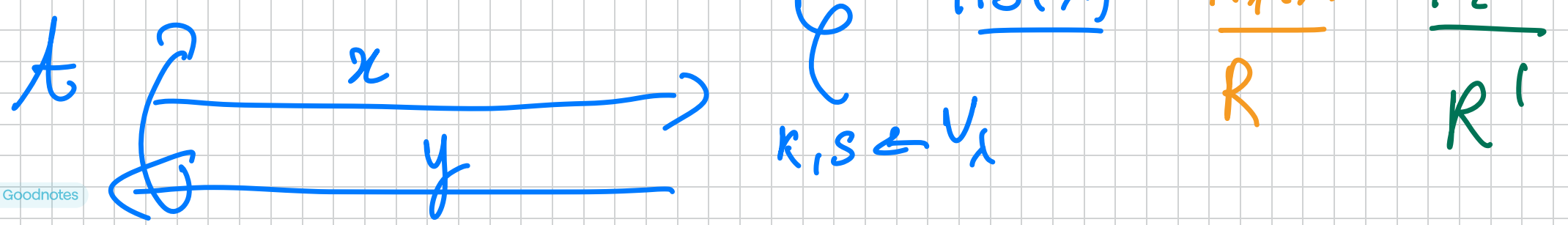
THM Assuming $\mathcal{F} = \{F_K : \{0,1\}^{m} \to \{0,1\}^{n}\}_K$ is a PRF family, and $\mathcal{H}$ is ε-AU for $\varepsilon = \text{negl}(\lambda)$, then

$$\mathcal{F}(H) = \{ F_k(h_s(\cdot)) \}_{(k,s)}$$

is a PRF family.

Cor $\mathcal{F}(H)$ is also UF-CMA MAC for FIL msgs of length $N >> n$.

Proof (of Thm). We need to show that $F_k(h_s(\cdot))$ is $\approx_c$ from $R^I : \{0,1\}^N \to \{0,1\}^n$ (random table).



b' ⟶ $y = f_k(h_s(x))$ $\quad y = R'(x)$

$y = R(h_s(x))$

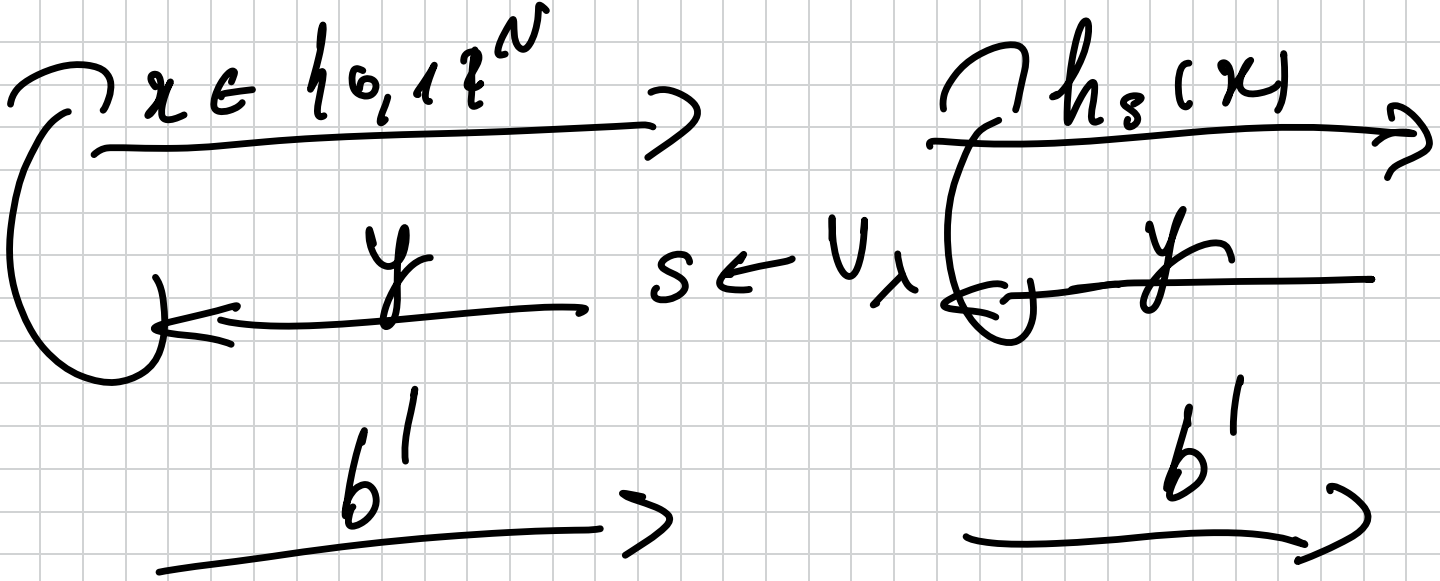
$R : \{0,1\}^n \to \{0,1\}^n$ and n a random table.

$R' : \{0,1\}^N \to \{0,1\}^n$ and N a random table.

LEMMA $H_0(\lambda) \approx_c H_1(\lambda)$.

Proof. Assume not: $\exists$ PPTA that can distinguish H_0, H_1. Build reduction $\mathcal{B}$ against security of f:

A ⟵ START $\mathcal{B}$ $\quad \mathcal{C}_{PRF}$



Analysis: By inspection. ∎

LEMMA $H_1(\lambda) \approx_s H_2(\lambda)$ so long as A asks $q = \text{poly}(\lambda)$ queries.

Proof. Define bad event E:

E: Becomes true if $\exists\, w, i$ s.t. $w \neq i$

$$h_s(x_i) = h_s(x_j)$$

$x_1, \ldots, x_q$ are the queries

Now, if $\overline{E}$ then $H_1(\lambda) \equiv H_2(\lambda)$ because $R(\cdot)$ is computed on the same values $y_1, y_2, y_3, \ldots, y_q$.

$$\Rightarrow SD(H_1(\lambda); H_2(\lambda)) \leq Pr[E]$$

We'd like to see:

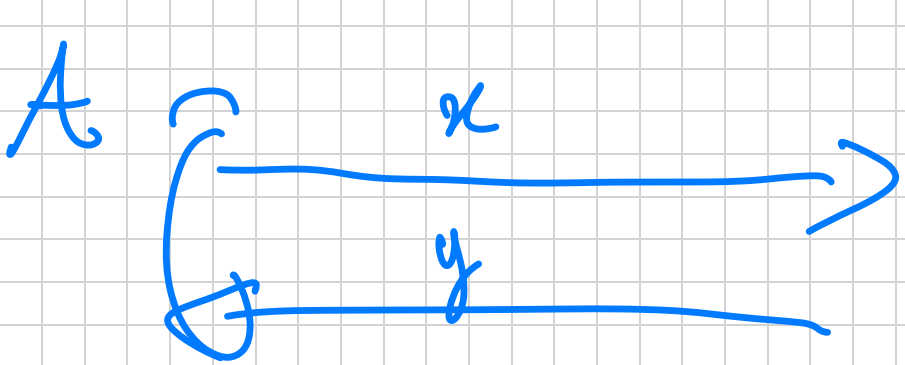
$$Pr[E] = Pr_s[\exists\, v, i : v \neq i \text{ and } h_s(x_i) = h_s(x_i)]$$

$$\leq \sum_{\substack{v,i \\ i \neq j}} Pr[h_s(x_i) = h_s(x_j)]$$

$$\leq \binom{q}{2} \cdot \varepsilon \leq \text{poly}(\lambda) \cdot \text{negl}(\lambda)$$ ∎

Issue: AV requires x to be independent of x. Not clear why/how for our $\mathcal{E}$.

Equivalent definition of $\mathcal{E}$:



$\mathcal{C}$

$y_1 = R(x_1)$

$\vdots$

$y_q = R(x_q)$

$\mathcal{E}'$: Sample $s \in U_s$, check if $\exists\, i, j$ $i \neq j$ s.t. $h_s(x_i) = h_s(x_j)$

The hash function:

1) The inner product. We have $N = n \cdot l$ of blocks of length n bits:

$$m = m_1, m_2, \dots, m_l$$

$m_i \in GF(2^n) = \mathbb{F}$ The finite field with 2^n elements. $(\mathbb{F}, +, \cdot)$ a field.

$s = (s_1, \dots, s_l)$; $s_i \in \mathbb{F}$

$$h_s(m) = \sum_{i=1}^{l} s_i \cdot m_i \in \mathbb{F}$$

where $\sum$ and $\cdot$ are over the field $\mathbb{F}$.

Why this works? Fix m_1, m' s.t.

$m \neq m'$ and $m = (m_1 \dots m_d)$

$m' = (m'_1 \dots m'_d)$

(wlog. $m_1 \neq m'_1$)

$$h_s(m) = h_s(m').$$

$$\Leftrightarrow \sum_i s_i \cdot m_i = \sum_i s_i \cdot m'_i$$

$$\Leftrightarrow s_1 (m_1 - m'_1) = \sum_{i=2}^{d} s_i \cdot (m'_i - m_i)$$

$$\Leftrightarrow s_1 = (m_1 - m'_1)^{-1} \cdot \sum_{i=2}^{d} s_i (m'_i - m_i)$$

$$\Rightarrow \Pr\left[h_s(m) = h_s(m')\right] \leq 2^{-n}$$

(PERFECT) UNIVERSAL

2) The above has very good $\varepsilon = 2^{-n}$, but $|S| = |M|$. Take $\mathbb{F} = GF(2^n)$.

$$M = (m_1, \ldots, m_d) \; ; \; m_i \in \mathbb{F}$$

$$s \in \mathbb{F}$$

Think of m_i as the coefficients of some polynomial and evaluate at s:

$$h_s(m) = \sum_{i=1}^{d} m_i \cdot s^{i-1} = q_m(s).$$

It works too:

$$h_s(m) = h_s(m') \iff q_m(s) = q_{m'}(s)$$

$$\iff q_m(s) - q_{m'}(s) = 0$$

$$\iff q_{m-m'}(s) = 0$$

$$\iff \sum_{i=1}^{d} (m_i - m_i')\, s^{i-1} = 0$$

$\Rightarrow$ Collision $\iff$ s is a root of the above polynomial with $\mathbb{F}$-coefficients.

The # roots is $d-1$

$$\Rightarrow \Pr[h_s(m) = h_s(m')] \leq \frac{d-1}{|\mathbb{F}|}$$

d is at most $poly(\lambda) \leq negl(\lambda)$

$|\mathbb{F}| \geq 2^n$; $n = poly(\lambda)$

3) In practice, even more efficient.
But some differences: (i) Only computational AU (not statistical) using the same PRF family $\mathcal{F}: (\mathcal{K}, \mathcal{M}) \to \mathcal{R}$ as MAC of $\mathcal{F}$.

Computational AU: $\forall$ PPT A:

$A \xrightarrow{x, x' \in \{0,1\}^N} \mathcal{C}$

$s \leftarrow U_\lambda$

Output 1 iff $x \neq x'$ and $h_s(x) = h_s(x')$.

Use some other PRF instantiation $F_s(\cdot)$ to construct $h_s(\cdot)$.

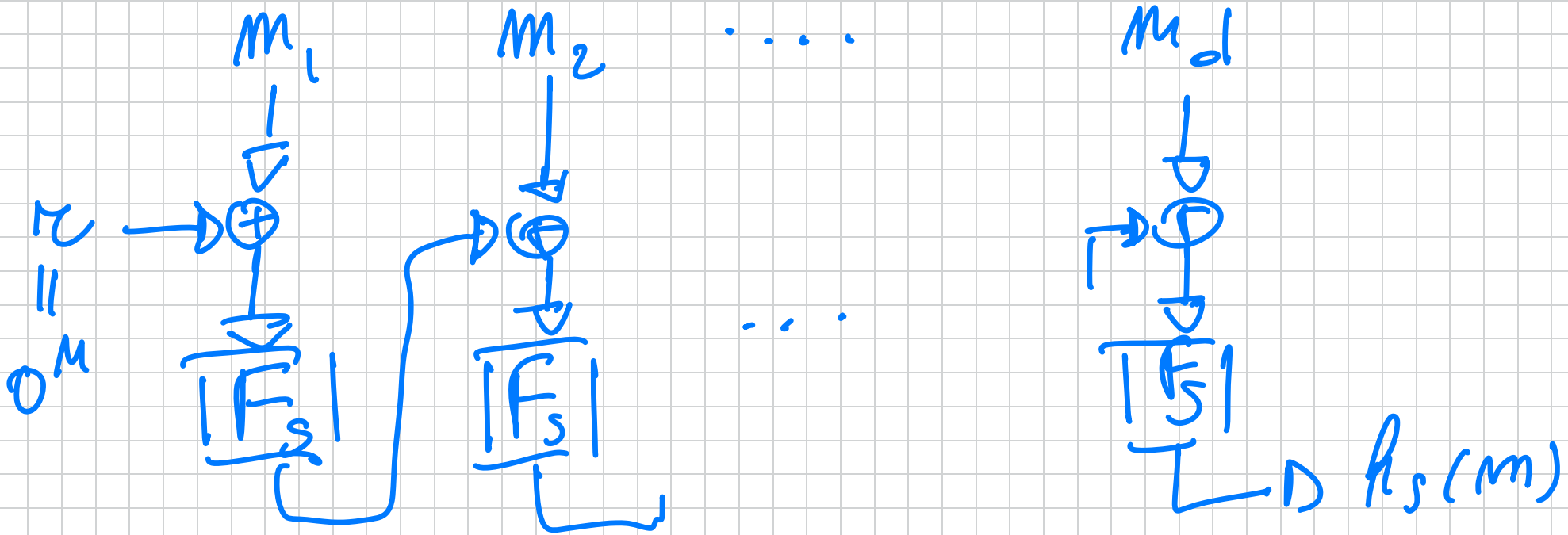
$$F_K(h_s(\cdot))$$

Optimization trick: instead of using k, s use just k (or s) and do

$F_k(0 \| \cdot)$ for $f_k(\cdot)$

$F_k(1 \| \cdot)$ for $f_s(\cdot)$.)

CBC-MAC



$$h_S(m_1 \cdots m_d) = F_S(m_d \oplus F_S(m_{d-1} \oplus \cdots \oplus F_S(m_2 \oplus F_S(m_1))\cdots))$$

THM If F a PRF then above $h_S(\cdot)$ is computational AU.

$\Rightarrow$ "Encrypted" CBC-MAC:

$$F_K(h_S(m_1))$$

THM In fact, E-CBC-MAC is UF CMA also for VIL.

XOR MAC

Instead of doing $F(r, v)$, you do: $(r, F_k(r) \oplus h_s(m))$ for random r. Actually this construction only gives a MAC and not a PRF.

Q: What h? A $\times$ U (Almost XOR UNIVERSAL): $\forall \varepsilon \in [0,1]$, $\forall m \neq m'$

$$\Pr_s [h_s(m) \oplus h_s(m') = \varepsilon] \leq \varepsilon.$$

ε-A $\times$ U

$AU \geq \varepsilon = 0^n$

Why? Because given $m, \tau = (r, v)$
Adv can output $m', \tau' = (r, v \oplus e)$
If $h_s(m) \oplus e = h_s(m')$ then we
e valid tag! Then, define h_s:

$$h_s(m_1 \cdots m_d) = F_s(m_1 \| 1) \oplus F_s(m_2 \| 2) \oplus \cdots \oplus F_s(m_d \| d \| 1)$$

Thm. Assuming F e PRF, the above
H is a CRH.

For VLL :

- E-CBC-MAC
- XOR MAC