CryptoVerif

Bruno Blanchet

INRIA Paris-Rocquencourt
bruno.blanchet@inria.fr

November 2014

Input and output of the tool

1. Prepare the input file containing
   - the specification of the protocol to study (initial game),
   - the security assumptions on the cryptographic primitives,
   - the security properties to prove.
2. Run CryptoVerif
3. CryptoVerif outputs
   - the sequence of games that leads to the proof,
   - a succinct explanation of the transformations performed between games,
   - an upper bound of the probability of success of an attack.

Process calculus for games: terms

Terms represent computations on messages (bitstrings).

$M ::=$	terms
$x, y, z, x[M_1, \dots, M_n]$	variable
$f(M_1, \dots, M_n)$	function application

Function symbols f correspond to functions computable by deterministic Turing machines that always terminate.

Process calculus for games: processes

$Q ::=$	input process
0	end
$Q \mid Q'$	parallel composition
$!^{i \leq N} Q$	replication N times
newChannel $c; Q$	restriction for channels
$c(x_1 : T_1, \dots, x_m : T_m); P$	input
$P ::=$	output process
yield	end
$\overline{c}\langle M_1, \dots, M_m \rangle; Q$	output
event $e(M_1, \dots, M_m); P$	event
new $x : T; P$	random number generation (uniform)
let $x : T = M$ **in** P	assignment
if M **then** P **else** P'	conditional
find $j \leq N$ **suchthat defined**$(x[j], \dots) \wedge M$ **then** P **else** P'	array lookup

Example: 1. symmetric encryption

We consider a probabilistic, length-revealing encryption scheme.

Definition (Symmetric encryption scheme SE)

- (Randomized) key generation function *kgen*.
- (Randomized) encryption function $enc(m, k, r')$ takes as input a message m, a key k, and random coins r'.
- Decryption function $dec(c, k)$ such that

$$dec(enc(m, kgen(r), r'), kgen(r)) = i_\perp(m)$$

The decryption returns a bitstring or $\perp$:

- $\perp$ when decryption fails,
- the cleartext when decryption succeeds.

The injection $i_\perp$ maps a bitstring to the same bitstring in bitstring $\cup \{\perp\}$.

Example: 2. MAC

Definition (Message Authentication Code scheme MAC)

- (Randomized) key generation function $mkgen$.
- MAC function $mac(m, k)$ takes as input a message m and a key k.
- Verification function $verify(m, k, t)$ such that

$$verify(m, k, mac(m, k)) = true.$$

A MAC is essentially a keyed hash function.

A MAC guarantees the integrity and authenticity of the message because only someone who knows the secret key can build the MAC.

Example: 3. encrypt-then-MAC

We define an authenticated encryption scheme by the encrypt-then-MAC construction:

$$enc'(m,(k,mk),r'') = e, mac(e,mk) \text{ where } e = enc(m,k,r'').$$

A basic example of protocol using encrypt-then-MAC:

- A and B initially share an encryption key k and a MAC key mk.
- A sends to B a fresh key k' encrypted under authenticated encryption, implemented as encrypt-then-MAC.

$$A \rightarrow B : e = enc(k',k,r''), mac(e,mk) \qquad k' \text{ fresh}$$

k' should remain secret.

Example: initialization

$$A \to B : e = enc(k', k, r''), mac(e, mk) \qquad k' \text{ fresh}$$

$$\begin{aligned} Q_0 = {} & start(); \textbf{new } r : keyseed; \textbf{let } k : key = kgen(r) \textbf{ in} \\ & \textbf{new } r' : mkeyseed; \textbf{let } mk : mkey = mkgen(r') \textbf{ in } \overline{c}\langle\rangle; (Q_A \mid Q_B) \end{aligned}$$

Initialization of keys:

1. The process Q_0 waits for a message on channel *start* to start running.
   The adversary triggers this process.
2. Q_0 generates encryption and MAC keys, k and mk respectively, using the key generation algorithms *kgen* and *mkgen*.
3. Q_0 returns control to the adversary by the output $\overline{c}\langle\rangle$.
   Q_A and Q_B represent the actions of A and B (see next slides).

Example: role of A

$$A \to B : e = enc(k', k, r''), mac(e, mk) \qquad k' \text{ fresh}$$

$$\begin{aligned} Q_A = {!}^{i \leq n} c_A(); &\textbf{new } k' : key; \textbf{new } r'' : coins; \\ &\textbf{let } e : bitstring = enc(k2b(k'), k, r'') \textbf{ in} \\ &\overline{c_A}\langle e, mac(e, mk)\rangle \end{aligned}$$

Role of A:

1. ${!}^{i \leq n}$ represents n copies, indexed by $i \in [1, n]$
   The protocol can be run n times (polynomial in the security parameter).
2. The process is triggered when a message is sent on c_A by the adversary.
3. The process chooses a fresh key k' and sends the message on channel c_A.

Example: role of B

$$A \rightarrow B : e = enc(k', k, r''), mac(e, mk) \qquad k' \text{ fresh}$$

$$\begin{aligned} Q_B = {} & !^{i' \leq n} c_B(e' : bitstring, ma : macstring); \\ & \textbf{if } verify(e', mk, ma) \textbf{ then} \\ & \textbf{let } i_\perp(k2b(k'')) = dec(e', k) \textbf{ in } \overline{c_B}\langle\rangle \end{aligned}$$

Role of B:

1. n copies, as for Q_A.
2. The process Q_B waits for the message on channel c_B.
3. It verifies the MAC, decrypts, and stores the key in k''.

Example: summary of the initial game

$$A \to B : e = enc(k', k, r''), mac(e, mk) \qquad k' \text{ fresh}$$

$$\begin{aligned} Q_0 = {} & start(); \textbf{new } r : keyseed; \textbf{let } k : key = kgen(r) \textbf{ in} \\ & \textbf{new } r' : mkeyseed; \textbf{let } mk : mkey = mkgen(r') \textbf{ in } \overline{c}\langle\rangle; (Q_A \mid Q_B) \end{aligned}$$

$$\begin{aligned} Q_A = {} & !^{i \leq n} c_A(); \textbf{new } k' : key; \textbf{new } r'' : coins; \\ & \textbf{let } e : bitstring = enc(k2b(k'), k, r'') \textbf{ in} \\ & \overline{c_A}\langle e, mac(e, mk)\rangle \end{aligned}$$

$$\begin{aligned} Q_B = {} & !^{i' \leq n} c_B(e' : bitstring, ma : macstring); \\ & \textbf{if } verify(e', mk, ma) \textbf{ then} \\ & \textbf{let } i_\perp(k2b(k'')) = dec(e', k) \textbf{ in } \overline{c_B}\langle\rangle \end{aligned}$$

Security assumptions on primitives

The most frequent cryptographic primitives are already specified in a library. The user can use them without redefining them.

In the example:

- The MAC is UF-CMA (unforgeable under chosen message attacks). An adversary that has access to the MAC and verification oracles has a negligible probability of forging a MAC (for a message on which the MAC oracle has not been called).

Security assumptions on primitives

The most frequent cryptographic primitives are already specified in a library. The user can use them without redefining them.

In the example:

- The MAC is UF-CMA (unforgeable under chosen message attacks).
  An adversary that has access to the MAC and verification oracles has a negligible probability of forging a MAC (for a message on which the MAC oracle has not been called).
- The encryption is IND-CPA (indistinguishable under chosen plaintext attacks).
  An adversary has a negligible probability of distinguishing the encryption of two messages of the same length.

Security assumptions on primitives

The most frequent cryptographic primitives are already specified in a library. The user can use them without redefining them.

In the example:

- The MAC is UF-CMA (unforgeable under chosen message attacks). An adversary that has access to the MAC and verification oracles has a negligible probability of forging a MAC (for a message on which the MAC oracle has not been called).
- The encryption is IND-CPA (indistinguishable under chosen plaintext attacks). An adversary has a negligible probability of distinguishing the encryption of two messages of the same length.
- All keys have the same length: **forall** $y : key; Z(k2b(y)) = Z_k$.

Security properties to prove

In the example:

- One-session secrecy of k'': each k'' is indistinguishable from a random number.
- Secrecy of k'': the k'' are indistinguishable from independent random numbers.

Demo

- CryptoVerif input file: enc-then-MAC.cv
- library of primitives
- run CryptoVerif
- output

Arrays

A variable defined under a replication is implicitly an array:

$$\begin{aligned} Q_A = {!}^{i \leq n} c_A(); \textbf{new } k'[i] : key; \textbf{new } r''[i] : coins; \\ \textbf{let } e[i] : bitstring = enc(k2b(k'[i]), k, r''[i]) \textbf{ in} \\ \overline{c_A}\langle e[i], mac(e[i], mk)\rangle \end{aligned}$$

Requirements:

- Only variables with the current indices can be assigned.
- Variables may be defined at several places, but only one definition can be executed for the same indices. (**if** . . . **then let** $x = M$ **in** P **else let** $x = M'$ **in** P' is ok)

So each array cell can be assigned at most once.

Arrays allow one to remember the values of all variables during the whole execution

Arrays (continued)

find performs an array lookup:

$$!^{i \leq N} \ldots \textbf{let } x = M \textbf{ in } P$$
$$\mid\, !^{i' \leq N'} c(y : T)\textbf{find } j \leq N \textbf{ suchthat defined}(x[j]) \wedge y = x[j] \textbf{ then } \ldots$$

Note that **find** is here used outside the scope of x.

This is the only way of getting access to values of variables in other sessions.

When several array elements satisfy the condition of the **find**, the returned index is chosen randomly, with uniform probability.

Arrays versus lists

Arrays replace lists often used in cryptographic proofs.

$$
\begin{aligned}
&!^{i \leq N} \ldots \textbf{let } x = M \textbf{ in let } y = M' \textbf{ in } P \\
&\mid\, !^{i' \leq N'} c(x' : T)\textbf{find } j \leq N \textbf{ suchthat defined}(x[j]) \wedge x' = x[j] \textbf{ then} \\
&\qquad\qquad P'(y[j])
\end{aligned}
$$

written by cryptographers

$$
\begin{aligned}
&!^{i \leq N} \ldots \textbf{let } x = M \textbf{ in let } y = M' \textbf{ in insert } (x, y) \textbf{ in } L; P \\
&\mid\, !^{i' \leq N'} c(x' : T)\textbf{get } (x, y) \textbf{ in } L \textbf{ suchthat } x' = x; P'(y)
\end{aligned}
$$

Arrays avoid the need for explicit list insertion instructions, which would be hard to guess for an automatic tool.

Indistinguishability as observational equivalence

Two processes (games) Q_1, Q_2 are observationally equivalent when the adversary has a negligible probability of distinguishing them: $Q_1 \approx Q_2$.

- The adversary is represented by an acceptable evaluation context $C ::= [\,] \quad C \mid Q \quad Q \mid C \quad \textbf{newChannel}\ c; C$.
- $C[Q]$ may execute events, collected in a sequence $\mathcal{E}$.
- A distinguisher D takes as input $\mathcal{E}$ and returns **true** or **false**.
  - Example: $D(\mathcal{E}) = \textbf{true}$ if and only if $e \in \mathcal{E}$.
- $\Pr[C[Q] \rightsquigarrow D]$ is the probability that $C[Q]$ executes $\mathcal{E}$ such that $D(\mathcal{E}) = \textbf{true}$.

Definition (Indistinguishability)

We write $Q \approx_p^V Q'$ when, for all evaluation contexts C acceptable for Q and Q' with public variables V and all distinguishers D,

$$|\Pr[C[Q] \rightsquigarrow D] - \Pr[C[Q'] \rightsquigarrow D]| \leq p(C, D).$$

Indistinguishability as observational equivalence

Lemma

1. *Reflexivity:* $Q \approx_0^V Q$.
2. *Symmetry:* $\approx_p^V$ *is symmetric.*
3. *Transitivity: if* $Q \approx_p^V Q'$ *and* $Q' \approx_{p'}^V Q''$, *then* $Q \approx_{p+p'}^V Q''$.
4. *Application of context: if* $Q \approx_p^V Q'$ *and* C *is an evaluation context acceptable for* Q *and* Q' *with public variables* V, *then* $C[Q] \approx_{p'}^{V'} C[Q']$, *where* $p'(C', D) = p(C'[C[\,]], D)$ *and* $V' \subseteq V \cup \mathrm{var}(C)$.

Proof technique

We transform a game G_0 into an observationally equivalent one using:

- observational equivalences $L \approx_p R$ given as axioms and that come from security assumptions on primitives. These equivalences are used inside a context:
$$G_1 \approx_0 C[L] \approx_{p'} C[R] \approx_0 G_2$$
- syntactic transformations: simplification, expansion of assignments, ...

We obtain a sequence of games $G_0 \approx_{p_1} G_1 \approx \ldots \approx_{p_m} G_m$, which implies $G_0 \approx_{p_1+\cdots+p_m} G_m$.

If some trace property holds up to probability p in G_m, then it holds up to probability $p + p_1 + \cdots + p_m$ in G_0.

MAC: definition of security (UF-CMA)

A MAC guarantees the integrity and authenticity of the message because only someone who knows the secret key can build the MAC.
More formally, $\mathrm{Succ}^{\mathrm{uf-cma}}_{\mathrm{MAC}}(t, q_m, q_v, l)$ is negligible if t is polynomial in the security parameter:

Definition (UnForgeability under Chosen Message Attacks, UF-CMA)

$$\mathrm{Succ}^{\mathrm{uf-cma}}_{\mathrm{MAC}}(t, q_m, q_v, l) =$$

$$\max_{\mathcal{A}} \Pr\left[\begin{array}{l} k \overset{R}{\leftarrow} \mathit{mkgen}; (m, s) \leftarrow \mathcal{A}^{\mathit{mac}(.,k),\mathit{verify}(.,k,.)} : \mathit{verify}(m, k, s) \wedge \\ m \text{ was never queried to the oracle } \mathit{mac}(., k) \end{array}\right]$$

where $\mathcal{A}$ runs in time at most t,
calls $\mathit{mac}(., k)$ at most q_m times with messages of length at most l,
calls $\mathit{verify}(., k, .)$ at most q_v times with messages of length at most l.

MAC: intuition behind the CryptoVerif definition

By the previous definition, up to negligible probability,

- the adversary cannot forge a correct MAC
- so, assuming $k \overset{R}{\leftarrow} \mathit{mkgen}$ is used only for generating and verifying MACs, the verification of a MAC with $\mathit{verify}(m, k, t)$ can succeed only if m is in the list (array) of messages whose $\mathit{mac}(\cdot, k)$ has been computed by the protocol
- so we can replace a call to *verify* with an array lookup:
  if the call to *mac* is $\mathit{mac}(x, k)$, we replace $\mathit{verify}(m, k, t)$ with

$$\textbf{find } j \leq N \textbf{ suchthat defined}(x[j]) \wedge \\ (m = x[j]) \wedge \mathit{verify}(m, k, t) \textbf{ then true else false}$$

MAC: CryptoVerif definition

$$\mathit{verify}(m, \mathit{mkgen}(r), \mathit{mac}(m, \mathit{mkgen}(r))) = \mathbf{true}$$

$$
\begin{array}{l}
!^{N''}\mathbf{new}\ r : \mathit{mkeyseed}; (\\
\quad !^{N}\mathit{Omac}(x : \mathit{bitstring}) := \mathit{mac}(x, \mathit{mkgen}(r)),\\
\quad !^{N'}\mathit{Overify}(m : \mathit{bitstring}, t : \mathit{macstring}) := \mathit{verify}(m, \mathit{mkgen}(r), t))\\
\approx\\
!^{N''}\mathbf{new}\ r : \mathit{mkeyseed}; (\\
\quad !^{N}\mathit{Omac}(x : \mathit{bitstring}) := \mathit{mac}(x, \mathit{mkgen}(r)),\\
\quad !^{N'}\mathit{Overify}(m : \mathit{bitstring}, t : \mathit{macstring}) :=\\
\qquad \mathbf{find}\ j \leq N\ \mathbf{suchthat}\ \mathbf{defined}(x[j]) \wedge (m = x[j]) \wedge\\
\qquad\qquad \mathit{verify}(m, \mathit{mkgen}(r), t)\ \mathbf{then\ true\ else\ false})
\end{array}
$$

MAC: CryptoVerif definition

$$\mathit{verify}(m, \mathit{mkgen}(r), \mathit{mac}(m, \mathit{mkgen}(r))) = \textbf{true}$$

$$
\begin{array}{l}
!^{N''}\textbf{new}\ r : \mathit{mkeyseed}; (\\
\quad !^{N}\mathit{Omac}(x : \mathit{bitstring}) := \mathit{mac}(x, \mathit{mkgen}(r)),\\
\quad !^{N'}\mathit{Overify}(m : \mathit{bitstring}, t : \mathit{macstring}) := \mathit{verify}(m, \mathit{mkgen}(r), t))\\
\approx_{N''\times\mathrm{Succ}^{\mathrm{uf-cma}}_{\mathrm{MAC}}(\textbf{time}+(N''-1)(\textbf{time}(\mathit{mkgen})+N\,\textbf{time}(\mathit{mac},\mathrm{maxl}(x))+N'\,\textbf{time}(\mathit{verify},\mathrm{maxl}(m))),N,N',\max(\mathrm{maxl}(x),\mathrm{maxl}(m)))}\\
!^{N''}\textbf{new}\ r : \mathit{mkeyseed}; (\\
\quad !^{N}\mathit{Omac}(x : \mathit{bitstring}) := \mathit{mac}'(x, \mathit{mkgen}'(r)),\\
\quad !^{N'}\mathit{Overify}(m : \mathit{bitstring}, t : \mathit{macstring}) :=\\
\qquad \textbf{find}\ j \leq N\ \textbf{suchthat}\ \textbf{defined}(x[j]) \wedge (m = x[j]) \wedge\\
\qquad\qquad \mathit{verify}'(m, \mathit{mkgen}'(r), t)\ \textbf{then true else false})
\end{array}
$$

CryptoVerif understands such specifications of primitives.

MAC: using the CryptoVerif definition

CryptoVerif applies the previous rule automatically in any context, perhaps containing several occurrences of *mac* and of *verify*:

- Each occurrence of *mac* is replaced with mac'.
- Each occurrence of *verify* is replaced with a **find** that looks in all arrays of computed MACs (one array for each occurrence of function *mac*).

Symmetric encryption: definition of security (IND-CPA)

An adversary has a negligible probability of distinguishing the encryption of two messages of the same length.

Definition (INDistinguishability under Chosen Plaintext Attacks, IND-CPA)

$$\mathrm{Succ}_{\mathrm{SE}}^{\mathrm{ind-cpa}}(t, q_e, l) =$$
$$\max_{\mathcal{A}} 2\,\mathrm{Pr}\left[b \overset{R}{\leftarrow}\{0,1\}; k \overset{R}{\leftarrow} kgen; b' \leftarrow \mathcal{A}^{enc(LR(.,.,b),k)} : b' = b\right] - 1$$

where $\mathcal{A}$ runs in time at most t,
calls $enc(LR(.,.,b),k)$ at most q_e times on messages of length at most l,
$LR(x,y,0) = x$, $LR(x,y,1) = y$, and $LR(x,y,b)$ is defined only when x and y have the same length.

Symmetric encryption: CryptoVerif definition

$$dec(enc(m, kgen(r), r'), kgen(r)) = i_\perp(m)$$

$$\begin{aligned}
&!^{N'}\textbf{new } r : keyseed; !^{N} Oenc(x : bitstring) := \\
&\qquad\qquad \textbf{new } r' : coins; enc(x, kgen(r), r') \\
&\approx \\
&!^{N'}\textbf{new } r : keyseed; !^{N} Oenc(x : bitstring) := \\
&\qquad\qquad \textbf{new } r' : coins; enc(Z(x), kgen(r), r')
\end{aligned}$$

$Z(x)$ is the bitstring of the same length as x containing only zeroes (for all $x : nonce$, $Z(x) = Znonce, \ldots$).

Symmetric encryption: CryptoVerif definition

$$dec(enc(m, kgen(r), r'), kgen(r)) = i_\perp(m)$$

$$!^{N'} \textbf{new } r : keyseed; !^{N} Oenc(x : bitstring) :=$$
$$\qquad \textbf{new } r' : coins; enc(x, kgen(r), r')$$
$$\approx_{N' \times \mathrm{Succ}_{\mathrm{SE}}^{\mathrm{ind-cpa}}(\textbf{time} + (N'-1)(\textbf{time}(kgen) + N\,\textbf{time}(enc, \mathrm{maxl}(x)) + N\,\textbf{time}(Z, \mathrm{maxl}(x))), N, \mathrm{maxl}(x))}$$
$$!^{N'} \textbf{new } r : keyseed; !^{N} Oenc(x : bitstring) :=$$
$$\qquad \textbf{new } r' : coins; enc'(Z(x), kgen'(r), r')$$

$Z(x)$ is the bitstring of the same length as x containing only zeroes (for all $x : nonce$, $Z(x) = Znonce$, $\ldots$).

Syntactic transformations (1)

Expansion of assignments: replacing a variable with its value. (Not completely trivial because of array references.)

Example

If mk is defined by

$$\textbf{let } mk = mkgen(r')$$

and there are no array references to mk, then mk is replaced with $mkgen(r')$ in the game and the definition of mk is removed.

Syntactic transformations (2)

Single assignment renaming: when a variable is assigned at several places, rename it with a distinct name for each assignment.
(Not completely trivial because of array references.)

Example

$$\begin{aligned}
&\mathit{start}(); \textbf{new } r_A : T_r; \textbf{let } k_A = \mathit{kgen}(r_A) \textbf{ in}\\
&\qquad\quad \textbf{new } r_B : T_r; \textbf{let } k_B = \mathit{kgen}(r_B) \textbf{ in } \overline{c}\langle\rangle; (Q_K \mid Q_S)\\
&Q_K = !^{i \leq n} c(h : T_h, k : T_k)\\
&\qquad \textbf{if } h = A \textbf{ then let } k' = k_A \textbf{ else}\\
&\qquad \textbf{if } h = B \textbf{ then let } k' = k_B \textbf{ else let } k' = k\\
&Q_S = !^{i' \leq n'} c'(h' : T_h);\\
&\qquad \textbf{find } j \leq n \textbf{ suchthat defined}(h[j], k'[j]) \wedge h' = h[j] \textbf{ then } P_1(k'[j])\\
&\qquad \textbf{else } P_2
\end{aligned}$$

Syntactic transformations (2)

Single assignment renaming: when a variable is assigned at several places, rename it with a distinct name for each assignment.
(Not completely trivial because of array references.)

Example

$$
\begin{aligned}
&\mathit{start}();\, \textbf{new } r_A : T_r; \textbf{let } k_A = \mathit{kgen}(r_A) \textbf{ in}\\
&\qquad \textbf{new } r_B : T_r; \textbf{let } k_B = \mathit{kgen}(r_B) \textbf{ in } \overline{c}\langle\rangle; (Q_K \mid Q_S)\\
&Q_K = !^{i \leq n} c(h : T_h, k : T_k)\\
&\qquad \textbf{if } h = A \textbf{ then let } k'_1 = k_A \textbf{ else}\\
&\qquad \textbf{if } h = B \textbf{ then let } k'_2 = k_B \textbf{ else let } k'_3 = k\\
&Q_S = !^{i' \leq n'} c'(h' : T_h);\\
&\qquad \textbf{find } j \leq n \textbf{ suchthat defined}(h[j], k'_1[j]) \wedge h' = h[j] \textbf{ then } P_1(k'_1[j])\\
&\qquad \textbf{orfind } j \leq n \textbf{ suchthat defined}(h[j], k'_2[j]) \wedge h' = h[j] \textbf{ then } P_1(k'_2[j])\\
&\qquad \textbf{orfind } j \leq n \textbf{ suchthat defined}(h[j], k'_3[j]) \wedge h' = h[j] \textbf{ then } P_1(k'_3[j])\\
&\qquad \textbf{else } P_2
\end{aligned}
$$

Syntactic transformations (3)

Move new: move restrictions downwards in the game as much as possible, when there is no array reference to them.
(Moving **new** $x : T$ under a **if** or a **find** duplicates it.
A subsequent single assignment renaming will distinguish cases.)

Example

$$\textbf{new } x : \mathit{nonce}; \textbf{if } c \textbf{ then } P_1 \textbf{ else } P_2$$

becomes

$$\textbf{if } c \textbf{ then new } x : \mathit{nonce}; P_1 \textbf{ else new } x : \mathit{nonce}; P_2$$

Syntactic transformations (4)

- Merge arrays: merge several variables $x_1, \dots, x_n$ into a single variable x_1 when they are used for different indices (defined in different branches of a test **if** or **find**).
- Merge branches of **if** or **find** when they execute the same code, up to renaming of variables with array accesses.

Syntactic transformations (5): manual transformations

Insert an instruction: insert a test to distinguish cases; insert a variable definition; ...

Preserves the semantics of the game (e.g., the rest of the code is copied in both branches of the inserted test).

Example

P becomes

$$\textbf{if } \mathit{cond} \textbf{ then } P \textbf{ else } P$$

Subsequent transformations can transform *P* differently, depending on whether *cond* holds.

Syntactic transformations (6): manual transformations

- Insert an event: to apply Shoup's lemma.
    - A subprocess P becomes **event** e.
    - The probability of distinguishing the two games is the probability of executing event e. It will be bound by a proof by sequences of games.
- Replace a term with an equal term. CryptoVerif verifies that the terms are really equal.

Simplification and elimination of collisions

- CryptoVerif collects equalities that come from:
  - Assignments: **let** $x = M$ **in** P implies that $x = M$ in P
  - Tests: **if** $M = N$ **then** P implies that $M = N$ in P
  - Definitions of cryptographic primitives
  - When a **find** guarantees that $x[j]$ is defined, equalities that hold at definition of x also hold under the find (after substituting j for the array indices at the definition of x)
  - Elimination of collisions: if x is created by **new** $x : T$, $x[i] = x[j]$ implies $i = j$, up to negligible probability (when T is large)
- These equalities are combined to simplify terms.
- When terms can be simplified, processes are simplified accordingly. For instance:
  - If M simplifies to **true**, then **if** M **then** P_1 **else** P_2 simplifies P_1.
  - If a condition of **find** simplifies to **false**, then the corresponding branch is removed.

Proof of security properties: one-session secrecy

One-session secrecy: the adversary cannot distinguish any of the secrets from a random number with one test query.

Definition (One-session secrecy)

Assume that the variable x of type T is defined in G under a single $!^{i \leq n}$.

G preserves the one-session secrecy of x up to probability p when, for all evaluation contexts C acceptable for $G \mid Q_x$ with no public variables that do not contain S, $2\Pr[C[G \mid Q_x] \rightsquigarrow D_{\mathsf{S}}] - 1 \leq p(C)$ where

$$\begin{aligned} Q_x = {} & c_0(); \mathbf{new}\ b : bool; \overline{c_0}\langle\rangle; \\ & (c(j : [1, n]); \mathbf{if\ defined}(x[j])\ \mathbf{then} \\ & \quad \mathbf{if}\ b\ \mathbf{then}\ \overline{c}\langle x[j]\rangle\ \mathbf{else\ new}\ y : T; \overline{c}\langle y\rangle \\ & \mid c'(b' : bool); \mathbf{if}\ b = b'\ \mathbf{then\ event}\ \mathsf{S}) \end{aligned}$$

$D_{\mathsf{S}}(\mathcal{E}) = (\mathsf{S} \in \mathcal{E})$, c_0, c, c', b, b', j, y, and S do not occur in G.

Proof of security properties: one-session secrecy

One-session secrecy: the adversary cannot distinguish any of the secrets from a random number with one test query.

Criterion for proving one-session secrecy of x:
x is defined by **new** $x[i] : T$ and there is a set of variables S such that only variables in S depend on x.
The output messages and the control-flow do not depend on x.

Proof of security properties: secrecy

Secrecy: the adversary cannot distinguish the secrets from independent random numbers with several test queries.

Criterion for proving secrecy of x: same as one-session secrecy, plus $x[i]$ and $x[i']$ do not come from the same copy of the same restriction when $i \neq i'$.

Proof strategy: advice

- One tries to execute each transformation given by the definition of a cryptographic primitive.
- When it fails, it tries to analyze why the transformation failed, and suggests syntactic transformations that could make it work.
- One tries to execute these syntactic transformations.
  (If they fail, they may also suggest other syntactic transformations, which are then executed.)
- We retry the cryptographic transformation, and so on.

Proof of the example: initial game

$$Q_0 = start(); \textbf{new } r : keyseed; \textbf{let } k : key = kgen(r) \textbf{ in}$$
$$\textbf{new } r' : mkeyseed; \textbf{let } mk : mkey = mkgen(r') \textbf{ in } \overline{c}\langle\rangle; (Q_A \mid Q_B)$$

$$Q_A = !^{i \leq n} c_A(); \textbf{new } k' : key; \textbf{new } r'' : coins;$$
$$\textbf{let } m : bitstring = enc(k2b(k'), k, r'') \textbf{ in}$$
$$\overline{c_A}\langle m, mac(m, mk)\rangle$$

$$Q_B = !^{i' \leq n} c_B(m' : bitstring, ma : macstring);$$
$$\textbf{if } verify(m', mk, ma) \textbf{ then}$$
$$\textbf{let } i_\perp(k2b(k'')) = dec(m', k) \textbf{ in } \overline{c_B}\langle\rangle$$

Proof of the example: remove assignments *mk*

$$\begin{aligned} Q_0 = {} & \mathit{start}(); \textbf{new } r : \mathit{keyseed}; \textbf{let } k : \mathit{key} = \mathit{kgen}(r) \textbf{ in} \\ & \textbf{new } r' : \mathit{mkeyseed}; \overline{c}\langle\rangle; (Q_A \mid Q_B) \end{aligned}$$

$$\begin{aligned} Q_A = {} & !^{i \le n} c_A(); \textbf{new } k' : \mathit{key}; \textbf{new } r'' : \mathit{coins}; \\ & \textbf{let } m : \mathit{bitstring} = \mathit{enc}(\mathit{k2b}(k'), k, r'') \textbf{ in} \\ & \overline{c_A}\langle m, \mathit{mac}(m, \mathit{mkgen}(r'))\rangle \end{aligned}$$

$$\begin{aligned} Q_B = {} & !^{i' \le n} c_B(m' : \mathit{bitstring}, \mathit{ma} : \mathit{macstring}); \\ & \textbf{if } \mathit{verify}(m', \mathit{mkgen}(r'), \mathit{ma}) \textbf{ then} \\ & \textbf{let } i_\perp(\mathit{k2b}(k'')) = \mathit{dec}(m', k) \textbf{ in } \overline{c_B}\langle\rangle \end{aligned}$$

Proof of the example: security of the MAC

$$Q_0 = start(); \textbf{new}\ r : keyseed; \textbf{let}\ k : key = kgen(r)\ \textbf{in}$$
$$\textbf{new}\ r' : mkeyseed; \overline{c}\langle\rangle; (Q_A \mid Q_B)$$

$$Q_A = !^{i \leq n} c_A(); \textbf{new}\ k' : key; \textbf{new}\ r'' : coins;$$
$$\textbf{let}\ m : bitstring = enc(k2b(k'), k, r'')\ \textbf{in}$$
$$\overline{c_A}\langle m, mac'(m, mkgen'(r'))\rangle$$

$$Q_B = !^{i' \leq n} c_B(m' : bitstring, ma : macstring);$$
$$\textbf{find}\ j \leq n\ \textbf{suchthat}\ \textbf{defined}(m[j]) \wedge m' = m[j] \wedge$$
$$verify'(m', mkgen'(r'), ma)\ \textbf{then}$$
$$\textbf{let}\ i_\perp(k2b(k'')) = dec(m', k)\ \textbf{in}\ \overline{c_B}\langle\rangle$$

Probability: $\mathrm{Succ}^{\mathrm{uf-cma}}_{\mathrm{MAC}}(\textbf{time} + \textbf{time}(kgen) + n\,\textbf{time}(enc, \mathrm{length}(key)) + n\,\textbf{time}(dec, \mathrm{maxl}(m')), n, n, \max(\mathrm{maxl}(m'), \mathrm{maxl}(m)))$.

Proof of the example: simplify

$$
\begin{array}{l}
Q_0 = \mathit{start}(); \textbf{new } r : \mathit{keyseed}; \textbf{let } k : \mathit{key} = \mathit{kgen}(r) \textbf{ in} \\
\qquad \textbf{new } r' : \mathit{mkeyseed}; \overline{c}\langle\rangle; (Q_A \mid Q_B)
\end{array}
$$

$$
\begin{array}{l}
Q_A = !^{i \leq n} c_A(); \textbf{new } k' : \mathit{key}; \textbf{new } r'' : \mathit{coins}; \\
\qquad \textbf{let } m : \mathit{bitstring} = \mathit{enc}(\mathit{k2b}(k'), k, r'') \textbf{ in} \\
\qquad \overline{c_A}\langle m, \mathit{mac}'(m, \mathit{mkgen}'(r'))\rangle
\end{array}
$$

$$
\begin{array}{l}
Q_B = !^{i' \leq n} c_B(m' : \mathit{bitstring}, \mathit{ma} : \mathit{macstring}); \\
\qquad \textbf{find } j \leq n \textbf{ suchthat defined}(m[j]) \wedge m' = m[j] \wedge \\
\qquad\quad \mathit{verify}'(m', \mathit{mkgen}'(r'), \mathit{ma}) \textbf{ then} \\
\qquad \textbf{let } k'' = k'[j] \textbf{ in } \overline{c_B}\langle\rangle
\end{array}
$$

$$\mathit{dec}(m', k) = \mathit{dec}(\mathit{enc}(\mathit{k2b}(k'[j]), k, r''[j]), k) = i_\perp(\mathit{k2b}(k'[j]))$$

Proof of the example: remove assignments k

$$Q_0 = \mathit{start}();\textbf{new } r : \mathit{keyseed};\textbf{new } r' : \mathit{mkeyseed};\overline{c}\langle\rangle;(Q_A \mid Q_B)$$

$$\begin{aligned}Q_A = !^{i\leq n} c_A();\textbf{new } k' : \mathit{key};\textbf{new } r'' : \mathit{coins};\\ \textbf{let } m : \mathit{bitstring} = \mathit{enc}(k2b(k'), \mathit{kgen}(r), r'') \textbf{ in}\\ \overline{c_A}\langle m, \mathit{mac}'(m, \mathit{mkgen}'(r'))\rangle\end{aligned}$$

$$\begin{aligned}Q_B = !^{i'\leq n} c_B(m' : \mathit{bitstring}, \mathit{ma} : \mathit{macstring});\\ \textbf{find } j \leq n \textbf{ suchthat defined}(m[j]) \wedge m' = m[j] \wedge\\ \mathit{verify}'(m', \mathit{mkgen}'(r'), \mathit{ma}) \textbf{ then}\\ \textbf{let } k'' = k'[j] \textbf{ in } \overline{c_B}\langle\rangle\end{aligned}$$

Proof of the example: security of the encryption

$$Q_0 = \mathit{start}(); \textbf{new } r : \mathit{keyseed}; \textbf{new } r' : \mathit{mkeyseed}; \overline{c}\langle\rangle; (Q_A \mid Q_B)$$

$$\begin{aligned}
Q_A = {}& !^{i \leq n} c_A(); \textbf{new } k' : \mathit{key}; \textbf{new } r'' : \mathit{coins};\\
& \textbf{let } m : \mathit{bitstring} = \mathit{enc}'(Z(k2b(k')), \mathit{kgen}'(r), r'') \textbf{ in}\\
& \overline{c_A}\langle m, \mathit{mac}'(m, \mathit{mkgen}'(r'))\rangle
\end{aligned}$$

$$\begin{aligned}
Q_B = {}& !^{i' \leq n} c_B(m' : \mathit{bitstring}, \mathit{ma} : \mathit{macstring});\\
& \textbf{find } j \leq n \textbf{ suchthat defined}(m[j]) \wedge m' = m[j] \wedge\\
& \quad \mathit{verify}'(m', \mathit{mkgen}'(r'), \mathit{ma}) \textbf{ then}\\
& \textbf{let } k'' = k'[j] \textbf{ in } \overline{c_B}\langle\rangle
\end{aligned}$$

Probability: $\mathrm{Succ}^{\mathrm{ind-cpa}}_{\mathrm{SE}}(\textbf{time} + (n + n^2)\textbf{time}(\mathit{mkgen}) + n\,\textbf{time}(\mathit{mac}, \mathrm{maxl}(m)) + n^2\,\textbf{time}(\mathit{verify}, \mathrm{maxl}(m')) + n^2\,\textbf{time}(= \mathit{bitstring}, \mathrm{maxl}(m'), \mathrm{maxl}(m)), n, \mathrm{length}(\mathit{key}))$

Proof of the example: security of the encryption

$$Q_0 = start(); \textbf{new } r : keyseed; \textbf{new } r' : mkeyseed; \overline{c}\langle\rangle; (Q_A \mid Q_B)$$

$$\begin{aligned}Q_A = {} & !^{i \leq n} c_A(); \textbf{new } k' : key; \textbf{new } r'' : coins;\\ & \textbf{let } m : bitstring = enc'(Z(k2b(k')), kgen'(r), r'') \textbf{ in}\\ & \overline{c_A}\langle m, mac'(m, mkgen'(r'))\rangle\end{aligned}$$

$$\begin{aligned}Q_B = {} & !^{i' \leq n} c_B(m' : bitstring, ma : macstring);\\ & \textbf{find } j \leq n \textbf{ suchthat defined}(m[j]) \wedge m' = m[j] \wedge\\ & \quad verify'(m', mkgen'(r'), ma) \textbf{ then}\\ & \textbf{let } k'' = k'[j] \textbf{ in } \overline{c_B}\langle\rangle\end{aligned}$$

Better probability: $\mathrm{Succ}_{\mathrm{SE}}^{\mathrm{ind-cpa}}(\textbf{time} + (n + n^2)\textbf{time}(mkgen) + n\,\textbf{time}(mac, \mathrm{maxl}(m)) + n^2\,\textbf{time}(verify, \mathrm{maxl}(m')) + n^2\,\textbf{time}(= bitstring, \mathrm{maxl}(m'), \mathrm{maxl}(m)), n, \mathrm{length}(key))$

Proof of the example: simplify

$$Q_0 = \mathit{start}(); \textbf{new } r : \mathit{keyseed}; \textbf{new } r' : \mathit{mkeyseed}; \overline{c}\langle\rangle; (Q_A \mid Q_B)$$

$$\begin{aligned}
Q_A = {} & !^{i \leq n} c_A(); \textbf{new } k' : \mathit{key}; \textbf{new } r'' : \mathit{coins};\\
& \textbf{let } m : \mathit{bitstring} = \mathit{enc}'({\color{red}Z_k}, \mathit{kgen}'(r), r'') \textbf{ in}\\
& \overline{c_A}\langle m, \mathit{mac}'(m, \mathit{mkgen}'(r'))\rangle
\end{aligned}$$

$$\begin{aligned}
Q_B = {} & !^{i' \leq n} c_B(m' : \mathit{bitstring}, ma : \mathit{macstring});\\
& \textbf{find } j \leq n \textbf{ suchthat defined}(m[j]) \wedge m' = m[j] \wedge\\
& \quad \mathit{verify}'(m', \mathit{mkgen}'(r'), ma) \textbf{ then}\\
& \textbf{let } k'' = k'[j] \textbf{ in } \overline{c_B}\langle\rangle
\end{aligned}$$

$$Z(k2b(k')) = Z_k$$

Proof of the example: secrecy

$$Q_0 = \mathit{start}(); \textbf{new } r : \mathit{keyseed}; \textbf{new } r' : \mathit{mkeyseed}; \overline{c}\langle\rangle; (Q_A \mid Q_B)$$

$$\begin{aligned} Q_A = {} & !^{i \leq n} c_A(); \textbf{new } k' : \mathit{key}; \textbf{new } r'' : \mathit{coins};\\ & \textbf{let } m : \mathit{bitstring} = \mathit{enc}'(Z_k, \mathit{kgen}'(r), r'') \textbf{ in}\\ & \overline{c_A}\langle m, \mathit{mac}'(m, \mathit{mkgen}'(r'))\rangle \end{aligned}$$

$$\begin{aligned} Q_B = {} & !^{i' \leq n} c_B(m' : \mathit{bitstring}, \mathit{ma} : \mathit{macstring});\\ & \textbf{find } j \leq n \textbf{ suchthat defined}(m[j]) \wedge m' = m[j] \wedge\\ & \quad \mathit{verify}'(m', \mathit{mkgen}'(r'), \mathit{ma}) \textbf{ then}\\ & \textbf{let } k'' = k'[j] \textbf{ in } \overline{c_B}\langle\rangle \end{aligned}$$

Preserves the one-session secrecy of k'' but not its secrecy.

Final result

Adding the probabilities, we obtain:

Result

The probability that an adversary that runs in time at most t, that executes n sessions of A and B and sends messages of length at most l_{mB} to B breaks the one-session secrecy of k'' is

$$2\mathrm{Succ}_{\mathrm{MAC}}^{\mathrm{uf-cma}}(t'_1, n, n, \max(l_{mB}, l_c)) + 2\mathrm{Succ}_{\mathrm{SE}}^{\mathrm{ind-cpa}}(t'_2, n, l_k)$$

where $t'_1 = t + \mathbf{time}(kgen) + n\,\mathbf{time}(enc, l_k) + n\,\mathbf{time}(dec, l_{mB})$
$t'_2 = t + (n + n^2)\mathbf{time}(mkgen) + n\,\mathbf{time}(mac, l_c) + n^2\,\mathbf{time}(verify, l_{mB}) + n^2\,\mathbf{time}(= bitstring, l_{mB}, l_c)$
l_k is the length of keys, l_c the length of encryptions of keys.

The factor 2 comes from the definition of secrecy.

Example of the FDH signature (joint work with D. Pointcheval)

hash hash function (in the random oracle model)
$\mathrm{f}(pk, m)$ one-way trapdoor permutation, with inverse $\mathrm{invf}(sk, m)$.

We define a signature scheme as follows:

- signature $\mathrm{sign}(m, sk) = \mathrm{invf}(sk, \mathrm{hash}(hk, m))$
- verification $\mathrm{verify}(m, pk, s) = (\mathrm{f}(pk, s) = \mathrm{hash}(hk, m))$

Our goal is to show that this signature scheme is UF-CMA
(secure against existential forgery under chosen message attacks).

Reminder: UF-CMA signatures

The advantage of the adversary:

$$\max_{\mathcal{A}} \Pr\left[\begin{array}{l}(pk, sk) \xleftarrow{R} kgen; (m, s) \leftarrow \mathcal{A}^{sign(.,sk)}(pk) : verify(m, pk, s) \wedge \\ m \text{ was never queried to the oracle } sign(., sk)\end{array}\right]$$

is small.

Formalizing the security of a signature scheme (1)

Key generation:

$start(); \mathbf{new}\ r : keyseed; \mathbf{let}\ pk = \mathrm{pkgen}(r)\ \mathbf{in\ let}\ sk = \mathrm{skgen}(r)\ \mathbf{in}\ \overline{c0}\langle pk\rangle$

Chooses a random seed uniformly in the set of bit-strings *keyseed* (consisting of all bit-strings of a certain length), generates a public key *pk*, a secret key *sk*, and outputs the public key.

Formalizing the security of a signature scheme (2)

Signature:

$$c1(m : \mathit{bitstring}); \overline{c2}\langle \mathrm{sign}(sk, m) \rangle$$

Formalizing the security of a signature scheme (2)

Signature:

$$c1(m : \mathit{bitstring}); \overline{c2}\langle \mathrm{sign}(sk, m)\rangle$$

This process can be called at most q_S times:

$$!^{i_S \leq q_S} c1(m : \mathit{bitstring}); \overline{c2}\langle \mathrm{sign}(sk, m)\rangle$$

Formalizing the security of a signature scheme (2)

Signature:

$$c1(m : \mathit{bitstring}); \overline{c2}\langle \mathrm{sign}(sk, m)\rangle$$

This process can be called at most q_S times:

$$!^{i_S \leq q_S} c1(m : \mathit{bitstring}); \overline{c2}\langle \mathrm{sign}(sk, m)\rangle$$

In fact, this is an abbreviation for:

$$!^{i_S \leq q_S} c1(m[i_S] : \mathit{bitstring}); \overline{c2}\langle \mathrm{sign}(sk, m[i_S])\rangle$$

The variables in repeated oracles are arrays, with one cell for each call, to remember the values used in each oracle call.
These arrays are indexed with the call number i_S.

Formalizing the security of a signature scheme (3)

Test:

$$c3(m' : \mathit{bitstring}, s : D); \textbf{if } \mathrm{verify}(m', pk, s) \textbf{ then}$$
$$\textbf{find } j \leq q_S \textbf{ suchthat defined}(m[j]) \wedge (m' = m[j])$$
$$\textbf{then yield else event } \mathrm{bad})$$

If s is a signature for m' and the signed message m' is not contained in the array m of messages passed to signing oracle, then the signature is a forgery, so we execute **event** bad.

Formalizing the security of a signature scheme (summary)

The signature and test oracles make sense only after the key generation oracle has been called, hence a sequential composition.

The signature and test oracles are simultaneously available, hence a parallel composition.

```
$start(); \textbf{new}\ r : keyseed; \textbf{let}\ pk = \mathrm{pkgen}(r)\ \textbf{in let}\ sk = \mathrm{skgen}(r)\ \textbf{in}\ \overline{c0}\langle pk\rangle;$
 ((* signature oracle *)
   $!^{is \leq q_S} c1(m : bitstring); \overline{c2}\langle \mathrm{sign}(sk, m)\rangle$
 | (* forged signature? *)
   $c3(m' : bitstring, s : D); \textbf{if}\ \mathrm{verify}(m', pk, s)\ \textbf{then}$
   $\textbf{find}\ j \leq q_S\ \textbf{suchthat defined}(m[j]) \wedge (m' = m[j])$
   $\textbf{then yield else event}$ bad)
```

The probability of executing **event** bad in this game is the probability of forging a signature.

Application to the FDH signature scheme

We add a hash oracle because the adversary must be able to call the random oracle (even though it cannot be implemented).

$start(); \ \textbf{new}\ hk : hashkey; \textbf{new}\ r : keyseed;$

$\textbf{let}\ sk = \mathrm{skgen}(r)\ \textbf{in let}\ pk = \mathrm{pkgen}(r)\ \textbf{in}\ \overline{c0}\langle pk\rangle;$

$(($(* hash oracle *)$)\ !^{i_H \leq q_H} hc1(x : bitstring); \ \overline{hc2}\langle \mathrm{hash}(hk, x)\rangle$

$\mid$ (* signature oracle *) $!^{i_S \leq q_S} c1(m : bitstring); \ \overline{c2}\langle \mathrm{invf}(sk, \mathrm{hash}(hk, m))\rangle$

$\mid$ (* forged signature? *)

$c3(m' : bitstring, s : D); \textbf{if}\ \mathrm{f}(pk, s) = \mathrm{hash}(hk, m')\ \textbf{then}$

$\textbf{find}\ j \leq q_S\ \textbf{suchthat defined}(m[j]) \wedge (m' = m[j])$

$\textbf{then yield else event}\ \mathrm{bad})$

Our goal is to bound the probability that **event** bad is executed in this game.

This game is given as input to the prover in the syntax above.

FDH: security of a hash function

A hash function is equivalent to a “random function”: a function that

- returns a new random number when it is a called on a new argument,
- and returns the same result when it is called on the same argument.

$!^{Nh}$ **new** $k : \mathit{hashkey}; !^{N} \mathit{Ohash}(x : \mathit{bitstring}) := \mathrm{hash}(k, x)$

$\approx_0$

$!^{Nh}$ **new** $k : \mathit{hashkey}; !^{N} \mathit{Ohash}(x : \mathit{bitstring}) :=$

```
    find $j \leq N$ suchthat defined$(x[j], r[j])$ && $(x = x[j])$
    then $r[j]$
    else new $r : D$; $r$
```

FDH: security of a hash function (optimized)

For a test $r' = h(x')$, we can avoid computing $h(x')$ explicitly:

- if x' has been passed to the hash function previously, compare r' with the previous result;
- otherwise, return false.

FDH: security of a hash function (optimized)

For a test $r' = h(x')$, we can avoid computing $h(x')$ explicitly:

- if x' has been passed to the hash function previously, compare r' with the previous result;
- otherwise, return false.

In the latter case, test indeed false, except when the fresh random number $h(x')$ collides with r' (probability $1/|D|$).

FDH: security of a hash function (optimized)

For a test $r' = h(x')$, we can avoid computing $h(x')$ explicitly:

- if x' has been passed to the hash function previously, compare r' with the previous result;
- otherwise, return false.

In the latter case, test indeed false, except when the fresh random number $h(x')$ collides with r' (probability $1/|D|$).

$$
\begin{array}{l}
!^{Nh}\ \textbf{new}\ k : \mathit{hashkey};\\
\quad (!^{N} \mathit{Ohash}(x : \mathit{bitstring}) := \mathrm{hash}(k, x),\\
\quad\ !^{Neq} \mathit{Oeq}(x' : \mathit{bitstring}; r' : D) := r' = \mathrm{hash}(k, x'))\\
\approx_{\#\mathit{Oeq}/|D|}\\
!^{Nh} (!^{N} \mathit{Ohash}(x : \mathit{bitstring}) := \textbf{find}\ j \leq N\ \textbf{suchthat}\\
\qquad\quad \textbf{defined}(x[j], r[j])\ \&\&\ (x = x[j])\ \textbf{then}\ r[j]\ \textbf{else new}\ r : D;\ \ r,\\
\quad\ !^{Neq} \mathit{Oeq}(x' : \mathit{bitstring}; r' : D) := \textbf{find}\ j \leq N\ \textbf{suchthat}\\
\qquad\quad \textbf{defined}(x[j], r[j])\ \&\&\ (x' = x[j])\ \textbf{then}\ r' = r[j]\ \textbf{else false})
\end{array}
$$

FDH: one-wayness

The adversary inverts f when, given the public key $pk = \mathrm{pkgen}(r)$ and the image of some x by $\mathrm{f}(pk, \cdot)$, it manages to find x (without having the trapdoor).

The function f is one-way when the adversary has negligible probability of inverting f.

Definition (One-wayness)

$$\mathrm{Succ}^{\mathrm{ow}}_{\mathcal{P}}(t) = \max_{\mathcal{A}} \Pr \left[\begin{array}{l} r \overset{R}{\leftarrow} keyseed, pk \leftarrow \mathrm{pkgen}(r), x \overset{R}{\leftarrow} D, \\ y \leftarrow \mathrm{f}(pk, x), x' \leftarrow \mathcal{A}(pk, y) : x = x' \end{array} \right]$$

where $\mathcal{A}$ runs in time at most t.

FDH: one-wayness (preliminary version)

$$\begin{array}{l}
!^{Nk}\ \mathbf{new}\ r : \mathit{keyseed};\ (\\
\quad \mathit{Opk}() := \mathrm{pkgen}(r),\\
\quad !^{Nf}\ \mathbf{new}\ x : D;\ (\\
\quad\quad \mathit{Oy}() := \mathrm{f}(\mathrm{pkgen}(r), x),\\
\quad\quad !^{N2}\mathit{Oeq}(x' : D) := (x' = x)))\\
\approx\\
!^{Nk}\ \mathbf{new}\ r : \mathit{keyseed};\ (\\
\quad \mathit{Opk}() := \mathrm{pkgen}(r),\\
\quad !^{Nf}\ \mathbf{new}\ x : D;\ (\\
\quad\quad \mathit{Oy}() := \mathrm{f}(\mathrm{pkgen}(r), x),\\
\quad\quad !^{N2}\mathit{Oeq}(x' : D) := \mathrm{false}))
\end{array}$$

FDH: one-wayness

$$\begin{array}{l}
!^{Nk}\ \mathbf{new}\ r : \mathit{keyseed};\ (\\
\quad \mathit{Opk}() := \mathrm{pkgen}(r),\\
\quad !^{Nf}\ \mathbf{new}\ x : D;\ (\\
\quad\quad \mathit{Oy}() := \mathrm{f}(\mathrm{pkgen}(r), x),\\
\quad\quad !^{N2}\mathit{Oeq}(x' : D) := (x' = x),\\
\quad\quad \mathit{Ox}() := x))\\
\approx\\
!^{Nk}\ \mathbf{new}\ r : \mathit{keyseed};\ (\\
\quad \mathit{Opk}() := \mathrm{pkgen}(r),\\
\quad !^{Nf}\ \mathbf{new}\ x : D;\ (\\
\quad\quad \mathit{Oy}() := \mathrm{f}(\mathrm{pkgen}(r), x),\\
\quad\quad !^{N2}\mathit{Oeq}(x' : D) := \mathbf{if}\ \mathbf{defined}(k)\ \mathbf{then}\ x' = x\ \mathbf{else}\ \mathrm{false},\\
\quad\quad \mathit{Ox}() := \mathbf{let}\ k : \mathit{bitstring} = \mathrm{mark}\ \mathbf{in}\ x))
\end{array}$$

FDH: one-wayness

```
$!^{Nk}$ $\mathbf{new}\ r : \mathit{keyseed};\ ($
  $Opk() := \mathrm{pkgen}(r),$
  $!^{Nf}$ $\mathbf{new}\ x : D;\ ($
    $Oy() := \mathrm{f}(\mathrm{pkgen}(r), x),$
    $!^{N2} Oeq(x' : D) := (x' = x),$
    $Ox() := x))$
$\approx_{Nk \times Nf \times \mathrm{Succ}^{\mathrm{ow}}_{\mathcal{P}}(\mathbf{time} + (Nk-1)\times\mathbf{time}(\mathrm{pkgen}) + (\#Oy-1)\times\mathbf{time}(\mathrm{f}))}$
$!^{Nk}$ $\mathbf{new}\ r : \mathit{keyseed};\ ($
  $Opk() := \mathrm{pkgen}'(r),$
  $!^{Nf}$ $\mathbf{new}\ x : D;\ ($
    $Oy() := \mathrm{f}'(\mathrm{pkgen}'(r), x),$
    $!^{N2} Oeq(x' : D) := \mathbf{if}\ \mathbf{defined}(k)\ \mathbf{then}\ x' = x\ \mathbf{else}\ \mathrm{false},$
    $Ox() := \mathbf{let}\ k : \mathit{bitstring} = \mathrm{mark}\ \mathbf{in}\ x))$
```

FDH: other properties of one-way trapdoor permutations (1)

invf is the inverse of f:

$$\forall r : \mathit{keyseed}, x : D; \mathrm{invf}(\mathrm{skgen}(r), \mathrm{f}(\mathrm{pkgen}(r), x)) = x$$

f is injective:

$$\forall k : \mathit{key}, x : D, x' : D; (\mathrm{f}(k, x) = \mathrm{f}(k, x')) = (x = x')$$

FDH: other properties of one-way trapdoor permutations (2)

We can replace a uniformly distributed random number y with $\mathrm{f}(\mathrm{pkgen}(r), x)$ where x is a uniformly distributed random number:

$$!^{Nf}\ \textbf{new}\ y : D; Oim() := y$$

$$\approx_0$$

$$!^{Nf}\ \textbf{new}\ x : D; Oim() := \mathrm{f}(\mathrm{pkgen}(r), x)$$

FDH: other properties of one-way trapdoor permutations (2)

We can replace a uniformly distributed random number y with $\mathrm{f}(\mathrm{pkgen}(r), x)$ where x is a uniformly distributed random number:

$$!^{Nf}\ \textbf{new}\ y : D; Oim() := y$$
$$\approx_0$$
$$!^{Nf}\ \textbf{new}\ x : D; Oim() := \mathrm{f}(\mathrm{pkgen}(r), x)$$

$\Rightarrow$ Can be applied too often!

FDH: other properties of one-way trapdoor permutations (2)

We can replace a uniformly distributed random number y with $\mathrm{f}(\mathrm{pkgen}(r), x)$ where x is a uniformly distributed random number:

$$
\begin{array}{l}
!^{Nk}\ \textbf{new}\ r : \mathit{keyseed}; (\\
\quad \mathit{Opk}() := \mathrm{pkgen}(r),\\
\quad !^{Nf}\ \textbf{new}\ y : D; (\mathit{Oant}() := \mathrm{invf}(\mathrm{skgen}(r), y), \mathit{Oim}() := y))\\
\approx_0\\
!^{Nk}\ \textbf{new}\ r : \mathit{keyseed}; (\\
\quad \mathit{Opk}() := \mathrm{pkgen}(r),\\
\quad !^{Nf}\ \textbf{new}\ x : D; (\mathit{Oant}() := x, \mathit{Oim}() := \mathrm{f}(\mathrm{pkgen}(r), x)))
\end{array}
$$

Demo

- CryptoVerif input file: examples/fdh
- library of primitives
- run CryptoVerif
- output

FDH: initial game

```
$\mathit{start}()$; new $hk : \mathit{hashkey}$; new $r : \mathit{keyseed}$;
let $sk : \mathit{key} = \mathrm{skgen}(r)$ in
let $pk : \mathit{key} = \mathrm{pkgen}(r)$ in $\overline{c0}\langle pk \rangle$;
( (* hash oracle *)
  $!^{i_H \leq q_H} hc1[i_H](x : \mathit{bitstring})$; $\overline{hc2[i_H]}\langle \mathrm{hash}(hk, x) \rangle$
| (* signature oracle *)
  $!^{i_S \leq q_S} c1[i_S](m : \mathit{bitstring})$; $\overline{c2[i_S]}\langle \mathrm{invf}(sk, \mathrm{hash}(hk, m)) \rangle$
| (* forged signature? *)
  $c3(m' : \mathit{bitstring}, s : D)$;
  if $\mathrm{f}(pk, s) = \mathrm{hash}(hk, m')$ then
  find $j \leq q_S$ suchthat defined$(m[j])$ && $(m' = m[j])$ then
    yield
  else
    event bad
)
```

FDH step 1: apply the security of the hash function

Replace each occurrence of hash(M) with a lookup in the arguments of previous calls to hash.

- If M is found, return the same result as the previous result.
- Otherwise, pick a new random number and return it.

For instance, $\overline{hc2[i_H]}\langle \mathrm{hash}(hk, x)\rangle$ is replaced with

```
find $@i1 \leq q_S$ suchthat defined$(m[@i1], r\_32[@i1])$
   $\&\& \ (x = m[@i1])$ then $\overline{hc2[i_H]}\langle r\_32[@i1]\rangle$
orfind $@i2 \leq q_H$ suchthat defined$(x[@i2], r\_34[@i2])$
   $\&\& \ (x = x[@i2])$ then $\overline{hc2[i_H]}\langle r\_34[@i2]\rangle$
else
   new $r\_34 : D$; $\overline{hc2[i_H]}\langle r\_34\rangle$
```

The test $\mathrm{f}(pk, s) = \mathrm{hash}(hk, m')$ uses Oeq. Probability difference $1/|D|$.

FDH step 2: simplify

```
(* forged signature? *)
$c3(m' : bitstring, s : D)$;
find $@i5 \leq q_S$ suchthat defined($m[@i5], r\_32[@i5]$) && ($m' = m[@i5]$) then
   if ($f(pk, s) = r\_32[@i5]$) then
   find $j \leq q_S$ suchthat defined($m[j]$) && ($m' = m[j]$) then yield else event bad
orfind $@i6 \leq q_H$ suchthat defined($x[@i6], r\_34[@i6]$) && ($m' = x[@i6]$) then
   if ($f(pk, s) = r\_34[@i6]$) then
   find $j \leq q_S$ suchthat defined($m[j]$) && ($m' = m[j]$) then yield else event bad
else
   if false then
   find $j \leq q_S$ suchthat defined($m[j]$) && ($m' = m[j]$) then yield else event bad
```

The red test always succeeds, so the blue part becomes **yield**.

The magenta part becomes **yield**.

FDH step 3: substitute *sk* with its value

The variable *sk* is replaced with skgen(r), and the assignment **let** $sk : key = \mathrm{skgen}(r)$ is removed.
This transformation is advised in order to able to apply the permutation property.

FDH step 4: permutation

```
(* signature oracle *)
$!^{i_S \leq q_S}$
$c1[i_S](m : bitstring);$
find $@i3 \leq q_S$ suchthat defined($m[@i3], r\_32[@i3]$) && ($m = m[@i3]$) then
   $\overline{c2[i_S]}\langle \mathrm{invf}(\mathrm{skgen}(r), r\_32[@i3])\rangle$
orfind $@i4 \leq q_H$ suchthat defined($x[@i4], r\_34[@i4]$) && ($m = x[@i4]$) then
   $\overline{c2[i_S]}\langle \mathrm{invf}(\mathrm{skgen}(r), r\_34[@i4])\rangle$
else
   new $r\_32 : D$;
   $\overline{c2[i_S]}\langle \mathrm{invf}(\mathrm{skgen}(r), r\_32)\rangle$
```

new $r\_i : D$ becomes **new** $y\_i : D$,

$\mathrm{invf}(\mathrm{skgen}(r), r\_i)$ becomes y_i,

r_i becomes $\mathrm{f}(\mathrm{pkgen}(r), y\_i)$

FDH step 5: simplify

```
(* forged signature? *)
$c3(m' : bitstring, s : D);$
find $@i5 \leq q_S$ suchthat defined($m[@i5], r\_32[@i5]$) && ($m' = m[@i5]$) then
  yield
orfind $@i6 \leq q_H$ suchthat defined($x[@i6], r\_34[@i6]$) && ($m' = x[@i6]$) then
  if ($\mathrm{f}(pk, s) = \mathrm{f}(\mathrm{pkgen}(r), y\_34[@i6])$) then
  find $j \leq q_S$ suchthat defined($m[j]$) && ($m' = m[j]$) then yield else event bad
```

$\mathrm{f}(pk, s) = \mathrm{f}(\mathrm{pkgen}(r), y\_i)$ becomes $s = y\_i$,
knowing $pk = pkgen(r)$ and the injectivity of f:
$\forall k : key, x : D, x' : D; (\mathrm{f}(k, x) = \mathrm{f}(k, x')) = (x = x')$

FDH step 6: one-wayness

```
(* forged signature? *)
$c3(m' : bitstring, s : D)$;
find $@i5 \leq q_S$ suchthat defined($m[@i5], r\_32[@i5]$) && ($m' = m[@i5]$) then
  yield
orfind $@i6 \leq q_H$ suchthat defined($x[@i6], r\_34[@i6]$) && ($m' = x[@i6]$) then
  if $s = y\_34[@i6]$ then
  find $j \leq q_S$ suchthat defined($m[j]$) && ($m' = m[j]$) then yield else event bad
```

$s = y\_i$ becomes **find** $@j\_i \leq q_H$ **suchthat defined**$(k\_i[@j\_i])$ **then** $s = y\_i$ **else** false,

In hash oracle, $\mathrm{f}(\mathrm{pkgen}(r), y\_i)$ becomes $\mathrm{f}'(\mathrm{pkgen}'(r), y\_i)$,

In signature oracle, y_i becomes **let** $k\_i : bitstring = \mathrm{mark}$ **in** y_i.

Difference of probability: $(q_H + q_S)\mathrm{Succ}_{\mathcal{P}}^{\mathrm{ow}}(\mathbf{time} + (q_H - 1)\mathbf{time}(f))$.

FDH step 7: simplify

```
(* forged signature? *)
$c3(m' : bitstring, s : D);$
find $@i5 \leq q_S$ suchthat defined($m[@i5], r\_32[@i5]$) && ($m' = m[@i5]$) then
  yield
orfind $@i6 \leq q_H$ suchthat defined($x[@i6], r\_34[@i6]$) && ($m' = x[@i6]$) then
  find $@j\_34 \leq q_S$ suchthat defined($k\_34[@j\_34]$) && ($@i4[@j\_34] = @i6$) then
  if $s = y\_34[@i6]$ then
  find $j \leq q_S$ suchthat defined($m[j]$) && ($m' = m[j]$) then yield else event bad
```

The test in red always succeeds, so **event** bad disappears, which proves the desired property.

FDH step 7: simplify (2)

```
(* forged signature? *)
$c3(m' : bitstring, s : D);$
...
orfind $@i6 \leq q_H$ suchthat defined$(x[@i6], r\_34[@i6])$ && $(m' = x[@i6])$ then
   find $@j\_34 \leq q_S$ suchthat defined$(k\_34[@j\_34])$ && $(@i4[@j\_34] = @i6)$ then
   if $s = y\_34[@i6]$ then
   find $j \leq q_S$ suchthat defined$(m[j])$ && $(m' = m[j])$ then yield else event bad
```

Definition of k_34:

```
$!^{i_S \leq q_S}$
$c1[i_S](m : bitstring);$
...
orfind $@i4 \leq q_H$ suchthat defined$(x[@i4], y\_34[@i4])$ && $(m = x[@i4])$ then
   let $k\_34 : bitstring =$ mark in ...
```

When $k\_34[@j\_34]$ is defined, $m[@j\_34]$ is defined and
$m[@j\_34] = x[@i4[@j\_34]] = x[@i6] = m'$
so the red test succeeds with $j = @j\_34$.

FDH: final result

Adding the probabilities, we obtain:

Result

The probability that an adversary that runs in time at most t and makes q_S signature queries and q_H hash queries forges a FDH signature is at most

$$1/|D| + (q_S + q_H)\mathrm{Succ}_{\mathcal{P}}^{\mathrm{ow}}(t + (q_H - 1)\mathbf{time}(f))$$

Experiments

Tested on the following protocols (original and corrected versions):

- Otway-Rees (shared-key)
- Yahalom (shared-key)
- Denning-Sacco (public-key)
- Woo-Lam shared-key and public-key
- Needham-Schroeder shared-key and public-key

Shared-key encryption is implemented as encrypt-then-MAC, using a IND-CPA encryption scheme.

(For Otway-Rees, we also considered a SPRP encryption scheme,
a IND-CPA + INT-CTXT encryption scheme,
a IND-CCA2 + IND-PTXT encryption scheme.)

Public-key encryption is assumed to be IND-CCA2.

We prove secrecy of session keys and correspondence properties.

Results (1)

In most cases, the prover succeeds in proving the desired properties when they hold, and obviously it always fails to prove them when they do not hold.

Only cases in which the prover fails although the property holds:

- Needham-Schroeder public-key when the exchanged key is the nonce N_A.
- Needham-Schroeder shared-key: fails to prove that $N_B[i] \neq N_B[i'] - 1$ with overwhelming probability, where N_B is a nonce

Results (2)

- Some public-key protocols need manual proofs.
  (Give the cryptographic proof steps and single assignment renaming instructions.)
- Runtime: 7 ms to 35 s, average: 5 s on a Pentium M 1.8 GHz.

Other case studies

- Full domain hash signature (with David Pointcheval)
  Encryption schemes of Bellare-Rogaway'93 (with David Pointcheval)
- Kerberos V, with and without PKINIT (with Aaron D. Jaggard, Andre Scedrov, and Joe-Kai Tsay).
- OEKE (variant of Encrypted Key Exchange, with David Pointcheval).
- SSH Transport Layer Protocol (with David Cadé).
- A part of an F# implementation of the TLS transport protocol (Microsoft Research and MSR-INRIA).

Conclusion

CryptoVerif can automatically prove the security of primitives and protocols.

- The security assumptions are given as observational equivalences (proved manually once).
- The protocol or scheme to prove is specified in a process calculus.
- The prover provides a sequence of indistinguishable games that lead to the proof and a bound on the probability of an attack.
- The user is allowed (but does not have) to interact with the prover to make it follow a specific sequence of games.

Future work: CryptoVerif extensions

- Support more primitives: Primitives with internal state, ...
- Extend the language for games
  - Loops
  - Mutable variables
  - Common code after a test
- Improvements in the proof strategy.
  More precise manual hints?
- More case studies.
  - Will suggest more extensions.
- Certify CryptoVerif; combine it with CertiCrypt.

Other research directions

- Proof of implementations of protocols in the computational model:
  - by generation of implementations from specifications (as CryptoVerif can also do)
  - by analysis of existing implementations (e.g. FS2CV, F*)
- Take into account side-channels.

Acknowledgments

- I warmly thank David Pointcheval for his advice and explanations of the computational proofs of protocols. This project would not have been possible without him.
- This work was partly supported by the ANR project FormaCrypt (ARA SSIA 2005).
- This work is partly supported by the ANR project ProSe (VERSO 2010).