MPRI SECURE: Proofs of Security Protocols
TD

Adrien Koutsos

2025/2026

Questions marked with a star ($\star$) can be skipped without impacting the rest of the exercise.

1 Negligibility

Question 1. *Show the following properties:*

- *If* $f \in \mathsf{negl}(\eta)$ *and* $g \in \mathsf{negl}(\eta)$ *then* $(f+g) \in \mathsf{negl}(\eta)$.
- *Idem, but for* $\max(f,g)$ *and* $\min(f,g)$.
- *Let* P *be a polynomial. If, for every* $1 \leq i \leq P(\eta)$, $f_i \in \mathsf{negl}(\eta)$, *then* $\sum_{1 \leq i \leq P(\eta)} f_i$ *is* ***not*** *necessarily negligible.*
- *Show that* $\sum_{1 \leq i \leq P(\eta)} f_i$ *is negligible if there exists* $f \in \mathsf{negl}(\eta)$ *uniformly bounding the* f_i*'s, i.e. s.t.* $f_i(\eta) \leq f(\eta)$ *for every* i, η.

2 Bi-Deduction

A *n-context* is a term C using distinguished variables $[]_1, \ldots, []_n$ called *holes*. If C is a n-context and $t_1, \ldots, t_n$ are terms, then $C[t_1, \ldots, t_n]$ is the term obtained by simultaneously substituting all variables $[]_i$ by t_i in C, i.e.

$$C[t_1, \ldots, t_n] \stackrel{\text{def}}{=} C\{[]_1 \mapsto t_1, \ldots, []_n \mapsto t_n\}$$

Example: if $C = \langle []_1, \langle []_1, []_2 \rangle \rangle$ then $C[a,b] = \langle a, \langle a, b \rangle \rangle$.

Consider the following rule schema:

$$\textsc{BiDeduce} \quad \frac{\vec{u} \sim \vec{v}}{C_1[\vec{u}], \ldots, C_l[\vec{u}] \sim C_1[\vec{v}], \ldots, C_l[\vec{v}]}$$

where $l \in \mathbb{N}$, $\vec{u}$ and $\vec{v}$ are vectors of terms of length n, and $C_1, \ldots, C_l$ are n-contexts.

Question 2. *Give an unsound instance of the* BiDeduce *rule schemata. Argue why your instance is unsound.*

Solution. E.g. take $l = 1$, $u = \mathsf{n}$, $v = \mathsf{n}'$ (where n and n' are two names), and $C_1 = \langle [], \mathsf{n} \rangle$. We clearly have that $u \sim v$ is valid, but $C[u] \sim C[v]$ is the formula

$$\langle \mathsf{n}, \mathsf{n} \rangle \sim \langle \mathsf{n}, \mathsf{n}' \rangle$$

which is not invalid, because the following program is a winning distinguisher with high probability:

$$\mathcal{B}(x) := \texttt{return } \pi_1(x) = \pi_2(x)$$

(as $\mathsf{n} \neq \mathsf{n}'$ with probability close to 1). ■

Question 3. *Give sufficient conditions on* $C_1, \ldots, C_l$ *under which an instance the* BiDeduce *rule schemata is sound. Care will be taken to restrict the rule applicability as little as possible.*

Solution. It is sufficient to require that:

- the only variables appearing in $C_1, \ldots, C_l$ are the hole variables $[]_1, \ldots, []_n$:

$$\mathsf{st}(C_1, \ldots, C_l) \cap \mathcal{X} \subseteq \{[]_1, \ldots, []_n\}$$

- the names of $\vec{u}$ and $\vec{v}$ do not occur in the context $C_1, \ldots, C_l$, i.e. if we let $\mathsf{names}(u) = \mathsf{st}(u) \cap \mathcal{N}$ then:

$$\mathsf{names}(\vec{u}, \vec{v}) \cap \mathsf{st}(C_1, \ldots, C_l) = \emptyset$$ ■

Question 4. *Prove that your restricted rule is sound using the usual rules of the logic.*

If your proof proceed by induction, your answer should include at least a precise statement of the induction hypothesis, as well as the key rules used in the inductive step.

Solution. We let $|t|$ be the size of a term t, i.e. the number of nodes in t (seeing a term as a tree).

We consider the following generalized rule:

$$\textsc{BiDeduce-G}\quad \frac{}{\begin{array}{l}\vec{u}, \mathsf{n}_1, \ldots, \mathsf{n}_n, C_1[\vec{u}], \ldots, C_l[\vec{u}] \\ \sim\ \vec{v}, \mathsf{n}_1, \ldots, \mathsf{n}_n, C_1[\vec{v}], \ldots, C_l[\vec{v}]\end{array}}$$

under the same conditions on $C_1, \ldots, C_l$ as BiDeduce, and where $\mathsf{n}_1, \ldots, \mathsf{n}_n$ is a sequence of distinct names containing at least all the names occurring in the contexts $C_1, \ldots, C_l$.

Assuming that the BiDeduce-G rules are sound, the BiDeduce rule schemata is sound (using the Restr rule).

Thus, it remains to prove that any instance of the BiDeduce-G rule is sound, which we do by induction on $|C_1| + \cdots + |C_l|$.

- If $l = 0$, the indistinguishability is exactly the premise of the BiDeduce rule extended with a finite sequence of distinct names that are different from the names of $\vec{u}$ and $\vec{v}$, and thus holds (using Fresh once for each name in $\mathsf{n}_1, \ldots, \mathsf{n}_n$).
- Take $l > 0$, we do a case-analysis on the root symbol of C_l:
  - **variable case**, $C_l = x$ where $x \in \mathcal{X}$. Using our restriction, we know that x is one of the hole variable $[]_i$. Then, $C_l[\vec{u}] = u_i$ and $C_l[\vec{v}] = v_i$, where u_i and v_i are the i-th term of, respectively, $\vec{u}$ and $\vec{v}$. As this is a duplicated entry, we conclude apply Dup and then the induction hypothesis.
  - **function symbol**, $C_l = f(D_1, ..., D_k)$ where $D_1, ..., D_k$ are contexts. We apply the FA rule to remove f. This yields a smaller rule (since we remove the f node), and $D_1, \ldots, D_k$ verify the same conditions as $C_1, \ldots, C_l$ (the conditions are stable by subterms). Thus, we conclude using the induction hypothesis.
  - **name symbol** $C_l = \mathsf{n}$ where $\mathsf{n} \in \mathcal{N}$. Using our restrictions, we know that n occurs among the names $\mathsf{n}_1, \ldots, \mathsf{n}_n$. Thus, we remove n using Dup, and conclude using the induction hypothesis. ■

3 Probabilistic Couplings

Reminder For any model $\mathbb{M}$, recall that $\mathbb{T}_{\mathbb{M},\eta}$ defines the randomness source used to give the semantics of a term. More precisely, for every η, we have $\mathbb{T}_{\mathbb{M},\eta} = \mathbb{T}^{\mathsf{a}}_{\mathbb{M},\eta} \times \mathbb{T}^{\mathsf{h}}_{\mathbb{M},\eta}$ where $\mathbb{T}^{\mathsf{a}}_{\mathbb{M},\eta}$ and $\mathbb{T}^{\mathsf{h}}_{\mathbb{M},\eta}$ are, respectively, the *adversary* and *honest* randomness. The adversary randomness $\mathbb{T}^{\mathsf{a}}_{\mathbb{M},\eta}$ must be of the form $\{0;1\}^{\mathsf{a}_\eta}$, i.e. the set of bitstrings of length a_η. Similarly, $\mathbb{T}^{\mathsf{a}}_{\mathbb{M},\eta}$ must be of the form $\{0;1\}^{\mathsf{h}_\eta}$. Finally, the semantics of the logic $[\![t]\!]^{\eta,\rho}_{\mathbb{M}}$ evaluates t using randomness $\rho = (\rho_{\mathsf{a}}, \rho_{\mathsf{h}}) \in \mathbb{T}_{\mathbb{M},\eta}$. Reformulating, if t has type τ, the function:

$$\left\{\begin{array}{ccc}\mathbb{T}_{\mathbb{M},\eta} & \to & [\![\tau]\!]^{\eta}_{\mathbb{M}} \\ \rho & \mapsto & [\![t]\!]^{\eta,\rho}_{\mathbb{M}}\end{array}\right.$$

defines the semantics of t as a η-indexed family of random variables using randomness source $\mathbb{T}_{\mathbb{M},\eta}$.

Question 5. *Prove that if, for every η, there exists a bijection $\beta : \mathbb{T}_{\mathbb{M},\eta} \to \mathbb{T}_{\mathbb{M},\eta}$ such that for all $\rho \in \mathbb{T}_{\mathbb{M},\eta}$:*

$$[\![u]\!]_{\mathbb{M}}^{\eta,\rho} = [\![v]\!]_{\mathbb{M}}^{\eta,\beta(\rho)} \qquad \text{and} \qquad \rho_{\mathsf{a}} = \rho'_{\mathsf{a}} \text{ where } \rho = (\rho_{\mathsf{a}}, \rho_{\mathsf{h}}) \text{ and } \beta(\rho) = (\rho'_{\mathsf{a}}, \rho'_{\mathsf{h}})$$

then $\mathbb{M} \models u \sim v$.

Question 6. *Deduce that:*

$$u, \mathsf{n} \sim u, \mathsf{n}'$$

is a valid formula whenever u is a ground term and $\mathsf{n}, \mathsf{n}' \notin \mathsf{st}(u)$.

Probabilistic Couplings Let $\mathbb{S}_1, \mathbb{S}_2$ be two sets implicitly equipped with probability distributions, resp., μ_1 and μ_2. For the sack of simplicity, we assume that $\mathbb{S}_1$ and $\mathbb{S}_2$ are both finite (though this is unnecessary). In that case, we can define the distributions μ_1 and μ_2 as follows:

$$\mu_i : \mathbb{S}_i \to [0;1] \quad \text{ such that } \sum_{x \in \mathbb{S}_i} \mu_i(x) = 1 \qquad \text{(for any } i \in \{1;2\})$$

A distribution c over $\mathbb{S}_1 \times \mathbb{S}_2$ is a *probabilistic coupling* of $(\mathbb{S}_1, \mu_1)$ and $(\mathbb{S}_2, \mu_2)$, which we write $c : (\mathbb{S}_1, \mu_1) \bowtie (\mathbb{S}_2, \mu_2)$[1], if c is such that its left and right marginals $\pi_1(c)$ and $\pi_2(c)$ follow, respectively, the distribution μ_1 and μ_2, where:

$$\forall e_1 \in \mathbb{S}_1,\ \pi_1(c)(e_1) = \sum_{e_2 \in \mathbb{S}_2} c(e_1, e_2)$$
$$\forall e_2 \in \mathbb{S}_2,\ \pi_1(c)(e_2) = \sum_{e_1 \in \mathbb{S}_1} c(e_1, e_2)$$

Question 7 (Independent and equality couplings)**.**

- *Show that $c(x,y) = \mu_1(x) \times \mu_2(y)$ is a probabilistic couplings $c : (\mathbb{S}_1, \mu_1) \bowtie (\mathbb{S}_2, \mu_2)$.*
- *Build a coupling $c : (\mathbb{S}, \mu) \bowtie (\mathbb{S}, \mu)$ such that $c(x,y) = 0$ whenever $x \neq y$.*

We let $\mathsf{supp}(\mu) = \{x \mid \mu(x) > 0\}$ be the support of a discrete distribution μ.

Question 8. *Let $\mathbb{M}$ be a model. Show that if, for every η, there exists $c : \mathbb{T}_{\mathbb{M},\eta} \times \mathbb{T}_{\mathbb{M},\eta}$ such that $\forall (\rho^1, \rho^2) \in \mathsf{supp}(c)$,*

$$\rho^1_{\mathsf{a}} = \rho^2_{\mathsf{a}} \qquad [\![u]\!]_{\mathbb{M}}^{\eta,\rho^1} = [\![v]\!]_{\mathbb{M}}^{\eta,\rho^2} \qquad \text{(where } \rho^i = (\rho^i_{\mathsf{a}}, \rho^i_{\mathsf{h}}) \text{ for } i \in \{1;2\})$$

then $\mathbb{M} \models u \sim v$.

Question 9. *Re-prove the rule of question 6 using a probabilistic couplings.*

Question 10. *Prove using a probabilistic coupling that*

$$\mathsf{if}\ \phi\ \mathsf{then}\ t_1[\mathsf{n}]\ \mathsf{else}\ t_2 \sim \mathsf{if}\ \phi\ \mathsf{then}\ t_1[\mathsf{n}']\ \mathsf{else}\ t_2$$

is a valid formula whenever ϕ, t_1 are ground terms and $\mathsf{n}, \mathsf{n}' \notin \mathsf{st}(\phi, t_1)$.
(Note that t_2 is unconstrained.)

Question 11. *Prove that without the condition that $\mathsf{n}, \mathsf{n}' \notin \mathsf{st}(\phi, t_1)$, the rule above is unsound.*

4 Relations Among Hash Cryptographic Assumptions

4.1 Hash Functions

Let $\Sigma = \{0,1\}$. A cryptographic hash function $\mathsf{H} : \Sigma^* \mapsto \Sigma^L$ allows to compute, for every message m, a *digest* $\mathsf{H}(m)$ – often called the *hash* – of fixed length L.[2] Examples of such functions are SHA-2, or the more recent SHA-3.

[1] We often let the distributions μ_1 and μ_2 be implicit.
[2] L is more or less the security parameter.

There are many security properties that we may want from a cryptographic hash function. A common property is to require that the hash function has no **collision**, where a collision is a pair of distinct messages m_0, m_1 such that $\mathsf{H}(m_0) = \mathsf{H}(m_1)$. Of course, for cardinality reasons, this cannot be achieved.

Therefore, we are going to slightly change the setting. A *keyed* cryptographic hash function $\mathsf{H} : \Sigma^* \times \Sigma^K \mapsto \Sigma^L$ takes as input a message m of any length and a key k of length K, and compute the hash of m under k. A keyed hash function could be implemented, for example, by taking $\mathsf{H}(m,k) \stackrel{\text{def}}{=} \mathsf{SHA\text{-}3}(k||m)$. To simplify things, we assume $K = L = \eta$ from now on.

4.2 Hardness Hypotheses on Hash Functions

We now present three different security notions for keyed hash functions.

Collision-Resistance A keyed cryptographic hash $\mathsf{H}(\_, \_)$ is computationally collision resistant if no PPTM adversary can built collisions, even when it has access to a hashing oracle.

Formally, a hash is *collision resistant under hidden key attacks* (CR-HK) iff. for every PPTM $\mathcal{A}$:

$$\mathsf{Pr}_{\mathsf{k}}\left(\mathcal{A}^{\mathcal{O}_{\mathsf{H}(\cdot,\mathsf{k})}}(1^\eta) = \langle m_1\,,\, m_2\rangle, m_1 \neq m_2 \text{ and } \mathsf{H}(m_1,\mathsf{k}) = \mathsf{H}(m_2,\mathsf{k})\right)$$

is negligible, where k is drawn uniformly in $\{0,1\}^\eta$.

Unforgeability A keyed hash function is computationally unforgeable when no adversary can forge new hashes, even when the adversary has access to a hashing oracle.

Formally, a hash is *unforgeable against chosen-message attacks* (EUF-CMA) iff. for every PPTM $\mathcal{A}$:

$$\mathsf{Pr}_{\mathsf{k}}\left(\mathcal{A}^{\mathcal{O}_{\mathsf{H}(\cdot,\mathsf{k})}}(1^\eta) = \langle m\,,\, \sigma\rangle,\ m \text{ not queried to } \mathcal{O}_{\mathsf{H}(\cdot,\mathsf{k})} \text{ and } \sigma = \mathsf{H}(m,\mathsf{k})\right)$$

is negligible, where k is drawn uniformly in $\{0,1\}^\eta$.

Pseudo-Random Function A keyed hash function $\mathsf{H}(\cdot,\mathsf{k})$ is a PRF if its outputs are computationally indistinguishable from the outputs of a random function.

Formally, a hash function is a *Pseudo Random Function* iff. for any PPTM $\mathcal{A}$:

$$\left|\mathsf{Pr}_{\mathsf{k}}(\mathcal{A}^{\mathcal{O}_{\mathsf{H}(\cdot,\mathsf{k})}}(1^\eta) = 1) - \mathsf{Pr}_g(\mathcal{A}^{\mathcal{O}_{g(\cdot)}}(1^\eta) = 1)\right|$$

is negligible, where:

- k is drawn uniformly in $\{0,1\}^\eta$.
- g is a random function from $\{0,1\}^*$ to $\{0,1\}^\eta$.

4.3 Relations Among Security Notions and Rule Schemata

Show that we have the following relations among keyed hash function security notions.

Question 12 ($\star$). *Show that PRF $\Rightarrow$ EUF-CMA $\Rightarrow$ CR-HK.*

We now consider the problem of designing sound rules of the indistinguishability logic capturing these different keyed hash function security notions.

Question 13. *Design and prove sound a rule schemata for CR-HK.*

Question 14. *Design and prove sound a rule schemata for PRF. In a first time, assume that there are at most two calls to the hash oracle. Then, generalize to any number of calls.*

4.4 EUF Rule and Variation

If H is an $\mathsf{EUF\text{-}CMA}$ keyed hash function, then the *ground* rule:

$$\frac{}{\left(s = \mathsf{H}(m,\mathsf{k}) \rightarrow \bigvee_{u \in \mathcal{S}} m = u\right) \sim \mathsf{true}}\ \mathsf{EUF}$$

is sound, when:

- $\mathcal{S} = \{u \mid \mathsf{H}(u,\mathsf{k}) \in \mathsf{st}(s,m)\}$;
- k appears only in H key positions in s, m, i.e. $\mathsf{k} \sqsubseteq_{\mathsf{H}(\_,\cdot)} s, m$.

We assume that the EUF rule given above is sound. We are now going to prove an improved, more precise, version of the rule.

Ignoring Hashes in Conditions We show that we can ignore some hashes appearing in conditions in s or m. To simplify matter, we only do it for a single condition.

Question 15. *Assume that H is $\mathsf{EUF\text{-}CMA}$. Show that the following rule is sound:*

$$\frac{}{(\mathsf{if}\ b\ \mathsf{then}\ s_0\ \mathsf{else}\ s_1) = \mathsf{H}(m,\mathsf{k}) \rightarrow \bigvee_{u \in \mathcal{S}_1 \cup \mathcal{S}_2} m = u \sim \mathsf{true}}\ \mathsf{EUF_{nc}}$$

when b, s_0, s_1, m are ground *terms, and:*

- $\mathcal{S}_i = \{u \mid \mathsf{H}(u,\mathsf{k}) \in \mathsf{st}(s_i,m)\}$, *for $i \in \{0,1\}$;*
- k *appears only in H key positions in s_0, s_1, m.*

Remark that we do not make *any* assumption on b, except that it is ground. E.g., we can have $b \equiv (\mathsf{att}(\mathsf{k}) = \mathsf{H}(0,\mathsf{k}))$.

Question 16 ($\star$)**.** *What is the relation between the advantage against $\mathsf{EUF_{nc}}$ and the advantage against the $\mathsf{EUF\text{-}CMA}$ security assumption? How would this advantage evolve if we generalized the $\mathsf{EUF_{nc}}$ rule to N conditions $b_1, \ldots, b_n$?*