Computationally Sound Mechanized Proofs of Correspondence Assertions

Bruno Blanchet
Inria, Paris
Bruno.Blanchet@inria.fr

Year 2023-24
Our goal: implement an automatic, \textit{computationally sound} prover for security protocols.

We have already discussed \textit{secrecy properties}.

In this talk, we show how to prove to \textit{correspondence assertions}, that is, properties of the style:

\begin{quote}
\textit{If some event has been executed, then some other events have been executed.}
\end{quote}

Basic application: \textit{authentication}. 

Differences with secrecy

For the proof of correspondences:

- The language for games is the same as for secrecy, except for the addition of events.
- The game transformations and the proof strategy are the same as for secrecy. (The events are left unchanged.)
- One needs a new algorithm for checking correspondences on the last game.
Example: a nonce challenge

Simple example inspired by the corrected Woo-Lam public-key protocol (1997)

\[ B \rightarrow A : (N, B) \]
\[ A \rightarrow B : \{ pk_A, B, N \}_{sk_A} \]

In our language:

\[ O_{\text{start}}() := rk_A \xleftarrow{R} \text{keyseed}; \]
\[ \text{let } pk_A = \text{pkgen}(rk_A) \text{ in let } sk_A = \text{skgen}(rk_A) \text{ in return}(pk_A); \]
\[ \text{foreach } i_A \leq n \text{ do } \]
\[ O_A(xN : \text{nonce}, xB : \text{host}) := \text{event } e_A(pk_A, xB, xN); \]
\[ r \xleftarrow{R} \text{seed}; \text{return}(\text{sign}(\text{concat}(pk_A, xB, xN), sk_A, r)) \]
\[ \text{foreach } i_B \leq n \text{ do } \]
\[ O_{B1}(xpk_A : \text{pkey}) := N \xleftarrow{R} \text{nonce}; \text{return}(N, B); \]
\[ O_{B2}(s : \text{signature}) := \text{if } \text{verify}(\text{concat}(xpk_A, B, N), xpk_A, s) \text{ then } \]
\[ \text{if } xpk_A = pk_A \text{ then event } e_B(xpk_A, B, N) \]
Arrays

All variables defined under replications are implicitly arrays. This allows us to store all values that occur during the executions of the game.

This replaces lists used by cryptographers and is key to cryptographic proofs.

\[ O_{\text{start}}() := rk_A \stackrel{R}{\leftarrow} \text{keyseed}; \]

let \( pk_A = \text{pkgen}(rk_A) \) in let \( sk_A = \text{skgen}(rk_A) \) in return(\( pk_A \));

foreground \( i_A \leq n \) do

\[ O_A[i_A](xN[i_A] : \text{nonce}, xB[i_A] : \text{host}) := \text{event} \ e_A(pk_A, xB[i_A], xN[i_A]); \]

\( r[i_A] \stackrel{R}{\leftarrow} \text{seed}; \) return(sign(concat(\( pk_A, xB[i_A], xN[i_A] \)), \( sk_A, r[i_A] \))

foreground \( i_B \leq n \) do

\[ O_{B1}[i_B](xpk_A[i_B] : \text{pkey}) := N[i_B] \stackrel{R}{\leftarrow} \text{nonce}; \) return(\( N[i_B], B \));

\[ O_{B2}[i_B](s[i_B] : \text{signature}) := \]

if verify(concat(xpk_A[i_B], B, N[i_B]), xpk_A[i_B], s[i_B]) then

if xpk_A[i_B] = pk_A then event e_B(xpk_A[i_B], B, N[i_B])
After game transformations

Using the unforgeability of signatures, the signature verification with $pk_A$ succeeds only for signatures generated with $sk_A$. After game transformations, we obtain the last game:

$$O_{\text{start}}() := rk_A \xleftarrow{\text{R}} \text{keyseed}; \text{let } pk_A = \text{pkgen}'(rk_A) \text{ in return}(pk_A);$$

$$\text{foreach } i_A \leq n \text{ do } O_A(xN : \text{nonce}, xB : \text{host}) :=$$

$$\text{event } e_A(pk_A, xB, xN); \text{let } m = \text{concat}(pk_A, xB, xN) \text{ in }$$

$$r \xleftarrow{\text{R}} \text{seed}; \text{return}(\text{sign}'(m, skgen'(rk_A), r))$$

$$\mid \text{foreach } i_B \leq n \text{ do } O_{B1}(xpk_A : \text{pkey}) := N \xleftarrow{\text{R}} \text{nonce}; \text{return}(N, B);$$

$$O_{B2}(s : \text{signature}) := \text{find } u \leq n \text{ such that}$$

$$\text{defined}(m[u], xB[u], xN[u]) \land (xpk_A = pk_A) \land (B = xB[u])$$

$$\land (N = xN[u]) \land \text{verify}'(\text{concat}(xpk_A, B, N), xpk_A, s) \text{ then }$$

$$\text{event } e_B(xpk_A, B, N)$$
A **non-injective correspondence** is a formula of the form $\psi \Rightarrow \phi$ where

\[
\phi ::= \begin{align*}
M \\
\text{event}(e(M_1, \ldots, M_m)) \\
\phi_1 \land \phi_2 \\
\phi_1 \lor \phi_2
\end{align*}
\]

and $\psi$ is a formula that contains only events and conjunctions.

**Example**

\[
\text{event}(e_B(x, y, z)) \Rightarrow \text{event}(e_A(x, y, z))
\]

means that, if $e_B(x, y, z)$ is executed, then $e_A(x, y, z)$ has also been executed (except in cases of negligible probability).
Let $\rho$ be an environment that maps variables to bitstrings. Let $\mathcal{E}$ be a sequence of events.

### Definition

| $\rho, \mathcal{E} \vdash M$ if and only if $M$ evaluates to $true$ in environment $\rho$ |
| $\rho, \mathcal{E} \vdash event(e(M_1, \ldots, M_m))$ if and only if |
| for all $j \leq m$, $M_j$ evaluates to $a_j$ in $\rho$ and $e(a_1, \ldots, a_m) \in \mathcal{E}$ |

### Definition

| $\mathcal{E} \vdash \psi \Rightarrow \phi$ if and only if |
| for all $\rho$ defined on $\text{var}(\psi)$ such that $\rho, \mathcal{E} \vdash \psi$, |
| there exists an extension $\rho'$ of $\rho$ to $\text{var}(\phi)$ such that $\rho', \mathcal{E} \vdash \phi$. |

### Definition

| $Q_0$ satisfies $\psi \Rightarrow \phi$ with public variables $V$ if and only if |
| for all evaluation contexts $C$ accessing only variables of $V$ in $Q_0$, $\Pr[C[Q_0] \text{ executes } \mathcal{E} \text{ and } \mathcal{E} \not\vdash \psi \Rightarrow \phi]$ is negligible. |
An injective correspondence also allows injective events \( \text{inj-event}(e(M_1, \ldots, M_m)) \).

Each execution of the injective events in \( \psi \) corresponds to distinct injective events in \( \phi \).

**Example**

\[
\text{inj-event}(e_B(x, y, z)) \Rightarrow \text{inj-event}(e_A(x, y, z))
\]

means that each execution of \( e_B(x, y, z) \) corresponds to a distinct execution of \( e_A(x, y, z) \).
Prove the correspondence \( \text{event}(e_B(x, y, z)) \Rightarrow \text{event}(e_A(x, y, z)) \) in the game

\[
\begin{aligned}
&\ldots \text{foreach } i_A \leq n \text{ do } \ldots \text{event } e_A(pk_A, xB, xN); \text{let } m = \ldots \text{ in } \ldots \\
&\quad \text{foreach } i_B \leq n \text{ do } \\
&\quad \quad \text{find } u \leq n \text{ such that } \text{defined}(m[u], xB[u], xN[u]) \land \\
&\quad \quad (xpk_A = pk_A) \land (B = xB[u]) \land (N = xN[u]) \land \\
&\quad \quad \text{verify}'(\text{concat}(xpk_A, B, N), xpk_A, s) \text{ then event } e_B(xpk_A, B, N)
\end{aligned}
\]
Prove the correspondence \(\text{event}(e_B(x, y, z)) \Rightarrow \text{event}(e_A(x, y, z))\) in the game

\[\ldots \text{foreach } i_A \leq n \text{ do } \ldots \text{event } e_A(pk_A, xB, xN); \text{let } m = \ldots \text{ in } \ldots\]
\[\mid \text{foreach } i_B \leq n \text{ do } \ldots \]

\[\text{find } u \leq n \text{ such that } \text{defined}(m[u], xB[u], xN[u]) \land (xpk_A = pk_A) \land (B = xB[u]) \land (N = xN[u]) \land \text{verify}'(\text{concat}(xpk_A, B, N), xpk_A, s) \text{ then event } e_B(xpk_A, B, N)\]

If \(\text{event}(e_B(x, y, z))\) has been executed, the program point \(\text{event } e_B(xpk_A, B, N)\) has been reached for some \(i_B = i'_B\), and \(e_B(x, y, z) = e_B(xpk_A[i'_B], B, N[i'_B]).\)

So \(m[u[i'_B]], xB[u[i'_B]], \text{ and } xN[u[i'_B]]\) are defined, \(xpk_A[i'_B] = pk_A, B = xB[u[i'_B]], \text{ and } N[i'_B] = xN[u[i'_B]].\)

Since \(m[u[i'_B]]\) is defined, the definition of \(m[i_A]\) has been executed for \(i_A = u[i'_B]\), so \(\text{event } e_A(pk_A, xB[i_A], xN[i_A])\) has been executed.
If \(\text{event}(e_B(x, y, z))\) has been executed, the program point \(\text{event} \ e_B(xpk_A, B, N)\) has been reached for some \(i_B = i'_B\), and \(e_B(x, y, z) = e_B(xpk_A[i'_B], B, N[i'_B])\).

So \(m[u[i'_B]], x_B[u[i'_B]],\) and \(xN[u[i'_B]]\) are defined, \(xpk_A[i'_B] = pk_A\), \(B = xB[u[i'_B]]\), and \(N[i'_B] = xN[u[i'_B]]\).

Since \(m[u[i'_B]]\) is defined, the definition of \(m[i_A]\) has been executed for \(i_A = u[i'_B]\), so \(\text{event} \ e_A(pk_A, xB[i_A], xN[i_A])\) has been executed.

We have
- \(x = xpk_A[i'_B] = pk_A\)
- \(y = B = xB[u[i'_B]] = xB[i_A]\)
- \(z = N[i'_B] = xN[u[i'_B]] = xN[i_A]\)

so \(e_A(pk_A, xB[i_A], xN[i_A]) = e_A(x, y, z)\) has been executed.
Intuition for the proof: injective correspondences (1)

Prove the correspondence

\( \text{inj-event}(e_B(x, y, z)) \Rightarrow \text{inj-event}(e_A(x, y, z)) \) in the game

\[
\text{... foreach } i_A \leq n \text{ do } \text{... event } e_A(pk_A, xB, xN); \text{ let } m = \ldots \text{ in } \ldots \\
\text{ | foreach } i_B \leq n \text{ do } \ldots \\
\text{ find } u \leq n \text{ suchthat defined}(m[u], xB[u], xN[u]) \wedge \\
(xpk_A = pk_A) \wedge (B = xB[u]) \wedge (N = xN[u]) \wedge \\
\text{verify'}(\text{concat}(xpk_A, B, N), xpk_A, s) \text{ then event } e_B(xpk_A, B, N)
\]
Intuition for the proof: injective correspondences (2)

Prove the correspondence
\[ \text{inj-event}(e_B(i, x, y, z)) \Rightarrow \text{inj-event}(e_A(i', x, y, z)) \]
in the game

\[
\text{foreach } i_A \leq n \text{ do } \text{event } e_A(i_A, pk_A, xB, xN); \text{let } m = \ldots \text{ in } \ldots
\]
\[
| \text{foreach } i_B \leq n \text{ do } \ldots
\]
\[
\text{find } u \leq n \text{ such that } \text{defined}(m[u], xB[u], xN[u]) \land
\]
\[
(xpk_A = pk_A) \land (B = xB[u]) \land (N = xN[u]) \land
\]
\[
\text{verify}'(\text{concat}(xpk_A, B, N), xpk_A, s) \text{ then event } e_B(i_B, xpk_A, B, N)
\]

In order to record in which session each event is executed, we add replication indices to events.
Intuition for the proof: injective correspondences (3)

Prove the correspondence

\[ \text{inj-event}(e_B(i, x, y, z)) \Rightarrow \text{inj-event}(e_A(i', x, y, z)) \]

in the game

\[ \text{foreach } i_A \leq n \text{ do . . . event } e_A(i_A, pk_A, xB, xN); \text{let } m = \ldots \text{ in . . .} \]

\[ \text{foreach } i_B \leq n \text{ do . . .} \]

\[ \text{find } u \leq n \text{ such that defined}(m[u], xB[u], xN[u]) \land (xpk_A = pk_A) \land (B = xB[u]) \land (N = xN[u]) \land \text{verify'}(\text{concat}(xpk_A, B, N), xpk_A, s) \text{ then event } e_B(i_B, xpk_A, B, N) \]

If event(e_B(i, x, y, z)) has been executed, the program point

\[ \text{event } e_B(i_B, xpk_A, B, N) \]

has been reached for some \( i_B = i'_B \), and

\[ e_B(i, x, y, z) = e_B(i'_B, xpk_A[i'_B], B, N[i'_B]). \]

So \( m[u[i'_B]], x_B[u[i'_B]], \) and \( x_N[u[i'_B]] \) are defined,

\[ xpk_A[i'_B] = pk_A, \ B = xB[u[i'_B]], \) and \( N[i'_B] = xN[u[i'_B]] \].

Since \( m[u[i'_B]] \) is defined, the definition of \( m[i_A] \) has been executed

for \( i_A = u[i'_B] \), so \text{event } e_A(i_A, pk_A, xB[i_A], xN[i_A]) \) has been executed.
As before, \( e_A(i_A, pk_A, xB[i_A], xN[i_A]) = e_A(i', x, y, z) \) for some \( i' \).

In order to show injectivity, we show that

- \( e_B \) executed twice, for \( i_B = i'_B \) and \( i_B = i''_B \), with \( i'_B \neq i''_B \)

\( \Rightarrow e_A \) executed twice, for \( i_A = u[i'_B] \) and \( i_A = u[i''_B] \), with \( u[i'_B] \neq u[i''_B] \)

By contraposition, we show \( u[i'_B] = u[i''_B] \Rightarrow i'_B = i''_B \).

\[
\begin{align*}
  u[i'_B] = u[i''_B] & \Rightarrow xN[u[i'_B]] = xN[u[i''_B]] \\
                   & \Rightarrow N[i'_B] = N[i''_B] \\
                   & \text{since } xN[u[i'_B]] = N[i'_B] \\
                   & \text{and } xN[u[i''_B]] = N[i''_B] \\
                   & \Rightarrow i'_B = i''_B \text{ up to negligible probability} \nonumber \\
                   & \text{by eliminating collisions.}
\end{align*}
\]
In order to prove $\psi \Rightarrow \phi$, two main steps:

1. **Collect the facts** that hold when the events in $\psi$ are executed.
2. Reason on these facts using an *equational prover* in order to show that the events in $\phi$ have been executed (and show injectivity when needed).

We shall now detail these points.
Collecting true facts

For each program point $P$, we collect a set of true facts at that point $\mathcal{F}_P$.

- We take into account assignments and tests above $P$.

**Example**

In `if M then P`, $M \in \mathcal{F}_P$.

- We take into account facts that hold at all definitions of variables.

**Example**

If `defined(x[\tilde{M}]) \in \mathcal{F}_P` and $M$ holds at all definitions of $x[\tilde{i}]$, then $M\{\tilde{M}/i\} \in \mathcal{F}_P$.

- We take into account that code is always executed up to the next output before switching to another thread.
Collecting true facts: example (1)

\[ O_{\text{start}}() := rk_A \overset{R}{\leftarrow} \text{keyseed}; \text{let } pk_A = pkgen'(rk_A) \text{ in return}(pk_A); \]

\[
\text{foreach } i_A \leq n \text{ do } O_A(xN : \text{nonce}, xB : \text{host}) := \]

\[
\text{event } e_A(pk_A, xB, xN); \text{let } m = \text{concat}(pk_A, xB, xN) \text{ in } \]

\[
\text{r } \overset{R}{\leftarrow} \text{seed}; \text{return}(\text{sign}'(m, skgen'(rk_A), r)) \]

\[
| \text{foreach } i_B \leq n \text{ do } O_{B1}(xpk_A : \text{pkey}) := N \overset{R}{\leftarrow} \text{nonce}; \text{return}(N, B); \]

\[
O_{B2}(s : \text{signature}) := \text{find } u \leq n \text{ such that } \]

\[
\text{defined}(m[u], xB[u], xN[u]) \land (xpk_A = pk_A) \land (B = xB[u]) \land (N = xN[u]) \land \text{verify}'(\text{concat}(xpk_A, B, N), xpk_A, s) \text{ then } \]

\[
\text{event } e_B(xpk_A, B, N) \]

At program point \( P = \text{event } e_B(xpk_A, B, N) \),

\[
\mathcal{F}_P = \{ \text{defined}(m[u[i_B]]), \text{defined}(xB[u[i_B]]), \text{defined}(xN[u[i_B]]), \]

\[
xpk_A[i_B] = pk_A, B = xB[u[i_B]], N[i_B] = xN[u[i_B]], \ldots \]
Collecting true facts: example (2)

\[ O_{\text{start}}() := rk_A \xleftarrow{\text{R}} \text{keyseed}; \text{let } pk_A = pkgen'(rk_A) \text{ in return}(pk_A); \]
\[ \text{foreach } i_A \leq n \text{ do } O_A(xN : \text{nonce}, xB : \text{host}) := \]
\[ \text{event } e_A(pk_A, xB, xN); \text{let } m = \text{concat}(pk_A, xB, xN) \text{ in} \]
\[ r \xleftarrow{\text{R}} \text{seed}; \text{return}(\text{sign'}(m, \text{skgen'}(rk_A), r)) \]

\[ | \ldots \]

At program point \( P = \text{event } e_B(xpk_A, B, N), \)

\[ \mathcal{F}_P = \{ \text{defined}(m[u[i_B]]), \text{defined}(xB[u[i_B]]), \text{defined}(xN[u[i_B]]), \]
\[ xpk_A[i_B] = pk_A, B = xB[u[i_B]], N[i_B] = xN[u[i_B]], \]
\[ \text{event}(e_A(pk_A, xB[u[i_B]], xN[u[i_B]])), \ldots \} \]

because \( \text{defined}(m[u[i_B]]) \in \mathcal{F}_P. \)
We use an algorithm inspired by the Knuth-Bendix completion algorithm to derive new equalities from known equalities.

The equational prover also eliminates collisions when they have negligible probability.

Details of this prover can be found in the paper: Blanchet, A Computationally Sound Mechanized Prover for Security Protocols, TDSC, 2008.

We say that $\mathcal{F}$ yields a contradiction when the prover starting from $\mathcal{F}$ derives false.
Proof of non-injective correspondences (1)

Let $\psi \Rightarrow \phi = F_1 \land \ldots \land F_m \Rightarrow \phi$ be a non-injective correspondence, with fresh variables.

Example

$\text{event}(e_B(x, y, z)) \Rightarrow \text{event}(e_A(x, y, z))$.

If $F_1, \ldots, F_m$ have been executed, then there exist $P_1, \ldots, P_m$ such that, for all $j \leq m$,

- $F_j = \text{event}(e_j(M_{j1}, \ldots, M_{jm_j}))$,
- $\text{event } e_j(M'_{j1}, \ldots, M'_{jm_j}); P_j$ occurs in $Q_0$, and
- $\text{event } e_j(M'_{j1}, \ldots, M'_{jm_j})$ has been executed with $F_j = \theta'_j \text{event}(e_j(M'_{j1}, \ldots, M'_{jm_j}))$, where $\theta'_j$ renames the replication indices at $P_j$ to fresh replication indices.

Example

$\text{event } e_B(xpk_A, B, N)$ has been executed, with $\text{event}(e_B(x, y, z)) = \text{event}(e_B(xpk_A[i'_B], B, N[i'_B])), \theta' = \{i'_B/i_B\}$.
Then the facts $\mathcal{F}_j = \theta'_j \mathcal{F}_{P_j} \cup \{\theta'_j M'_j = M_{j1}, \ldots, \theta'_j M'_{jm_j} = M_{jm_j}\}$ hold.

Example

$\mathcal{F}_P \{i'_B/i_B\} \cup \{x = xpk_A[i'_B], y = B, z = N[i'_B]\}$ hold, where $\mathcal{F}_P$ has been described previously.

For each such $P_1, \ldots, P_m$, we show that

$\mathcal{F} = \mathcal{F}_1 \cup \ldots \cup \mathcal{F}_m$ implies $\theta \phi$

for some $\theta$ equal to the identity on $\text{var}(\psi)$, by the equational prover.

(For this proof, we prove atomic facts contained in $\theta \phi$, we choose $\theta$ by matching facts to prove with elements of $\mathcal{F}$, and we show that $\mathcal{F}$ implies $M$ by showing that $\mathcal{F} \cup \{\neg M\}$ yields a contradiction.)

Example

$x = xpk_A[i'_B], y = B, z = N[i'_B], xpk_A[i'_B] = pk_A, B = xB[u[i'_B]],$

$N[i'_B] = xN[u[i'_B]], \text{event}(e_A(pk_A, xB[u[i'_B]], xN[u[i'_B]]))$

imply $\text{event}(e_A(x, y, z))$. 